Studies on the Bacterial Potassium Efflux System Kef in *Shewanella denitrificans Sd*Kef

Samuel Charles Grayer

The Queen's College, University of Oxford

DPhil Dissertation

Organic Chemistry

(DPhil Systems Approaches to Biomedical Science IDC)

Submitted Trinity 2015

Supervisors: Professor Stuart Conway and Associate Professor Garrett Morris

Studies on the Bacterial Potassium Efflux System Kef in *Shewanella denitrificans Sd*Kef

Samuel Charles Grayer

The Queen's College

DPhil Systems Approaches to Biomedical Science IDC – Organic Chemistry

Submitted Trinity 2015

The glutathione-gated potassium efflux system (Kef) is a K^+/H^+ antiporter found in the majority of Gram-negative pathogens, which shows potential as a novel antibiotic target. Kef plays a vital role in the protection of bacteria against toxic electrophiles through the regulation of cytoplasmic pH. Kef is inhibited by glutathione (**GSH**), ɣ-L-glu-L-cys-gly, and activated by glutathione S-conjugates (**GSX**). Healy *et al.* have quantified the affinities of **GSH** and a range of **GSX** for *Shewanella denitrificans* Kef (*Sd*Kef). **GSH** was found to have a weak affinity of 900 µM, whereas the strongest **GSX**, **^{t}BuSG**, was 400 nM.

This dissertation looks to understand the potency shown by **^{t}BuSG** for *Sd*Kef by exploring the binding contributions from each group of the tripeptide. Truncated analogues of **^{t}BuSG** were synthesised and evaluated using a competition fluorescence assay and ^{1}H CPMG NMR. In summary, removal of the glycine unit caused a complete loss in affinity for *Sd*Kef, whereas removing the glutamate unit resulted in a negligible loss. Interactions made by the Glu-Cys amide carbonyl oxygen, the Cys-Gly amide N*H* and a directional interaction of the conjugated thiol were also found to be important contributors to affinity.

The information obtained during this work allowed the development of a membrane permeant, truncated analogue of **^{t}BuSG**, which lacks the majority of the glutamate, for use in *in vivo* studies. The truncated analogue is able to activate *Sd*Kef to elicit K^+ efflux, demonstrating that the majority of the glutamate is not essential for activity. Furthermore, application of this truncated analogue to *Escherichia coli* cells expressing the *sdkef* gene in a Kirby-Bauer disc diffusion assay has demonstrated for the first time that small molecules activating *Sd*Kef can elicit inhibition of growth / cell death. Kef thus shows promise as a target for the development of novel antibacterial agents.

Acknowledgements

I would like to start by thanking my supervisors Professor Stuart Conway and Associate Professor Garrett Morris. Stuart – thank you for introducing me to the world of research and the encouragement you provided from my first day as a Part II student through to the last day of my PhD. Garrett – thank you for your continued support during my time at InhibOx and throughout the rest of my PhD. I'm really pleased about your academic appointment you deserve it! I would also like to thank Dr. Paul Finn for helping with the computational aspects of my project at InhibOx and Dr. David Staunton for his continued support in the Molecular Biophysics Suite. Also to our collaborators in the Miller Group and the Naismith Group, thank you for your input and support.

As life is short I shall move on to Lingbing so he doesn't have to read everything. My PhD-long affair with the semi-prep would not have been as smooth sailing without you. You have been a legend and I maintain that Lingbing is King. I will stop getting emotional now, as we all know what 1 mg of a hormone can do! Jess thanks for getting the ball rolling with the project and passing on your expertise. Anthony, when Bonnie Tyler was holding out for a hero 'til the end of the night I imagine the street-wise Hercules who turned up was much like you. I cannot express enough how much of a difference your arrival made to my PhD. You may not be my wildest fantasy but you were "for damn sure!" the superman that swept me off my feet. Corinna you almost managed to drill organisation into me, although I never quite got the hang of which side of the lab was blue and which was red. Sonia I really appreciate your help with STD NMR in my time of need. Jaideep – I can't think of a more fitting head Postdoc. Happy Birthday Julia!

Amjad – in the words of Shakespeare: "some are born great, some achieve greatness, and some have greatness thrust upon them." When I went down to the NMR room one cold winter day in my second year I had no idea greatness was to be thrust upon me! You were a great addition to team Kef and I come away from my and Anthony's time in the basement with you with 700 precious memories.

Dave and Tim thank you for always taking time out to help me with chemistry problems, even when you were busy writing your theses! Di you looked after me from our early days in team hypoxia through to the final year of your PhD. Alas, the circle of life has turned another cycle and Saunders is all that remains of the trio. Although the speed he's sprinting to the finish line, he'll probably be submitted soon, preferably without a hole in his bucket! Laura you were a great person to troubleshoot with when things were not as straightforward as anticipated. Brian the antiques roadshow has been great fun, although I really hope that there will be no more antiques from you in the future. Michelle – you go girl! Was totes amaze experiencing this with you! Michael I always looked forward to when you went home because it meant that there would be lots of Swiss chocolate in the lab the following week! Pour mon amie Amelie. Thank you for tolerating me calling you Am due to my incapability to handle multisyllabic words first thing in the morning (and the rest of the day). It was a pleasure to have you as a neighbour! Sarah and Charles you were a much-needed addition to the Conway group and it's a shame you didn't arrive sooner! Long live trips to the Old Bookbinders! Liam I'm not sure the acknowledgement is an appropriate setting for what there is to be said. All I will say is that I didn't expect to meet such a good friend when I started. To the Part IIs, Masters and summer students in no particular order: Tim, Andy, Amy, Calvin, Katie, Maya, Abbie, Jay, Sophie, Candise, Alexander and Eva. You all contributed to the experience in your own special ways. Of note: Ich liebe Alexander.

Finally thank you to my family for their support over the last four years. Susie this experience has made me realise how lucky I am. Whenever I was at my grumpiest you always managed to turn that frown upside down ☹ → ☺.

Oh yeah and I think there was someone called Angie in the group, but she changed her name at some point. Hopefully she's doing OK.

Table of Contents - Chapter 1

Table of Contents – Chapter 2

Table of Contents – Chapter 3

Table of Contents - Chapter 4

Table of Contents – Chapter 5

Table of Contents – Chapter 6

Table of Contents – Chapter 7

List of Abbreviations

°C	Degrees Celsius
Å	Angstrom
Au	Absorbance units
*Ab*Kef	*Acinetobacter baumanii* Kef
Ac	Acetyl
ADP	Adenosine diphosphate
ADT	AutoDockTools
AMP	Adenosine monophosphate
aq.	Aqueous
Ar	Aromatic
ATP	Adenosine triphosphate
Boc	*tert*-Butyloxycarbonyl
c	Concentration
CAP	Catabolite activator protein
cal	Calories
CCG	Chemical Computing Group
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CLIP-HSQC	Clean inphase-HSQC
conc.	Concentrated
COSY	Correlation spectroscopy
CPMG	Carr-Purcell-Meiboom-Gill
CPU	Central processing unit
CSR	Chiral Shape Recognition
Da	Daltons
DCC	*N,N'*-Dicyclohexylcarbodiimide
DIBAL	Diisobutylaluminium hydride
DIPEA	*N,N*-Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNGSH	*S*-((5-(Dimethylamino)naphthalen-1-yl)sulfonylaminopropyl) L-glutathione

DSF	Differential scanning fluorimetry
e.e.	Enantiomeric excess
*Ec*Kef	*Escherichia coli* Kef
*Ec*KefCCTD	*Escherichia coli* Kef C-terminal domain
EDC.HCl	*N*-(3-Dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
eq.	Equivalents
ES	Electrospray
ESG	*N*-Ethylsucciniimido glutathione
ESI	Electrospray ionisation
ESP	Electrostatic potential surface
g	Grams
GE	Group efficiency
GSH	Glutathione
GSNO	*S*-Nitrosoglutathione reductase
GSNOR	*S*-Nitrosoglutathione
GSX	Glutathione S-conjugates
h	Hours
HBTU	*O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
Hm1SdKefCTD	Homology model version 1 of *Sd*KefCTD using 3L9W as template
Hm2SdKefCTD	Homology model version 2 of *Sd*KefCTD using 3L9X as template
HMBC	Heteronuclear multiple bond correlation
HPLC	High-performance liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence
Hz	Hertz
IPA	2-Propanol
IR	Infra red spectroscopy
ITC	Isothermal titration calorimetry
K	Kelvin
Kef	K^+ efflux systems
KTN	K^+ transport and nucleotide-binding

LB	Luria broth
LC	Liquid chromatography
m/*z*	Mass divided by charge ratio
Me	Methyl
min	Minutes
MOE	Molecular operating environment
Mol	Moles
mp	Melting point
MST	Microscale Thermophoresis
NAD^+	Oxidised nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NEM	*N*-Ethylmaleimide
NHA	Number of heavy atoms
NHS/EDC	*N*-Hydroxysuccinimide / ethyl(dimethylaminopropyl) carbodiimide
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
OD	Optical density
ppm	Parts per million
PDB	Protein data bank
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
R_f	Retention factor
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
RP	Reverse phase
rpm	Rotations per minute
RT	Room temperature
RU	Response units
s	Seconds
SAR	Structure activity relationship
*Sd*Kef	*Shewanella denitrificans* Kef
*Sd*KefCTD	*Shewanella denitrificans* Kef C-terminal domain
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
S_NAr	Nucleophilic aromatic substitution

SPR	Surface Plasmon Resonance
STD NMR	Saturation transfer difference NMR
^{t}Bu	*tert*-Butyl
^{t}BuSG	*S*-(*N*-*tert*-Butylsuccinimido) L-glutathione
TCEP.HCl	Tris(2-carboxyethyl)phosphine hydrochloride
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin-layer chromatography
Tm	Inflection point of the transition curve
TMS	Tetramethylsilane
Trt	Trityl
USR	Ultra-fast shape recognition
UV/Vis	Ultraviolet/visible
WT	Wild-type

Amino Acid Codes

G	glycine	Gly
A	alanine	Ala
L	leucine	Leu
M	methionine	Met
F	phenylalanine	Phe
W	tryptophan	Trp
K	lysine	Lys
Q	glutamine	Gln
E	glutamic acid	Glu
S	serine	Ser
P	proline	Pro
V	valine	Val
I	isoleucine	Ile
C	cysteine	Cys
Y	tyrosine	Tyr
H	histidine	His
R	arginine	Arg
N	asparagine	Asn
D	aspartic acid	Asp
T	threonine	Thr

Chapter 1:

Introduction

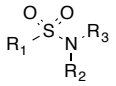
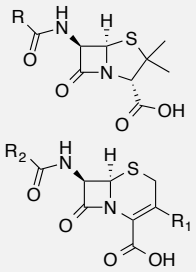
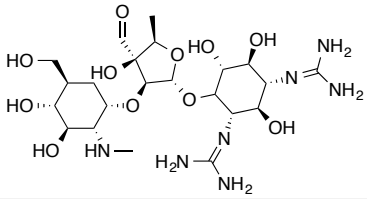
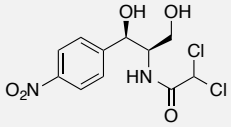
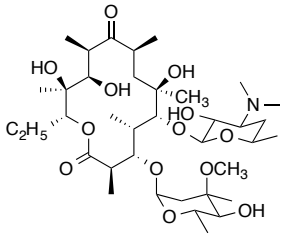
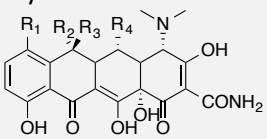
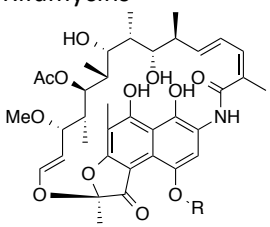
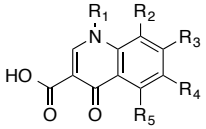
1. Introduction

Antibiotic resistance is an escalating problem, and so it is essential that antibiotics demonstrating novel mechanisms of action be developed. Glutathione-gated K^+ efflux systems (Kef), which are common to most Gram-negative pathogens and have no known homologues in humans act as feasible targets for the development of such agents. This introduction shall provide a very brief overview of antibiotic discovery, including past and current discovery platforms, before discussing Kef in more detail.

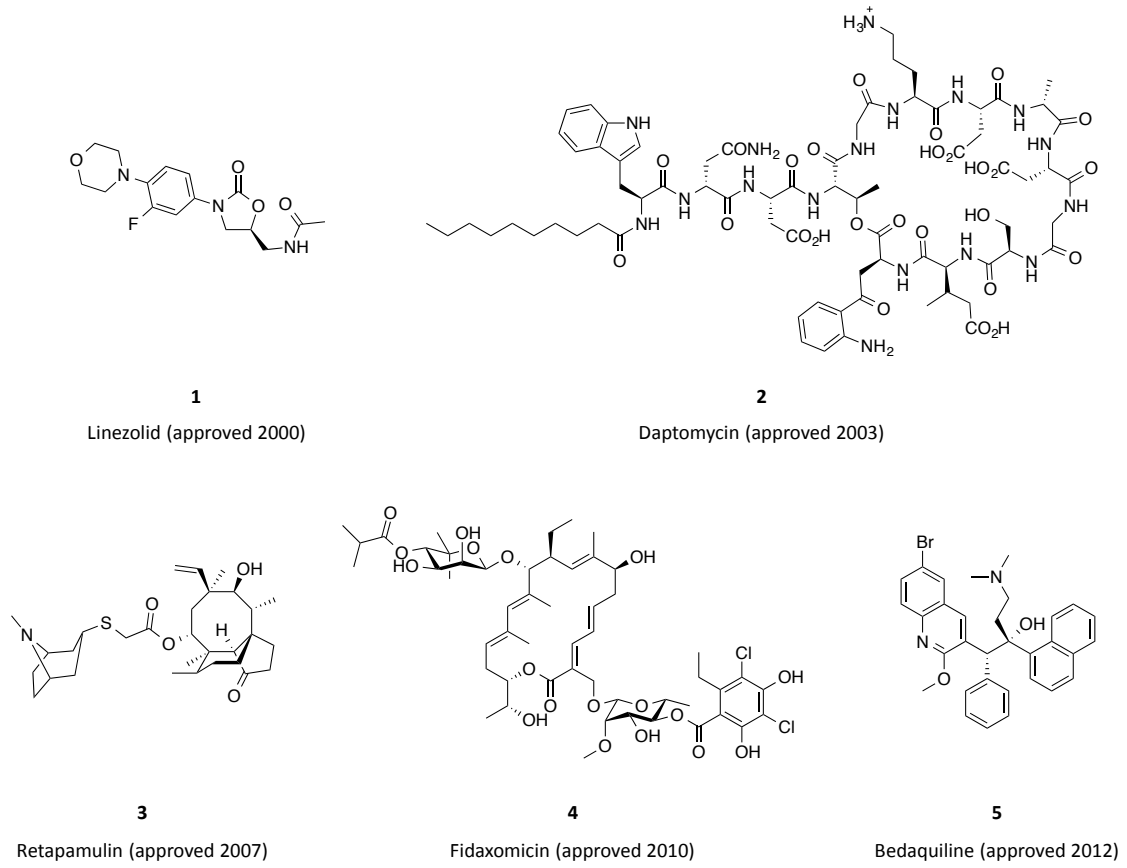
1.1 The golden era of antibiotic discovery and the rise in antibacterial resistance

The discovery of the sulfonamides§ and β-lactams in the 1930s saw the birth of an antibiotic revolution. The 1930s-1960s have been heralded as the golden era of this revolution, with half of the antibiotics that were in use in 2006 discovered between 1950-1960 (Table 1.1).[1] Unfortunately, this success was not sustained during the 1970s-1990s, during which time only analogues of existing antibiotics were released and mupirocin, a topical antibiotic for Gram-positive bacteria was discovered.[2]

Table 1.1 Examples of antibiotic classes discovered in the 1930s-1960s.[3]

Class	Introduced	Mechanism of action
Sulfonamides 	1936	Prevent growth and multiplication by inhibiting nucleic acid synthesis
β-lactams 	1938	Inhibit bacteria cell wall biosynthesis
Aminoglycosides 	1946	Inhibit the synthesis of proteins by binding to the 30S ribosomal subunit to prevent translation
Chloramphenicol 	1948	Inhibit the synthesis of proteins by binding to the 50S ribosomal subunit to prevent translation
Macrolides 	1951	Inhibit the synthesis of proteins by binding to the 50S ribosomal subunit to prevent translation
Tetracyclines 	1952	Inhibit the synthesis of proteins by binding to the 30S ribosomal subunit to prevent translation
Rifamycins 	1958	Binds to the RNA polymerase β-subunit
Glycopeptides	1958	Inhibit bacteria cell wall biosynthesis
Quinolones 	1968	Interfere with bacteria DNA replication and transcription

Fortunately, the decline in novel antibiotic innovation witnessed during the 1970s-1990s has started to reverse, with the release of five new structural classes since 2000: linezolid, **1**; daptomycin, **2**; retapamulin, **3**; fidaxomicin, **4**; and bedaquiline, **5**; (Figure 1.1).[2] These new classes, however, only combat Gram-positive infections and so significant challenges remain to counteract the intensifying problem of antibiotic resistance, especially in Gram-negative bacteria.[2]



1
Linezolid (approved 2000)

2
Daptomycin (approved 2003)

3
Retapamulin (approved 2007)

4
Fidaxomicin (approved 2010)

5
Bedaquiline (approved 2012)

Figure 1.1 The structures of linezolid, **1**, daptomycin, **2**, retapamulin, **3**, fidaxomicin, **4**, and bedaquiline, **5**.[2]

The dilemma posed by antibiotic resistance was highlighted by a 2014 World Health Organisation survey.[4] This survey collated data on the resistance of the most globally concerning strains of bacteria to antibiotics commonly used to treat them (Figure 1.2). In summary: *Escherichia coli* showed high levels of resistance to third-generation cephalosporins, **6**, and fluoroquinolones, **7**; *Klebsiella pneumonia* were found to be resistant to third-generation cephalosporins, **6**, and carbapenems, **8**; *Streptococcus pneumonia* demonstrated reduced

susceptibility to penicillin, **9**; *Neisseria gonorrhea* were less susceptible to third-generation cephalosporins, **6**; *Staphylococcus aureus* were reported to show high levels of resistance to methicillin, **10**.[4] Given the extent of escalation in antibiotic resistance, it is essential that antibacterial agents exhibiting novel modes of action are developed. This is imperative not only to treat discrete bacterial infections but also to safeguard modern medical advancements, such as major surgery, organ transplantation and cancer chemotherapy, all of which rely on access to effective antibacterial drugs.[5]

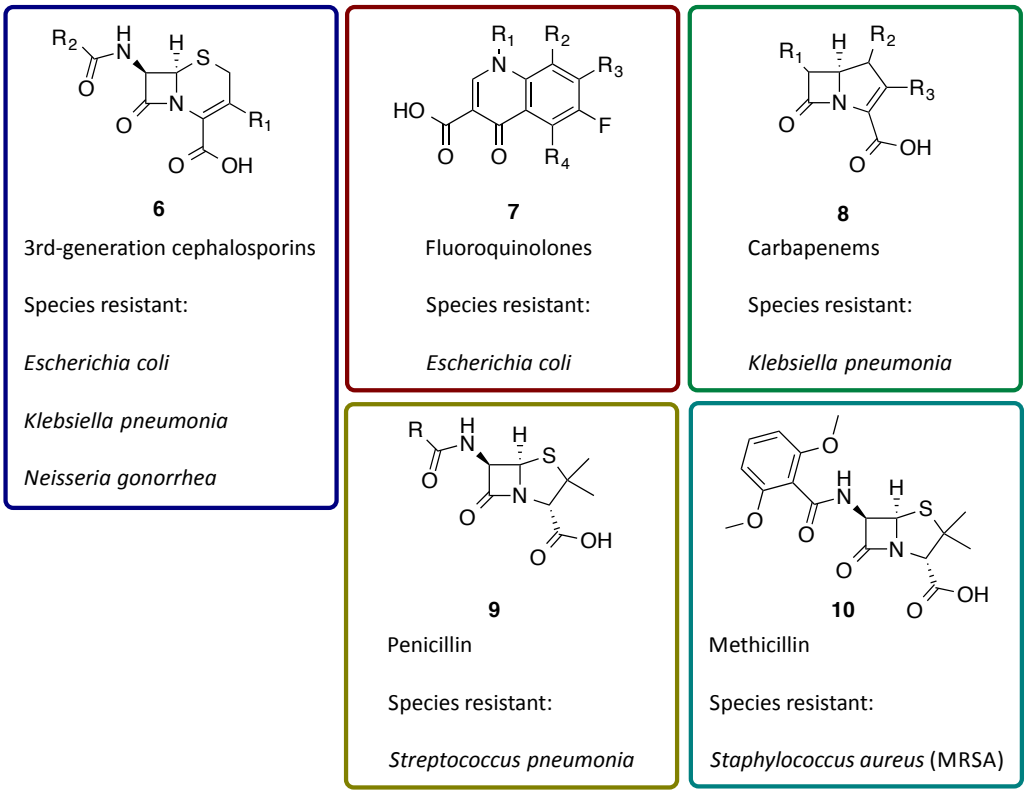


Figure 1.2 The general structures of antibiotics and the species that have shown resistance to them.

The current pipeline for new antibiotics does not, however, correspond with the growing concern shown for antibiotic resistance. A review commissioned by the UK Prime Minister into antimicrobial resistance analysed the antibiotic pipeline published by The Pew Charitable Trusts' in December 2014.[6] The review deduced that the pipeline was inadequate to meet current and future clinical requirements. Only 16 of the 41 antibiotics in development had significant efficacy

against multi-drug resistant Gram-negative species, and only 3 showed the potential to treat bacterial strains currently considered to the be most resistant. Furthermore, most of the candidates were still 10-15 years away from market.[6]

A number of factors have contributed to the decline in antibiotic discovery, including: complacency; lack of incentive to invest; competition from generic medicines; increased regulation; and fewer effective discovery platforms. Discussing each of these factors in detail would lead to a review in itself, and so the remaining focus will be on effective discovery platforms.

1.2 Discovery platforms

The success of the golden era of antibiotic discovery was largely attributed to the Nobel Prize-winning discovery platform developed by Selman Waksman.[7] This platform allowed the identification of natural products with antimicrobial activity through the screening of soil-derived *streptomycetes* against bacteria on an overlay plate and detecting zones of growth inhibition. The approach led to the discovery of streptomycin, an aminoglycoside that was the first compound to show efficacy against tuberculosis. The platform was thus adopted by the wider pharmaceutical industry and applied to other soil-derived *actinomycetes*. However, after 20 years of success, the viability of this approach diminished due to a plateau in the discovery of novel, efficacious compounds.[3] With the 'low-hanging fruit' offered by the 'Waksman platform' harvested during the golden era, scientists have since explored alternative platforms for developing novel antibacterial agents. A review by O'Connell *et al.* highlights some of the promising approaches taken.[8]

Although the number of natural products discovered *via* the 'Waksman platform' had been saturated, the majority had arisen from screening soil-derived *actinomycetes*. It is possible that the next generation of natural products could come from the systematic evaluation of different ecological niches of microbial flora. This approach might require screening bacteria that are unculturable under standard laboratory conditions, a problem that could be addressed by: developing nontraditional cultivation technologies; expressing the genes responsible for antibacterial biosynthesis in a culturable host; circumventing laboratory cultivation by evaluating the environmental DNA of microorganisms for the production of novel antibacterial agents.[8]

Other interesting proposals for the development of novel antibiotics include Lantibiotics, antimicrobial peptides and hybrid antimicrobials.[8] Lantibiotics are peptides with multiple macrocyclic rings that arise from numerous posttranslational modifications. They are ultimately toxins that are produced by bacteria to inhibit the growth of similar bacteria. Antimicrobial peptides comprise around 50 amino acids or fewer. A number of these residues are lysine or arginine and more than half are hydrophobic residues. The lysine and arginine residues impart an overall positive charge, which is fundamental to the proposed mechanism of action: initial targeting of the anionic phosphate head groups of bacterial phospholipid membranes. Hybrid antimicrobials, as their name suggests, comprise antimicrobial agents with distinct modes of action that have been fused together. The resulting agent should therefore act on two or more targets, with the overall aim of hindering the development of bacterial resistance.[8]

The final approach outlined by O'Connell *et al.* involves targeting an intracellular signaling phenomenon called quorum sensing.[8] Quorum sensing allows bacteria to communicate *via* signaling molecules called autoinducers. Autoinducers bind to receptor proteins to coordinate the activation or repression of gene expression. Pathogenic bacteria, such as *Pseudomonas*

aeruginosa, use quorum sensing to coordinate assaults on their hosts through the regulation of virulence factor production. Inhibition of quorum sensing is thus a promising approach to developing novel antibacterial agents. The reader is directed to two additional reviews on quorum sensing for a more detailed insight into the phenomenon.[9,10]

In summary, promising strides are being made towards the development of novel antibiotics. Three key discovery platforms have been highlighted: identifying novel natural products that elicit antibacterial behavior; modifying and exploiting known natural products that elicit antibacterial behavior; and probing a specific target, such as quorum sensing, to develop agents with novel mechanisms of action. The latter approach will be implemented in this dissertation through the targeting of glutathione-gated K^+ efflux systems (Kef).

1.3 The role and regulation of potassium ions (K^+) in bacteria

Potassium ions (K^+) are known to have four key functions in bacteria: 1) acting as an osmotic solute to maintain turgor pressure; 2) activating intracellular enzymes; 3) regulating intracellular pH; 4) acting as a secondary messenger.[11] It is therefore essential that the intracellular concentration of K^+ be carefully regulated to ensure that K^+ is able to fulfill each of these roles. The efflux of K^+ in bacteria is controlled by Kef, while K^+ influx is controlled by uptake systems, such as: Trk (**Tr**ansport **K^+**), Kdp (**K^+ de**p**e**ndent) and Kup (**K^+ up**take).[11] It should be noted that the mechanosensitive channels of bacteria can also provide a nonspecific path for K^+ efflux by opening upon osmotic downshock.[11] With the key roles played by K^+ and its regulation in bacteria outlined, it is possible to perceive how targeting Kef could inhibit the growth of, or kill, bacteria. As such, the potential consequences of disrupting the regulation of intracellular pH by permanently inhibiting or activating Kef shall be explored in more detail later.

1.4 Glutathione-gated K^+ efflux systems (Kef) as novel antibiotic targets

Ferguson *et al.* first proposed Kef as promising targets for novel antibacterial agents in 1993.[12] Kef are found in the majority of gram-negative bacteria,[13] and have no known homologues in humans, presenting the opportunity to develop a universal agent that can treat a range of pathogenic infections. Furthermore, a homology search performed by Dr. Anthony Chan (unpublished data) has established that homologues of Kef are present in a number of the ESKAPE pathogens (***E****nterococcus faecium*, ***S****taphylococcus aureus*, ***K****lebsiella pneumoniae*, ***A****cinetobacter baumanii*, ***P****seudomonas aeruginosa*, and ***E****nterobacter* species). These pathogens have been highlighted as problematic because they cause the majority of US hospital infections reported to date (Table 1.2).[14]

Table 1.2 Homology search of Kef homologues performed using NCBI protein blast (blastp).[15] The amino acid primary sequence of *Shewanella denitrificans* Kef (GenBank ID: ABE53663.1) was used as a query sequence for the blastp search. For algorithm selection, DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)[16] was chosen for the blastp search against available genomes of ESKAPE pathogens [i.e. *Enterococcus faecium* (taxid:1352), *Staphylococcus aureus* (taxid:1280), *Klebsiella pneumoniae* ATCC BAA-2146 (taxid:1263871), *Acinetobacter baumanii* BIDMC 56 (taxid:1439316), *Pseudomonas aeruginosa* PAO1 (taxid:208964), Enterobacteria (taxid:91347)]. The Kef homologues from human pathogens have been selected and tabulated (Dr. Anthony Chan, unpublished data).

GeneBank ID	Bacterial species	Cover (%)	Identity (%)	Positives (%)	Gaps (%)
AHI34262.1	*Klebsiella pneumoniae* ATCC® BAA-2146™	98	43	62	1
KDF03402.1	*Acinetobacter baumanii* BIDMC 56	98	39	59	2
NP_249898.1	*Pseudomonas aeruginosa* POA1	97	40	58	2
Q8XA20.1	*Escherichia coli* O157:H7	98	44	62	1
YP_006780926.1	*Escherichia coli* O104:H4	98	44	62	1
Q3Z5W2.1	*Shigella sonnei* Ss046	98	44	62	1
WP_003829616.1	*Citrobacter freundii*	98	44	62	1
B4TJ42.1	*Salmonella enterica*	98	44	61	1
WP_032715999.1	*Enterobacter aerogenes*	98	44	62	1
WP_035896028.1	*Kluyvera ascorbata*	98	44	62	1
A7MIA0.1	*Cronobacter sakazakii* ATCC BAA-894	98	45	62	1

1.5 The structure of Kef

The Kef systems from *E. coli* KefB (*Ec*KefB) and KefC (*Ec*KefC) are the most extensively studied, and so form the basis for this section. When cloning and sequencing the *kefC* gene of *Ec*KefC, Munro *et al.* identified a single open reading frame that encodes for a protein with 620 amino acids.[17] The first 180 amino acids of the N-terminus were found to be highly hydrophobic, suggesting that this region corresponds to the transmembrane channel, which is predicted to have 12 transmembrane spans.[11] The remaining residues 181-620 of the C-terminus were found to be much more hydrophilic, with the exception of the regions 181-260 and 320-340, which are very hydrophobic.[17]

The significance of the C-terminus' high hydrophilic residue content is that it encompasses a cytoplasmic region, where ligand-gating of *Ec*KefC occurs. This cytoplasmic gating region corresponds to a common protein fold called the K^+ transport and nucleotide-binding (KTN) domain. KTN domains are ubiquitous to a number of prokaryotic and eukaryotic K^+ channels and contain a structural motif called a Rossmann fold to which nucleotides bind.[18]

KTN domains consist of dimers that can combine to form higher assembly oligomers (tetramers and octamers), through hydrophobic patches.[18,19] Each of the subunits that form these dimers contains a Rossmann fold of a six-stranded parallel β-sheet with three α-helices on either side of it. A tight association is formed between the two subunits due to their C-terminal α-6 helices crossing-over with one another (Figure 1.3).[18] The Rossmann folds are connected by loops joining the C-terminal β-6 strands to the exchanged α-6 helices to form a hinge region.[18] This hinge region is conformationally flexible and the KTN domain is proposed to undergo conformation changes upon ligand binding.[18] The KTN and transmembrane domains in *Ec*KefC are in close proximity, connected *via* a flexible 19-residue linker that potentially provides a

physical means of communication between the two domains. Conformational changes at the KTN domain upon ligand association could thus be translated to pore opening or closing at the transmembrane domain.[18] The conformational changes induced by ligand binding to the *Ec*KefC KTN domain will be explored in more detail later.

In addition to the N-terminal transmembrane and C-terminal KTN domains in *Ec*KefC, there is an ancillary subunit called KefF (KefG for *Ec*KefB), which is required for maximum activity of *Ec*KefC.[20] Roosild *et al.* have published a 2.4 Å resolution X-ray crystal structure of the KTN-bearing C-terminal domain of *Ec*KefC (*Ec*KefCCTD) in complex with a tethered KefF subunit (PDB code: 3EYW; Figure 1.3).[21]

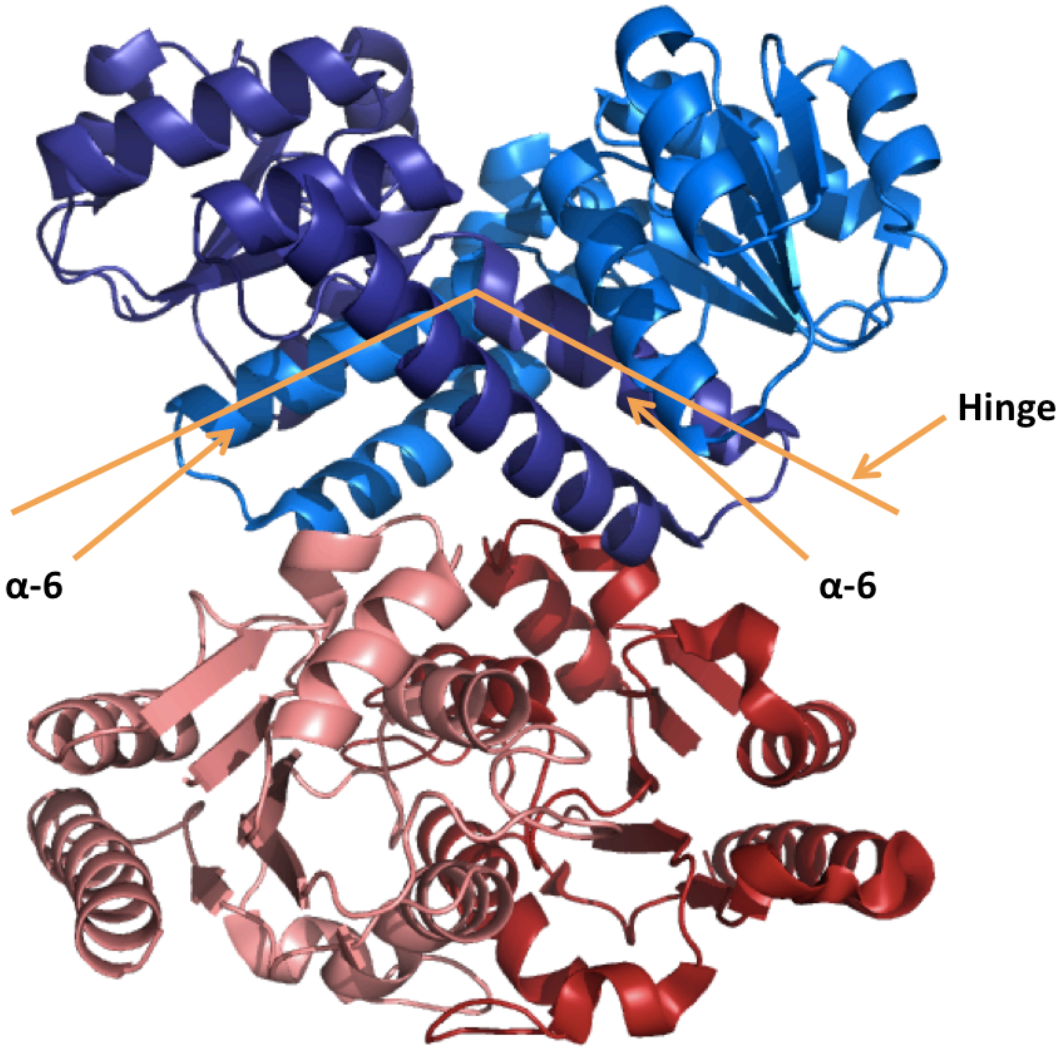
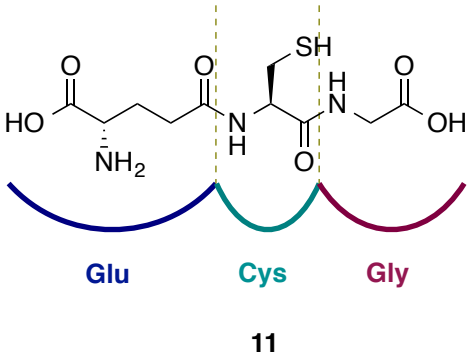


Figure 1.3 The 2.4 Å resolution X-ray crystal structure of the KTN-bearing C-terminal domain of *Ec*KefC (*Ec*KefCCTD) (shown as cartoon; blue) in complex with KefF (shown as cartoon; red) (PDB code: 3EYW).[21]

As can be observed from the X-ray crystal structure (PDB code: 3EYW) in Figure 1.3, the KTN domain *Ec*KefCCTD and its ancillary subunit KefF each form dimers. The overall *Ec*KefCCTD-KefF complex is consequently a dimer of dimers. The KefF dimer binds at the hinge interface of the two subunits of *Ec*KefCCTD, holding it at an angle of approximately 120°.[21] KefF has been proposed to help to stabilise a conductive state of the channel by interacting with *Ec*KefCCTD to form this angle.[21]

1.6 The role of glutathione and Kef in the detoxification of electrophiles

One of the ligands that gates Kef is glutathione (**GSH**; **11**) a ubiquitous tripeptide comprising ɣ-L-Glu-L-Cys-Gly (Figure 1.4), which exists at high cytosolic concentrations in both eukaryotic and prokaryotic cells (≈ 10-20 mM). The defining feature of this tripeptide is the presence of a ɣ-linked peptide bond between the residues Glu and Cys. This ɣ-linked peptide is thought to protect glutathione from hydrolysis by intracellular peptidases.[22]



11

Figure 1.4 The structure of glutathione.

Glutathione has a number of functions in cells, including: REDOX buffer; coenzyme; co-substrate; cysteine store; antioxidant activity; and detoxification of electrophiles.[22] The role of glutathione and Kef in the detoxification of electrophiles in bacteria will be explored here.

So far Kef have been described as regulators of K^+ efflux. More specifically, they are K^+/H^+ antiporters that use the energy associated with the movement of K^+ out of the cell, down their concentration gradient, to transport H^+ into the cell, against their concentration gradient. Evidence for this functionality comes from the treatment of *E. coli* KefB and KefC with the electrophile *N*-ethylmaleimide, **12**, (**NEM**; Figure 1.7) which was found to induce a rapid but reversible loss of K^+, along with an exchange of K^+ for H^+.[23-25] The overall outcome of this antiport activity is acidification of the intracellular environment. Bacteria exploit the drop in

cytoplasmic pH associated with the activation of Kef to provide protection against toxic electrophiles, such as metabolic methylglyoxal, **13**, (Figure 1.5).[12,24,26-29]

Lowering the intracellular pH is thought to result in protonation of nucleophilic sites that are susceptible to electrophilic assault, such as amino groups in proteins and DNA. Protonation of these sites affords a protective shield that reduces their reactivity towards electrophilic species.[29,30] Activators or inhibitors of Kef could therefore be used to disrupt a vital safeguard against electrophilic assault. Furthermore, the regulation of cytoplasmic pH is essential for cell growth, and so interfering with this process through the activation or inhibition of Kef may also contribute to the inhibition of growth and ultimate death of bacteria.[31]

Glutathione regulates K^+ efflux by inhibiting Kef in the absence of electrophiles to prevent K^+ leakage.[24,32] In the presence of electrophiles, glutathione S-conjugates form that activate Kef, triggering K^+ efflux. Glutathione S-conjugates are formed when the free thiol on glutathione couples to an assaulting electrophile. The extent of activation elicited by a given glutathione S-conjugate is subject to the following degree of subtlety: when the conjugation adduct is large and hydrophobic, all Kef transporters are activated; when the adduct is small and hydrophilic, only a subset of Kef are activated.[26] For example, *S*-lactoylglutathione, **16**, the glutathione S-conjugate that arises from methylglyoxal, **13**, detoxification is a small adduct that strongly activates *Ec*KefB, but only weakly activates *Ec*KefC.[12]

Glutathione thus plays a dual role in the detoxification of electrophiles in bacteria: firstly it acts as an electrophile scavenger; secondly the glutathione S-conjugates arising from electrophile scavenging activate Kef to elicit K^+ efflux and concomitant acidification of the cytosol. This process is exemplified in Figure 1.5, which demonstrates the process of glutathione S-conjugate formation and detoxification for the electrophile methylglyoxal, **13**. Direct reaction of

glutathione with methylglyoxal, **13**, forms a hemithioacetal, **14**, which is isomerised to *S*-lactoylglutathione, **16**, by Glyoxalase I. Once the levels of intracellular methylglyoxal have lowered, Glyoxalase II hydrolyses *S*-lactoylglutathione, **16**, back to glutathione, forming D-lactate, **17**, in the process.[33]

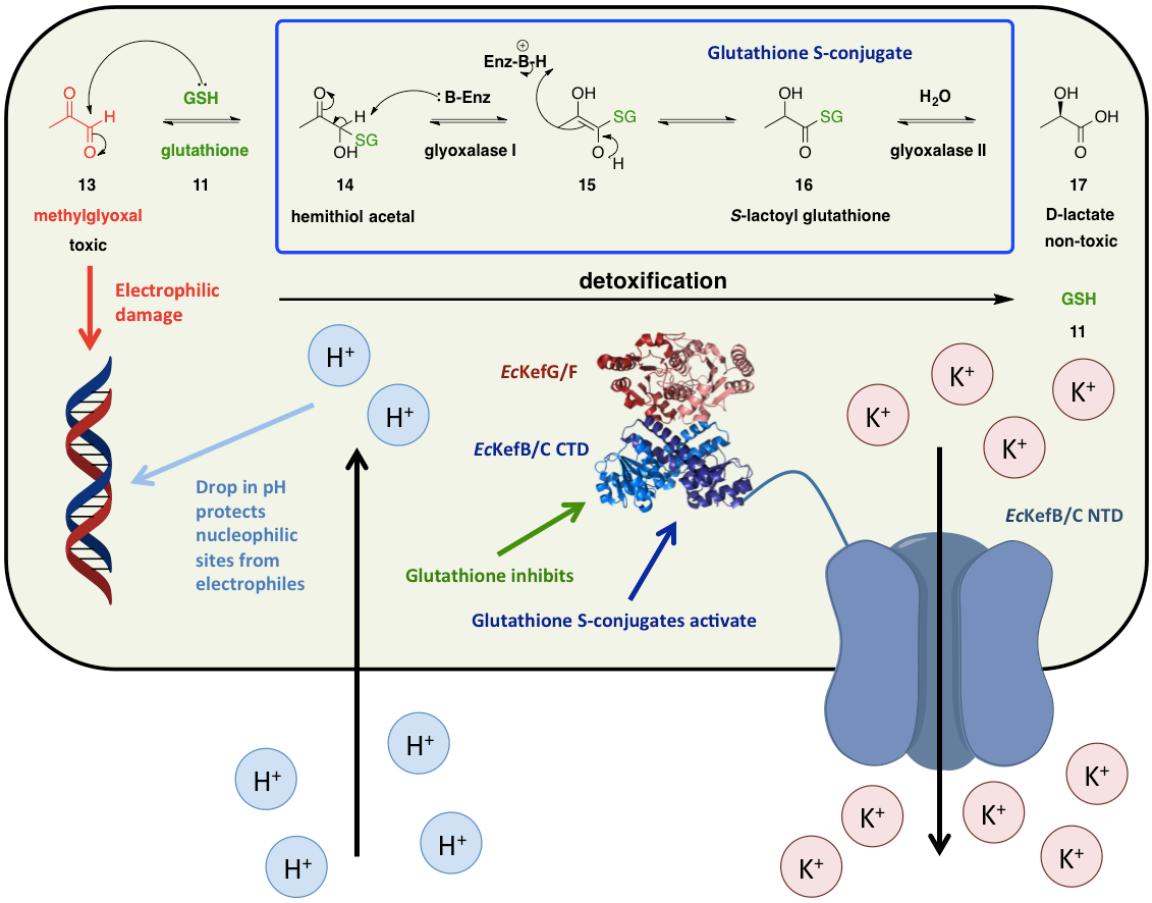


Figure 1.5 Schematic of the synergistic actions of the Kef system in *E. coli* in response to electrophilic assault. Glutathione (**GSH**) acts as an inhibitory ligand of Kef. In the presence of electrophiles, **GSH** forms glutathione S-conjugates with these species to activate the Kef system. Activation of Kef results in an efflux of K^+ and concomitant H^+ influx. The decrease in intracellular pH causes nucleophiles to become protonated, protecting them from electrophilic assault. The glutathione S-conjugates undergo enzyme-mediated metabolism to less toxic species.[34]

1.7 Identifying the glutathione-binding site

The mechanism by which glutathione and glutathione S-conjugates are able to inhibit or activate Kef, respectively, has been the subject of hot debate. Before forming an opinion on this, it is imperative to understand the glutathione-binding site. Mutational studies performed by Miller *et al.*[27] and Roosild *et al.*[21] on *Ec*KefC provided the first direct insight into the location of the glutathione-binding site (Figure 1.6).

The mutants R416S, R516C (identified by Miller *et al.*) and N551D (identified by Roosild *et al.*) were found to have reduced sensitivity to glutathione inhibition, suggesting that these mutations disrupt glutathione binding to *Ec*KefC. All three of the mutated residues are located on the cytosolic C-terminal KTN domain *Ec*KefCCTD. Subsequent mutational studies performed by Roosild *et al.* established that the residue Q412 was also critical to glutathione binding, with Q419K showing reduced sensitivity to inhibition by glutathione that was equivalent to introducing the triple mutation of R416A, R516A, and N551A.[31] Mapping the key binding residues identified above onto the X-ray crystal structure of *Ec*KefCCTD (PDB code: 3EYW)[21] finds that they are all located along the interface between the two subunits of the KTN domain dimer. Interestingly, R416 and Q412 are located on one of the subunits (Chain B; Figure 1.6), with R516 and N551 located on the other subunit (Chain A; Figure 1.6).

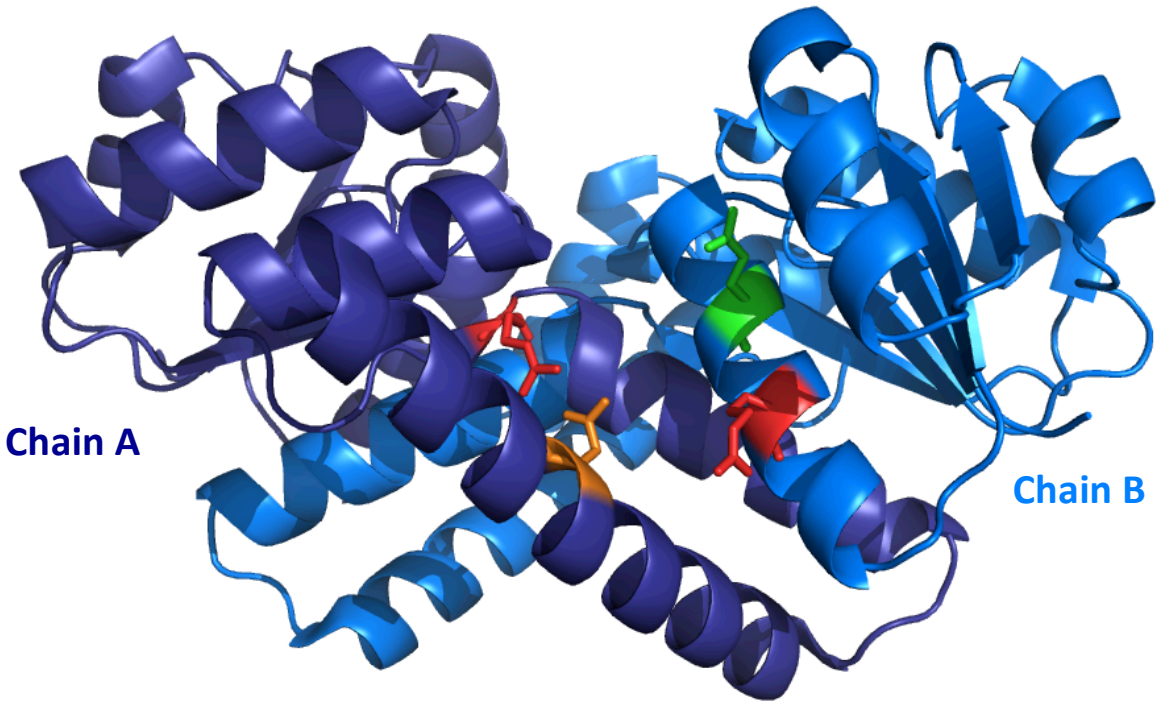


Figure 1.6 The 2.4 Å resolution X-ray crystal structure of *Ec*KefCCTD (shown as cartoon; blue; PDB code: 3EYW).[21] Residues identified to have reduced sensitivity to glutathione inhibition when mutated are highlighted as coloured sticks: R416 (red, Chain B); R516 (red, Chain A); N551 (orange, Chain A); Q412 (green, Chain B).

The positioning of these key-binding residues along this interface of the KTN domain suggests that the glutathione-binding site is located here, and supports the hypothesis of a ligand-mediated hinge mechanism of activation. Furthermore, as the KTN domain is a dimer, the opposite side of the protein has a similar arrangement of these residues, and so could act as a second location for glutathione binding. To provide structural evidence for these hypotheses, Roosild *et al.* set out to co-crystallise *Ec*KefCCTD with glutathione and the glutathione S-conjugate **ESG**, **18**.[31] **ESG** is formed when glutathione reacts *via* a conjugate addition at its thiol position with the electrophile *N*-ethylmaleimide (**NEM**), **12**, (Figure 1.7)

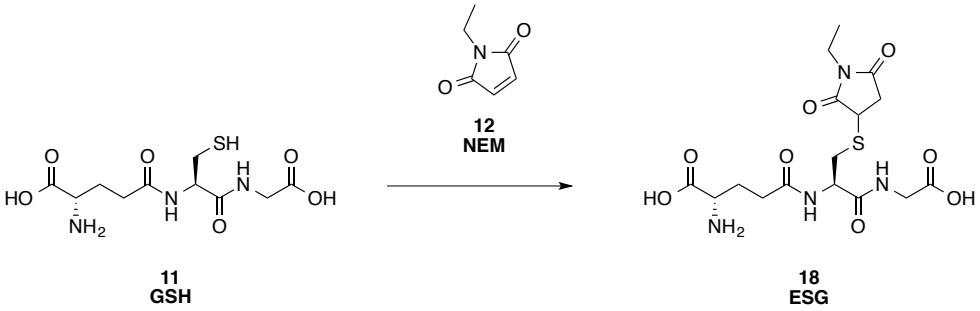


Figure 1.7 The structure of **ESG** and its formation from the reaction of glutathione (**GSH**) and **NEM**.

A 2.10 Å resolution X-ray crystal structure of *Ec*KefCCTD in complex with KefF and **ESG** (PDB code: 3L9X) was solved.[31] Obtaining the glutathione-bound structure, however, required the use of reducing agents to convert *in crystal* the co-crystallised **ESG** in 3L9X to glutathione. The result was a 1.75 Å resolution X-ray crystal structure of *Ec*KefCCTD in complex with KefF and glutathione (PDB code: 3L9W).[31] Both structures are globally similar in assembly to the ligand-free structure 3EYW[21] and had additional electron density in the proposed glutathione-binding site that was attributed to glutathione or **ESG**. The second glutathione-binding site on the opposite side of the dimer was unoccupied. The absence of a ligand in this location could, however, be an artifact of crystal packing due to the presence of crystal contacts near a mobile region of chain B in the model.

Analysis of the binding-site found that the binding modes of glutathione (Figure 1.8 A) and **ESG** (Figure 1.8 B) were very similar. Identical polar contacts were predicted to form between *Ec*KefCCTD and the backbones of both tripeptides, which were in fitting with the mutational studies: R416 and R516 co-ordinate directly to the Gly-carboxylate of the tripeptide backbone; N551 interacts indirectly with the Gly-carboxylate of the tripeptide backbone *via* a well-structured water molecule; and Q412 forms hydrogen-bonding interactions with the Cys-Gly amide N*H* and the Glu-Cys amide carbonyl oxygen of the tripeptide backbone. In

addition to these key binding residues, R498, D499 and the backbone carbonyl of V500 were also predicted to make key interactions with the glutathione and **ESG** tripeptide backbones.

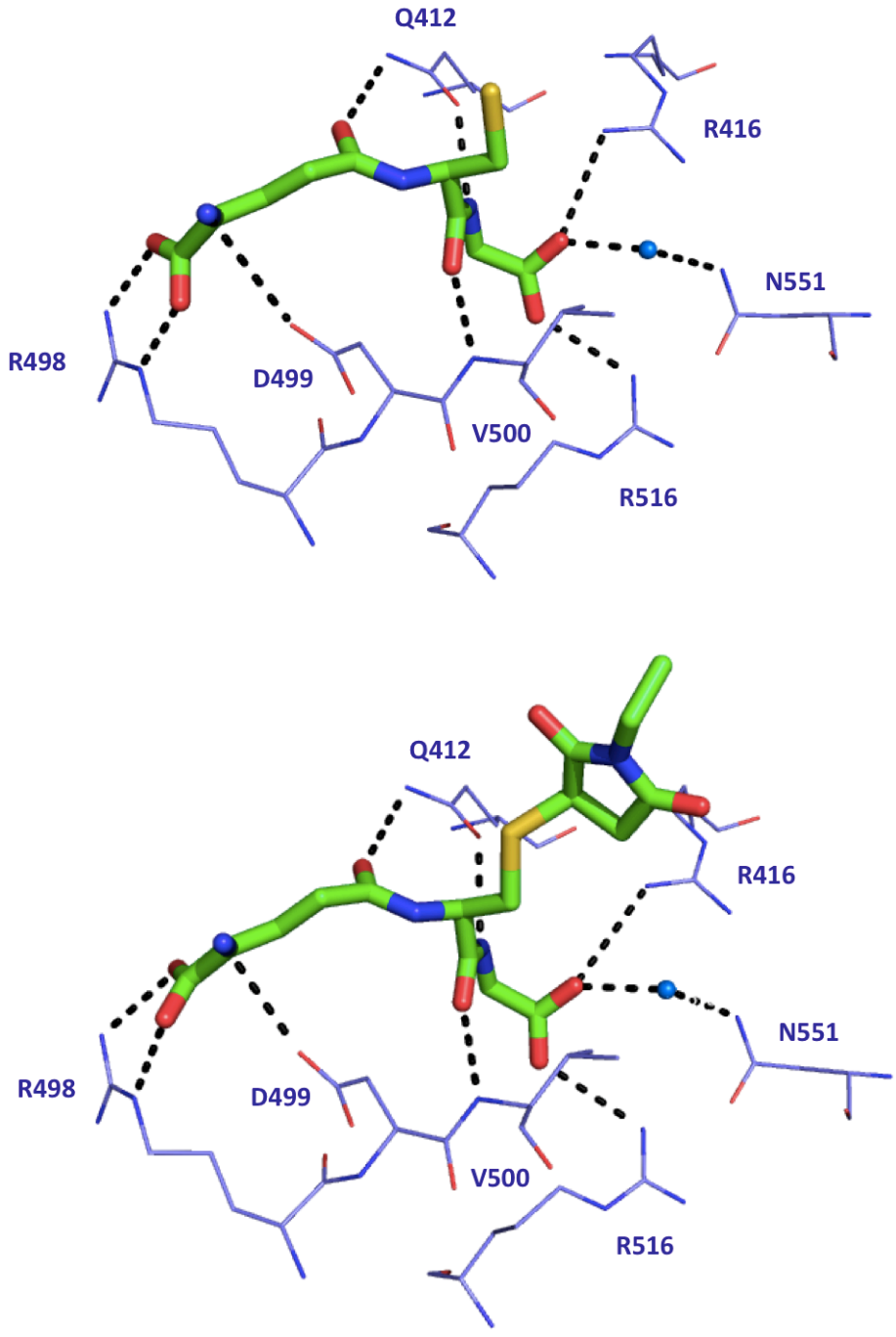


Figure 1.8 The glutathione-binding site of *Ec*KefCCTD. (A) Interactions between glutathione (shown as sticks; carbon = green) and key residues of *Ec*KefCCTD (PDB code: 3L9W) are highlighted by black dashed lines; (B) Interactions between the glutathione S-conjugate **ESG** (shown as sticks; carbon = green, one diastereomer shown) and key residues of *Ec*KefCCTD (PDB code: 3L9X) are highlighted by black dashed lines. Images generated in PyMOL.[35]

1.8 Proposed mechanism of action of Kef inhibition / activation

Given that the tripeptide backbones of glutathione and **ESG** were both making identical polar contacts with *Ec*KefCCTD, their difference in activity must arise from contrasting interactions made by the free thiol in glutathione and the succinimido ring on **ESG**. Comparison of the glutathione and **ESG** bound X-ray crystal structures 3L9W and 3L9X enabled Roosild *et al.* to elucidate the contrasting interactions made by the free thiol and the succinimido ring. A mechanism of action was thus proposed to account for the differing inhibitory and activatory effects that glutathione and glutathione S-conjugates respectively exert on Kef (Figure 1.9).[31]

The glutathione-bound structure 3L9W, indicated that there is a hydrophobic pocket formed by the α-7 and α-8 helices from one of the KTN subunits, which is occupied by the thiol group of glutathione and the residue F441 from the α-2 helix of the other subunit (Figure 1.9 A). Comparison of the ligand-free structure 3EYW[21] with 3L9W suggests that this conformation is stabilised by glutathione, as its thiol group is predicted to displace water that would otherwise be positioned near the phenyl ring of F441. Overlaying the **ESG** bound structure 3L9X with 3L9W suggests that the succinimido ring of **ESG** would sterically clash with the phenyl ring of F441 (Figure 1.9 B). As such, a different conformation of the KTN domain is observed in 3L9X, whereby F441 has been displaced from the hydrophobic pocket, and consequently the α-2 helix and other helices, such as α-8, have been repositioned (Figure 1.9 C). It was proposed that this conformational change in the KTN domain, when glutathione S-conjugates bind to displace F441, might induce pore opening at the N-terminal transmembrane domain to enable K^+ efflux (Figure 1.9).[31]

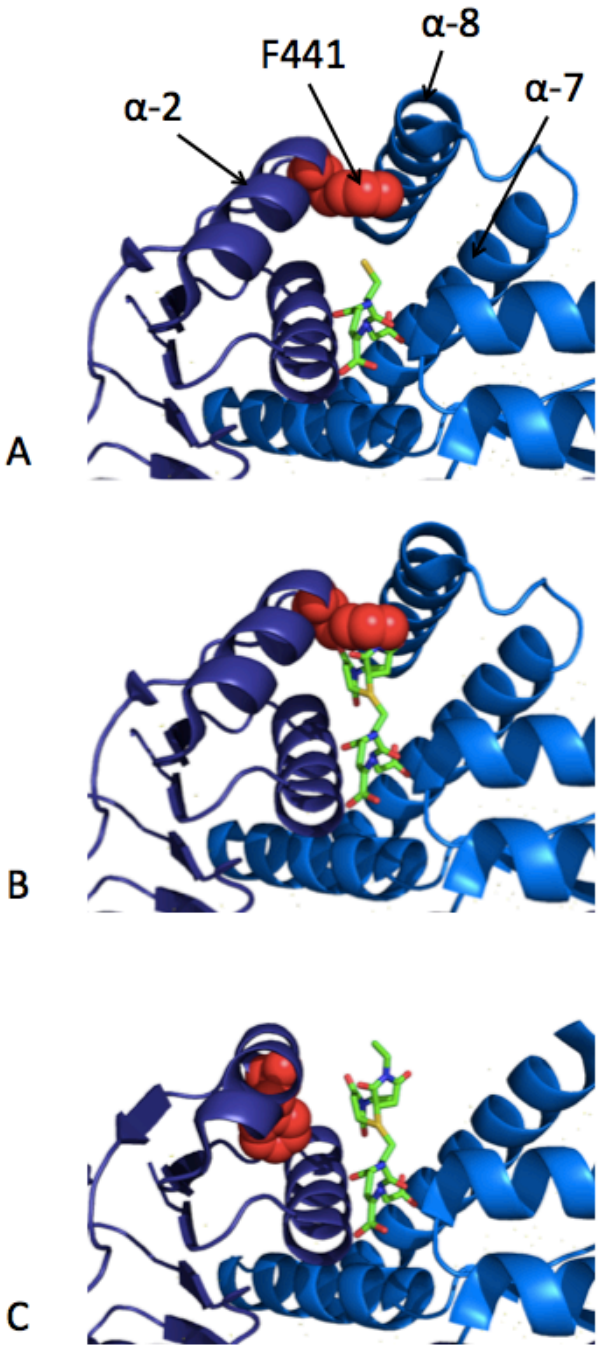


Figure 1.9 The mechanism of glutathione S-conjugate modulation of *Ec*KefCCTD conformation. (A) The reduced thiol of glutathione (shown as sticks; carbon = green) stabilises a binding interaction between F441 (shown as red spheres) on α-2 helix of one subunit with a hydrophobic pocket formed by α-7 and α-8 helices from the opposing subunit in 3L9W; (B) **ESG** (two diastereomers shown as sticks; carbon = green) overlaid with the conformation adopted in 3L9W to demonstrate the steric clash between F441 (shown as red spheres) and the succinimido ring of **ESG**; (C) The succinimido ring of **ESG** is proposed to displace F441 (shown as red spheres) from the hydrophobic pocket, yielding a different conformation of *Ec*KefCCTD in 3L9X. The disruption of this protein-protein interaction leads to the α-8 helix being repositioned and no longer resolved crystallographically.[31]

The proposed mechanism of action was subjected to mutational analysis, whereby the residue F441 was systematically mutated to bulkier residues (F441W and F441Y) and smaller, more flexible residues (F441L and F441D). These mutations were assessed for their ability to affect **ESG**-elicited K^+ efflux. The bulkier residues F441W and F441Y saw no significant loss in activity, whereas the activity was lower in the more flexible F441L mutant. Reducing the size of the mutant residue further to afford F441D caused the activity to decrease more.[31] There therefore appears to be a correlation between the size and rigidity of the amino acid at the 441-position and the ability of the corresponding *Ec*KefC mutant to be activated by **ESG**. These mutational studies are in good agreement with the proposed displacement of the residue F441, which acts as a single amino acid switch to induce conformational changes that trigger K^+ efflux.

It is worth noting here that the paper published by Roosild *et al.* has been corrected to note that the positioning of the succinimido ring of **ESG** in 3L9X is not clear from the observed electron density. Furthermore, there is additional electron density adjacent to the sulfur of glutathione in 3L9W that is sufficient to have a succinimido ring modeled into it. It is therefore possible that glutathione and **ESG** could have been modeled into the wrong X-ray crystal structures.[31]

In light of this, it is unlikely that the mechanism of action of Kef activation is as simplistic as a single amino acid switch at the F441 position. However, one cannot detract from the conformational changes observed regardless of the X-ray crystal structure to which glutathione and **ESG** were bound. The concept that glutathione and **ESG** are able to stabilise alternative conformations of the KTN domain, which in turn could stabilise open or closed forms of the N-terminal transmembrane channel, remains intact. The succinimido ring is clearly the important factor in implementing conformational changes and further work needs to be done to either further validate the mechanism proposed by *Roosild et al.* or provide an alternative model.

1.9 Structural requirements for activators of Kef

With the glutathione-binding site of *Ec*KefC established and an insight provided into the mechanism of action for Kef inhibition/activation, Healy *et al.* set out to determine the affinities of glutathione and glutathione S-conjugates for a homologue of the Kef KTN domain (Figure 1.10).[34] These studies were performed using a homologue of Kef from *Shewanella denitrificans* (*Sd*Kef), and binding was assessed both qualitatively and quantitatively using an isolated form of its C-terminal KTN domain (*Sd*KefCTD) and the following assays: ITC; competition fluorescence (described in Chapter 3; Section 3.2.1); and differential scanning fluorimetry (described in Chapter 2; Section 2.4).

The reason that the study was performed using *Sd*Kef was because this homologue does not require a KefF-like ancillary protein for full activity, simplifying purification, and the interpretation of biophysical assays. Healy *et al.* quantified the affinity of glutathione for *Sd*KefCTD to be 900 μM (fluorescence emission spectra) and **ESG** to be 12 μM (fluorescence emission spectra) and 23 μM (ITC) using a single-site binding model. These results are interesting, as they help explain how glutathione, which is abundant in cells at concentrations of ≈ 10-20 mM, is overridden by lower concentrations of glutathione S-conjugates to achieve rapid gating. The difference in affinity between glutathione and **ESG** is attributed to the succinimido group, which was speculated to do one or more of the following: form hydrophobic interactions with *Sd*KefCTD; cause a conformational change to *Sd*KefCTD that results in enthalpic gains; or cause a conformational change to *Sd*KefCTD that results in entropic gains due to the release of structured water molecules. In order to explore the role of the electrophilic component in glutathione S-conjugates further, Healy *et al.* synthesised a range of adducts and assessed their affinity for *Sd*KefCTD qualitatively and in some cases quantitatively (Figure 1.10).[34]

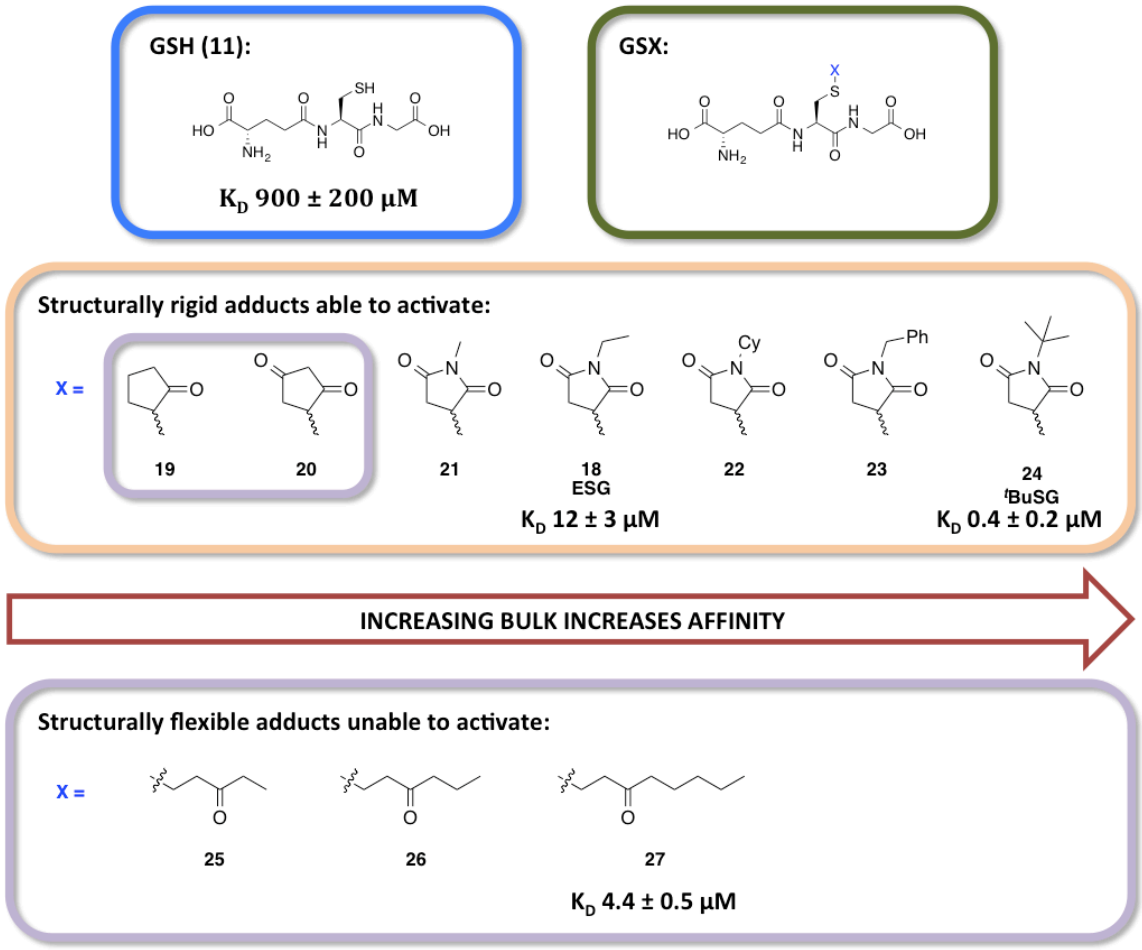


Figure 1.10 Summary of the findings of Healy *et al.* on the structural requirements for activators of the *Shewanella denitrificans* Kef system.[34]

Figure 1.10 summarises the trends extracted from these experiments. Increasing the steric bulk of the structurally rigid adducts was found to cause a consequent increase in affinity for SdKef. **^{t}BuSG**, **24**, an analogue of **ESG** which has a *tert*-butyl group in place of the ethyl group on the succinimido ring, was identified as the strongest binder of *Sd*KefCTD, with an affinity of 0.4 µM (fluorescence emission spectra). The same trend was observed for the structurally flexible adducts, with compound **27**, the glutathione S-conjugate of 1-octen-3-one, having a dissociation constant of 4.4 µM (fluorescence emission spectra).[34]

In order to evaluate whether each of the adducts was able to activate full-length *Sd*Kef, K^+ efflux studies were performed using their precursor electrophiles. The structurally rigid glutathione S-conjugates, **18** and **21**-**24**, were found to activate *Sd*Kef, whereas **19** and **20** were not. As both **19** and **20** lack the substituted nitrogen on the succinimido ring, it is likely that a requirement for activation is the presence of a bulky substituent at this position. None of the structurally flexible glutathione S-conjugates, **25**-**27**, were able to activate *Sd*Kef, despite showing affinities similar to the more structurally rigid compounds **18**-**24**. The conclusion drawn was that the structural properties of the electrophile are important in determining whether the glutathione S-conjugate formed is able to activate *Sd*Kef, with both sterically bulky and structurally rigid electrophiles required for efficient activation of *Sd*Kef. With the structure activity relationship (SAR) around the thiol position of glutathione S-conjugates explored, investigating the SAR of the tripeptide backbone is the next logical step to further the understanding of glutathione and glutathione S-conjugate regulation of Kef.

1.10 Summary

Antibiotic resistance is an escalating problem, with multidrug resistant pathogens becoming commonplace. It is, therefore, essential that antibiotics demonstrating novel mechanisms of action are developed to circumvent this resistance. K^+ efflux systems (Kef), which are common to most Gram-negative pathogens and have no known homologues in humans act as feasible targets for the development of such agents. Transport of K^+ is fundamental to the regulation of cell turgor and cytoplasmic pH. Kef provide bacteria with essential protection against toxic electrophiles, through the regulation of their cytoplasmic pH. Removal of this protective system renders bacteria unable to survive electrophilic stress. Inhibitors or activators of Kef could, therefore, eliminate a vital safeguard against toxic electrophiles.

The opportunity to disrupt Kef induced K^+ efflux is provided by a cytosolic C-terminal domain that contains a common protein fold called the K^+ transport and nucleotide-binding (KTN) domain. Kef KTN domains are regulated by glutathione, a ubiquitous tripeptide, that inhibits Kef, and glutathione S-conjugates, formed by glutathione reacting with electrophiles, that activate Kef. Published X-ray crystal structures of the KTN domain from *E. coli* KefC co-crystallised with either glutathione or the glutathione S-conjugate **ESG** have provided an insight into how each of these tripeptides can induce conformational changes in the KTN domain to inhibit or activate Kef. Assessment of a range of glutathione S-conjugates' affinity and ability to elicit K^+ efflux has found that glutathione S-conjugates need to be sterically bulky and structurally rigid in order to activate Kef.

1.11 Aim

The aim of this project is to develop membrane permeant chemical probes that can be used to provide a proof of concept of Kef as a novel target for antibiotic development. This aim will be approached in a number of ways, including using *in silico* screening to identify novel binders of the KTN domain of a homologue of Kef, and dissecting a potent glutathione S-conjugate, **^{t}BuSG**, **24**. These probes will be tested *in vitro* using a number of biophysical assays and *in vivo* using a K^+ efflux assay and a Kirby-Bauer disc diffusion assay.[36]

1.12 Chapter 1 References

[1] J. Davies, *Can J Infect Dis Med Microbiol* **2006**, *17*, 287–290.

[2] M. S. Butler, M. A. Blaskovich, M. A. Cooper, *The Journal of Antibiotics* **2013**, *66*, 571–591.

[3] K. Lewis, *Nat Rev Drug Discov* **2013**, *12*, 371–387.

[4] World Health Organization, *Antimicrobial Resistance: Global Report on Surveillance*, World Health Organization, **2014**.

[5] O. Cars, L. D. Högberg, M. Murray, O. Nordberg, S. Sivaraman, C. S. Lundborg, A. D. So, G. Tomson, *BMJ* **2008**, *337*, a1438.

[6] *Review on Antimicrobial Resistance Chaired by Jim O'Neill. Securing New Drugs for Future Generations: the Pipeline of Antibiotics*, **2015**.

[7] S. A. Waksman, *Science* **1953**, *118*, 259–266.

[8] K. M. G. O'Connell, J. T. Hodgkinson, H. F. Sore, M. Welch, G. P. C. Salmond, D. R. Spring, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 10706–10733.

[9] W. R. J. D. Galloway, J. T. Hodgkinson, S. Bowden, M. Welch, D. R. Spring, *Trends in Microbiology* **2012**, *20*, 449–458.

[10] W. R. J. D. Galloway, J. T. Hodgkinson, S. D. Bowden, M. Welch, D. R. Spring, *Chem. Rev.* **2011**, *111*, 28–67.

[11] W. Epstein, *Prog. Nucleic Acid Res. Mol. Biol.* **2003**, *75*, 293–320.

[12] G. P. Ferguson, A. W. Munro, R. M. Douglas, D. McLaggan, I. R. Booth, *Molecular Microbiology* **1993**, *9*, 1297–1303.

[13] G. Ferguson, *Trends in Microbiology* **1999**, *7*, 242–247.

[14] L. B. Rice, *J Infect Dis.* **2008**, *197*, 1079–1081.

[15] M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, T. L. Madden, *Nucl. Acids Res.* **2008**, *36*, W5–W9.

[16] G. M. Boratyn, A. A. Schäffer, R. Agarwala, S. F. Altschul, D. J. Lipman, T. L. Madden, *Biol. Direct* **2012**, *7*, 12.

[17] A. W. Munro, G. Y. Ritchie, A. J. Lamb, R. M. Douglas, I. R. Booth, *Molecular Microbiology* **1991**, *5*, 607–616.

[18] T. P. Roosild, S. Miller, I. R. Booth, S. Choe, *Cell* **2002**, *109*, 781–791.

[19] H. J. Kim, H. H. Lim, S. H. Rho, S. H. Eom, C.-S. Park, *J. Biol. Chem.* **2006**, *281*, 38573–38581.

[20] S. Miller, L. S. Ness, C. M. Wood, B. C. Fox, I. R. Booth, *J. Bacteriol.* **2000**, *182*, 6536–6540.

[21] T. P. Roosild, S. Castronovo, S. Miller, C. Li, T. Rasmussen, W. Bartlett, B. Gunasekera, S. Choe, I. R. Booth, *Structure* **2009**, *17*, 893–903.

[22] A. Krezel, W. Bal, *Org. Biomol. Chem.* **2003**, *1*, 3885–3890.

[23] E. P. Bakker, W. E. Mangerich, *FEBS Lett.* **1982**, *140*, 177–180.

[24] J. Meury, S. Lebail, A. Kepes, *Eur. J. Biochem.* **1980**, *113*, 33–38.

[25] J. Meury, A. Kepes, *EMBO J.* **1982**, *1*, 339–343.

[26] M. J. Elmore, A. J. Lamb, G. Y. Ritchie, R. M. Douglas, A. Munro, A. Gajewska, I. R. Booth, *Molecular Microbiology* **1990**, *4*, 405–412.

[27] S. Miller, R. M. Douglas, P. Carter, I. R. Booth, *J. Biol. Chem.* **1997**, *272*, 24942–24947.

[28] L. S. Ness, I. R. Booth, *J. Biol. Chem.* **1999**, *274*, 9524–9530.

[29] G. P. Ferguson, D. McLaggan, I. R. Booth, *Molecular Microbiology* **1995**, *17*, 1025–1033.

[30] G. P. Ferguson, Y. Nikolaev, D. McLaggan, M. Maclean, I. R. Booth, *J. Bacteriol.* **1997**, *179*, 1007–1012.

[31] T. P. Roosild, S. Castronovo, J. Healy, S. Miller, C. Pliotas, T. Rasmussen, W. Bartlett, S. J. Conway, I. R. Booth, *PNAS* **2010**, *107*, 19784–19789.

[32] M. J. Elmore, A. J. Lamb, G. Y. Ritchie, R. M. Douglas, A. Munro, A. Gajewska, I. R. Booth, *Molecular Microbiology* **1990**, *4*, 405–412.
[33] P. J. Thornalley, *Biochem. J.* **1990**, *269*, 1–11.
[34] J. Healy, S. Ekkerman, C. Pliotas, M. Richard, W. Bartlett, S. C. Grayer, G. M. Morris, S. Miller, I. R. Booth, S. J. Conway, et al., *Biochemistry* **2014**, *53*, 1982–1992.
[35] The PyMOL Molecular Graphics System, Version 1.7.4 Schrodinger, LLC, **2010**.
[36] A. W. Bauer, W. M. Kirby, J. C. Sherris, M. Turck, *Am. J. Clin. Pathol.* **1966**, *45*, 493–496.

Chapter 2:

Expression and purification of the C-terminal domain of *Shewanella denitrificans* Kef (*Sd*KefCTD) and evaluation of the nucleotide natively bound to it

2. Expression and purification of the C-terminal domain of *Shewanella denitrificans* Kef (*Sd*KefCTD) and evaluation of the nucleotide natively bound to it

2.1 Introduction and aims

The overall aim of my DPhil research was to develop probes that could be used to validate Kef as a novel antibiotic target. In order to achieve this goal, biophysical assays were required that would allow ligands of Kef to be ranked by affinity. Healy *et al.* made the first strides towards such an assay with the development of a competition fluorescence assay (described in Chapter 3).[1] Healy *et al.* commenced by selecting a suitable domain and homologue of Kef for use in the assay. This selection process required a detailed understanding of the molecular biology of Kef.

Mutational and X-ray crystallographic studies have established that glutathione and glutathione S-conjugates exert their effects of inhibition and activation, respectively, through binding to a C-terminal domain of Kef.[1-4] This domain is known as a K^+ transport and nucleotide-binding domain (KTN), which binds a nucleotide molecule in a structural motif called a Rossmann fold.[3,4] Full-length Kef proteins are homodimeric, and the binding site of glutathione and glutathione S-conjugates is at the interface of the two C-terminus KTN domains, which are located in a non-membrane region of the channel.[1] Isolated C-terminal KTN domains of Kef were therefore selected as a suitable model protein for performing biophysical studies.

It was decided that the previously characterised KefC homologue of Kef found in *Escherichia coli* (*Ec*KefC) was not suitable for biophysical studies due to it requiring an ancillary protein, KefF (*Ec*KefF) for full activity.[3-5] *Ec*KefC and *Ec*KefF are both dimers that have a weak affinity for one another and the instability of the resulting *Ec*KefFC dimer of dimers was predicted to interfere

with biochemical studies.[3] Therefore, a homologue of Kef was selected that was not dependent on a KefF-like protein. Healy *et al.* identified the homologue of Kef from *Shewanella denitrificans* (*Sd*Kef), which has no KefF-like protein, and mutational analysis found that the key binding residues of it are similar to *Ec*KefC.[1]

Having selected *Sd*Kef as a suitable homologue of Kef to use in biophysical assays, Healy *et al.* set about developing a soluble protein construct of the C-terminal KTN domain of *Sd*Kef (*Sd*KefCTD). A plasmid construct encoding the gene for *Sd*KefCTD was obtained using the expression construct for the full-length *Sd*Kef channel, pTrcSdKefH$_6$ (Figure 2.1).

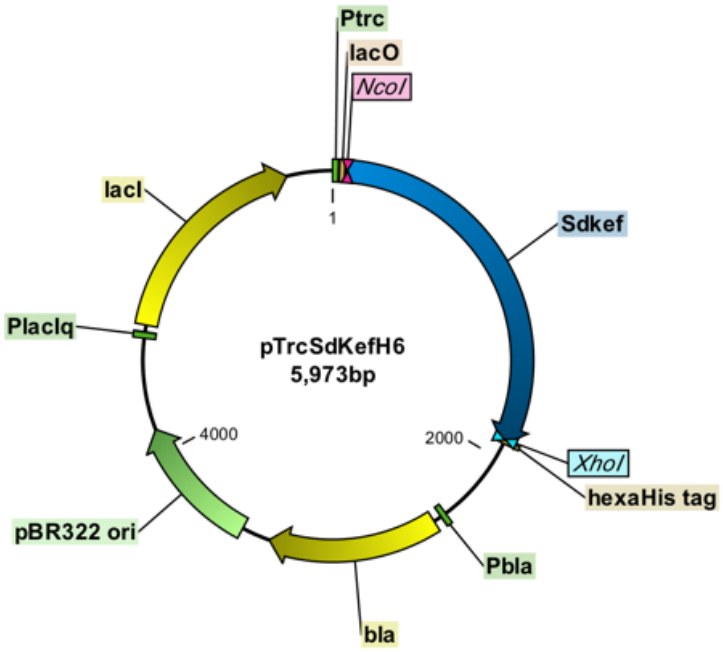


Figure 2.1 Plasmid map of pTrcSdKefH$_6$ plasmid, where *Ptrc* is the promoter for the gene *sdkef*; *lacO* is the lac operator; *Nco*I is the *Nco*I restriction site; *Sdkef* is the region containing the gene *sdkef*; *Xho*I is the *Xho*I restriction site; hexaHis tag is the HIS$_6$ tag for *Sd*Kef; *Pbla* is the promoter for the gene *bla*; *bla* is the gene for β-lactamase for ampicillin resistance; pBR322 origin is the origin of replication that ensures the pTrcSdKefH$_6$ is replicated by the bacteria; *PlacIq* is the promoter for the gene *lacI*; *lacI* is the gene for the LacI repressor, which repressed the transcription of the *sdkef* gene by binding to *lacO*. Image constructed using CLC Main Workbench 7.6.2.

pTrcSdKefH_6 was constructed from a pTrc99A plasmid backbone, which has the gene encoding for the full-length *Sd*Kef channel (GenBank ID: ABE53663.1) and a C-terminal LEH_6 tag cloned into it *via* the *Nco*I and *Xho*I restriction sites. The plasmid construct encoding the gene for *Sd*KefCTD was obtained by using site-directed mutagenesis to introduce an *Nco*I site into pTrcSdKefH_6 at the 3' end of the membrane domain-encoding region. Digestion of this *Nco*I site, followed by ligation yielded the desired plasmid construct designated pTrcSdKefQCTD (Figure 2.2).

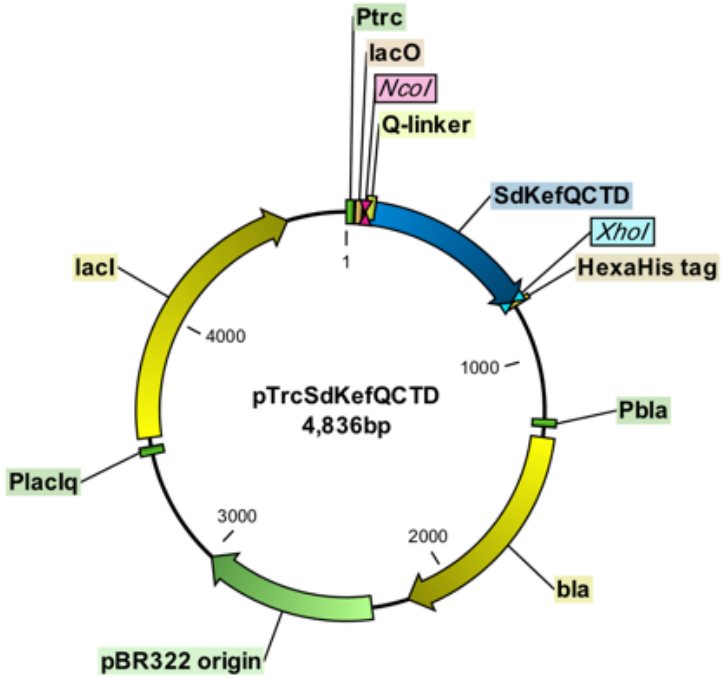


Figure 2.2 Plasmid map of pTrcSdKefQCTD plasmid, where *Ptrc* is the promoter for the gene *sdkef*; *lacO* is the lac operator; *Nco*I is the *Nco*I restriction site; *SdkefQCTD* is the region containing the gene *sdkefQCTD*; *Xho*I is the *Xho*I restriction site; hexaHis tag is the HIS_6 tag for *Sd*KefQCTD; *Pbla* is the promoter for the gene *bla*; *bla* is the gene for β-lactamase for ampicillin resistance; pBR322 origin is the origin of replication that ensures the pTrcSdKefQCTD is replicated by the bacteria; *PlacIq* is the promoter for the gene *lacI*; *lacI* is the gene for the LacI repressor, which repressed the transcription of the *sdkefQCTD* gene by binding to *lacO*. Image constructed using CLC Main Workbench 7.6.2.

Expression of pTrcSdKefQCTD afforded the 237 amino acid *Sd*KefCTD protein, which contains the C-terminal KTN domain of *Sd*Kef, as well as, the Q-linker (GHELEVDIEP). The Q-linker is a sequence comprising the regulatory loop from the membrane domain (residues H266-P274). The regulatory loop from the membrane domain was included as it had previously been found to improve stability and solubility.[3] Healy *et al.* went on to show that purified *Sd*KefCTD exists in solution as dimers by static multiangle light scattering analysis, which detected 52.0 ± 2.5 kDa particles (26.4 kDa per monomer).[1]

With a soluble protein construct of *Sd*KefCTD, in hand, it was possible to start developing biophysical assays to identify novel ligands of Kef (as discussed in Chapters 3 and 4). It was also possible to commence crystallisation trials to obtain an X-ray crystal structure of *Sd*KefCTD for use in *in silico* protein-ligand docking studies (Chapter 3). Our collaborator Dr. Christos Pliotas in the Naismith group, University of St Andrews, obtained an X-ray crystal structure of *Sd*KefCTD with a resolution of 2.9 Å (Figure 2.3; Pliotas and Naismith unpublished data). This structure was found to consist of two monomers that form an archetypal KTN dimer,[6-8] consistent with the multiangle light scattering observations of Healy *et al.*[1]

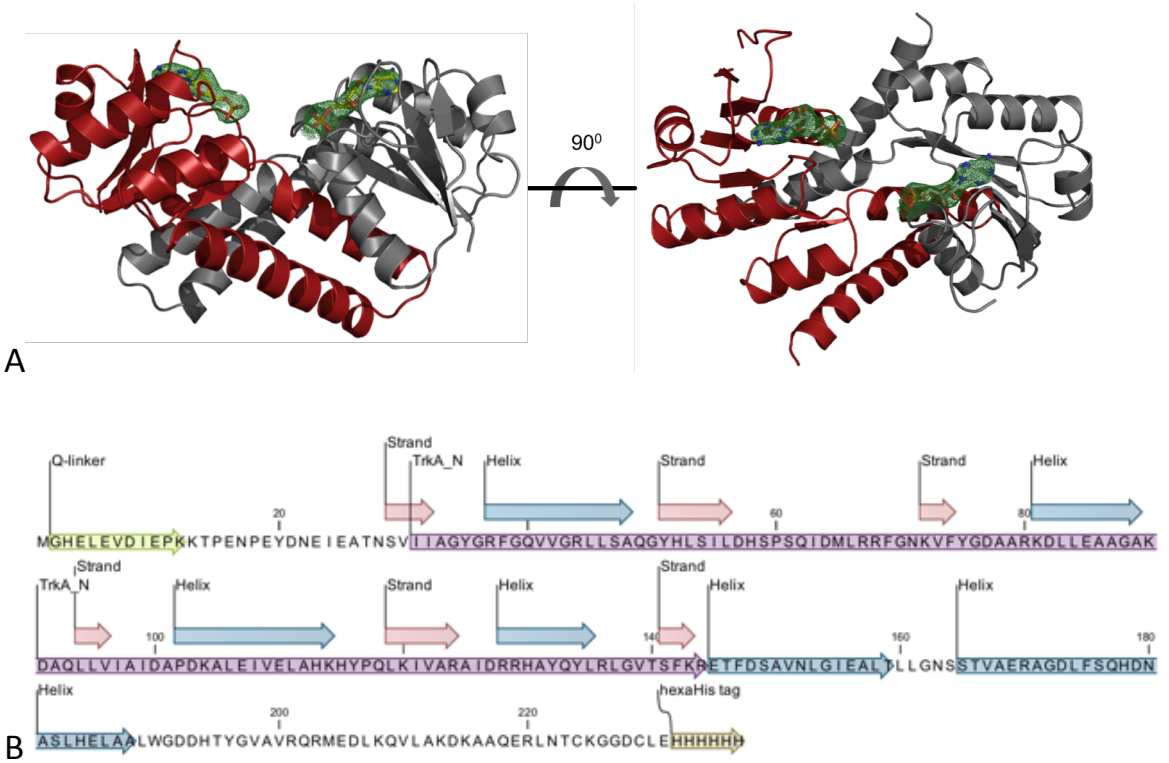
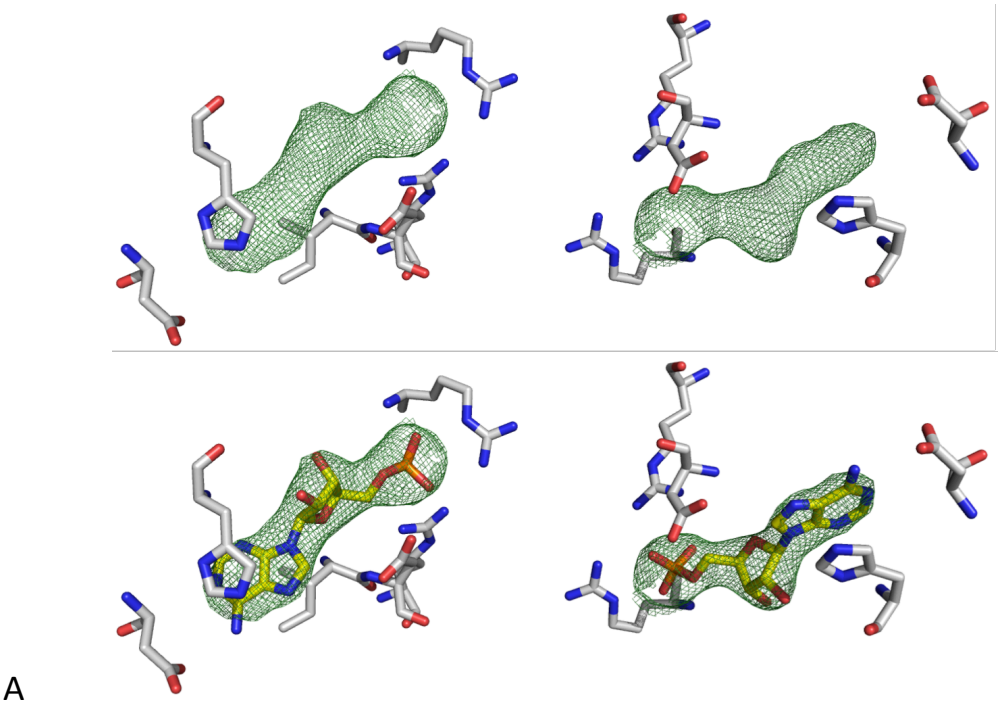
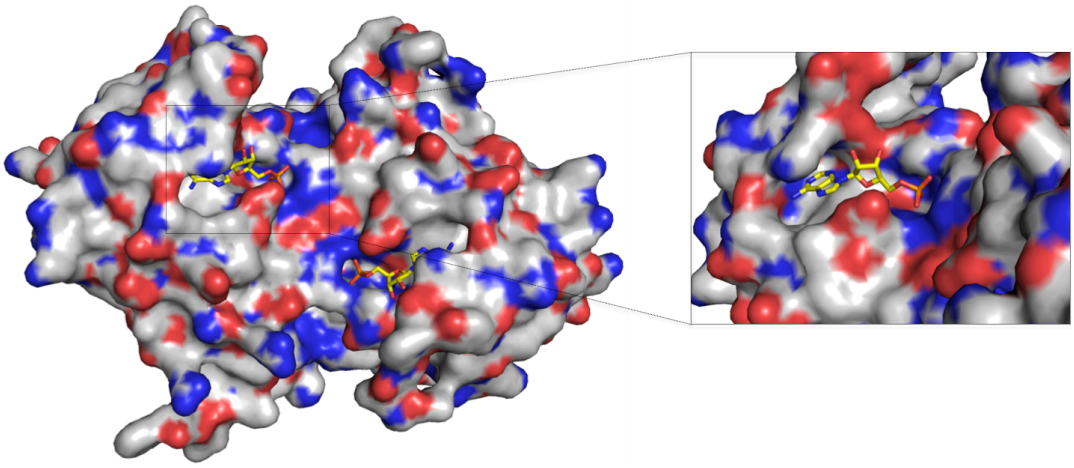


Figure 2.3 (A) X-ray crystal structure of *Sd*KefCTD highlighting the two monomers of the KTN dimer as red and grey cartoons. The additional electron densities observed in the nucleotide-binding pockets are shown as green meshes, with **AMP** molecules modeled in as sticks. Image constructed using PyMOL.[9] (B) The overlaid primary and secondary structures of the X-ray crystal structure of *Sd*KefCTD. α-helixes are shown as blue arrows labelled 'Helix' and β-sheets are shown as pink arrows labelled 'Strand'. Image constructed using CLC Main Workbench 7.6.2

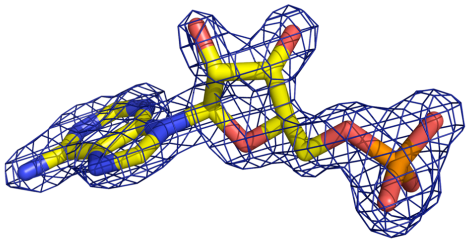
Unfortunately, at the time that the *in silico* protein-ligand docking studies were performed the X-ray crystal structure of *Sd*KefCTD had not been obtained. However, when the X-ray crystal structure of *Sd*KefCTD was resolved, Dr. Christos Pliotas made an interesting observation regarding the occupancy of the nucleotide-binding pockets in the Rossmann folds. These pockets were observed to contain electron density that corresponded to a bound nucleotide. The nucleotide that Dr. Christos Pliotas was able to model into this electron density was adenosine monophosphate (**AMP**; Figure 2.4 A & B). The modeling of **AMP** into *Sd*KefCTD was in accord with the X-ray crystal structure of *Ec*KefC (pdb code: 3L9W) published by Roosild *et al.*,[4] which also had additional electron density in the nucleotide-binding pocket that was attributed to **AMP** (Figure 2.4 C).



A



B



C

Figure 2.4 (A) Unbiased electron density for both nucleotide-binding pockets of the X-ray crystal structure of *Sd*KefCTD, contoured at σ = 1.5. **AMP** has been fitted into the electron density (sticks; carbon = yellow) (B) Surface representation the X-ray crystal structure of *Sd*KefCTD. The **AMP** molecules are shown as sticks (carbon = yellow); (C) 2 Fo – Fc map contoured at 2σ of **AMP** (sticks; carbon = yellow) identified in the nucleotide-binding pocket of the X-ray crystal structure 3L9W. Images constructed using PyMOL.[9]

The presence of the **AMP** molecules was curious for two main reasons. Firstly, **AMP** had not been present in the crystallisation mix, and so if the electron density in the nucleotide-binding pocket did correspond to **AMP**, then it must have been retained in the protein during the purification process. Secondly, an analysis of the 17 structures of KTN domains in the literature (excluding the wider RCK superfamily) finds that 5 have **NADH** bound and the remainder have **AMP**, **ADP** or **ATP** bound. Given the range of nucleotides observed to occupy KTN domains, questions were raised as to whether the electron density in the X-ray crystal structure of *Sd*KefCTD corresponded to **AMP**. Furthermore, if the electron density was found to arise from a bound **AMP** molecule, what role did the nucleotide play in Kef activity.

It was, therefore, proposed to assess the ability of a range of nucleotides to stabilise *Sd*KefCTD using differential scanning fluorimetry (DSF).[10] The information obtained from this experiment was used to guide ^{1}H Carr-Purcell-Meiboom-Gill (CPMG) NMR and analytical HPLC experiments to establish the identity of the bound nucleotide. Once the nucleotide had been identified, biological and chemical mutation studies were performed to rationalise the high affinity of the natively bound nucleotide and understand its role in *Sd*KefCTD.

2.2 Expression and purification of *Sd*KefCTD

In order to characterise the natively bound nucleotide in *Sd*KefCTD, the protein first needed to be expressed and purified. I performed the initial expression and purifications of *Sd*KefCTD, which was also required for biophysical assays (Chapters 3, 4 and 5), and Dr. Anthony Chan carried out subsequent purifications. The procedure followed to express and purify *Sd*KefCTD was essentially the same as the one published by Healy *et al.*,[1] however, a few minor modifications were made to adapt it to the laboratory in Oxford. These changes included: altering the media from LB (containing 0.1% *w/v* glucose) to 2× TY media (containing 0.2% *w/v* glucose, 16 g L^{-1} Oxoid^{TM}

Tryptone, 10 g L^{-1} Oxoid^{TM} Yeast Extract, 5 g L^{-1} NaCl); lysing the cells using sonication instead of a French press; using a manual purification column containing NOVAGEN® His·Bind® resin for the purification step.

The expression commenced with the pTrcSdKefQCTD vector, provided by the Miller group, University of Aberdeen, being transformed into the *E. coli* strain MJF373.[1] The transformed MJF373 cells were grown at 37 °C until an OD_{600} of 0.8 was reached. After cooling to 30 °C, expression was induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG; final concentration of 0.8 mM) for 4 h. After this time, the cells were lysed and the lysate was clarified by centrifugation, with the resulting supernatant being applied to a manual purification column containing NOVAGEN® His·Bind® Resin. The column was washed with buffer containing 30 mM imidazole, before being eluted with buffer containing 500 mM imidazole. The peak fractions, as determined by UV/Vis spectroscopy, were finally subjected to buffer exchange using a PD-10 column (GE Healthcare) into 50 mM phosphate buffer, 150 mM NaCl, pH 7.4. The protein was obtained in good purity, running on the SDS-PAGE gels with a mass corresponding to that of the monomer (26.4 kDa; Figure 2.5).

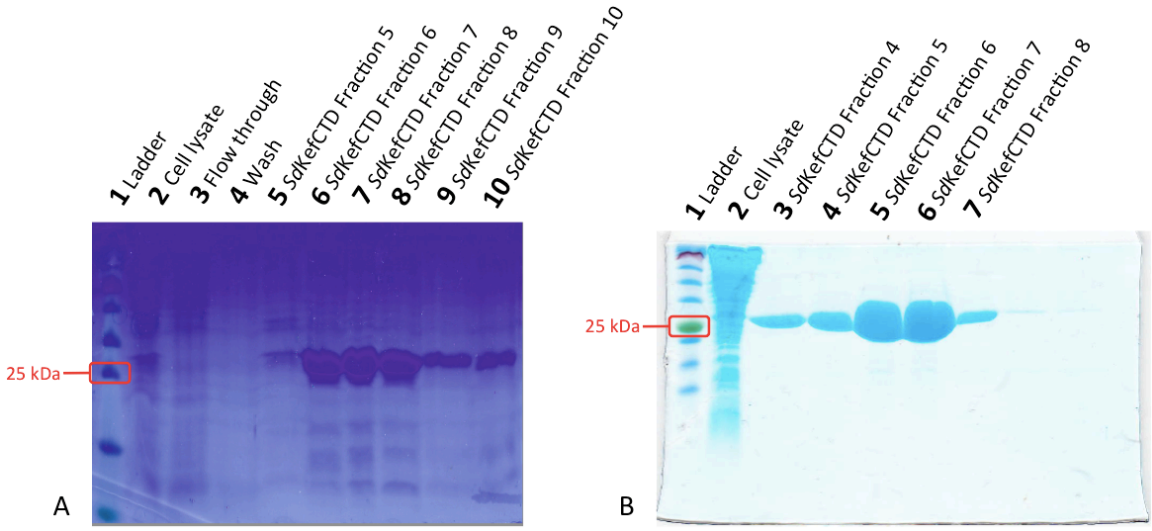


Figure 2.5 (A) SDS-PAGE gel analysis of *Sd*KefCTD purification performed by myself: Lane 1: Ladder, PageRuler™ Prestained Protein Ladder, 10 to 180 kDa; Lane 2: Unpurified cell lysate released from sonication of bacterial cells; Lanes 3: Flow through from loading HisTrap™ column; Lane 4: Wash of HisTrap™ column with wash buffer; Lanes 5-10: Elution fractions collected. (B) SDS-PAGE gel analysis of *Sd*KefCTD purification performed by Dr. Anthony Chan: Lane 1: Ladder, Color Prestained Protein Standard, Broad Range (New England BioLabs, Inc.); Lane 2: Unpurified cell lysate released from sonication of bacterial cells; Lanes 3-7: Elution fractions collected.

2.3 Differential scanning fluorimetry studies to assess the ability of a range of nucleotides to stabilise *Sd*KefCTD

With *Sd*KefCTD in hand, it was possible to explore the ability of the nucleotides **AMP**, **28**, **ADP**, **29**, **ATP**, **30**, **NADP**, **31**, **NAD$^+$**, **32**, and **NADH**, **33**, (Figure 2.6) to bind to and stabilise *Sd*KefCTD in the presence and absence of the regulatory peptides glutathione and glutathione S-conjugate **ESG** using differential scanning fluorimetry (DSF).[10]

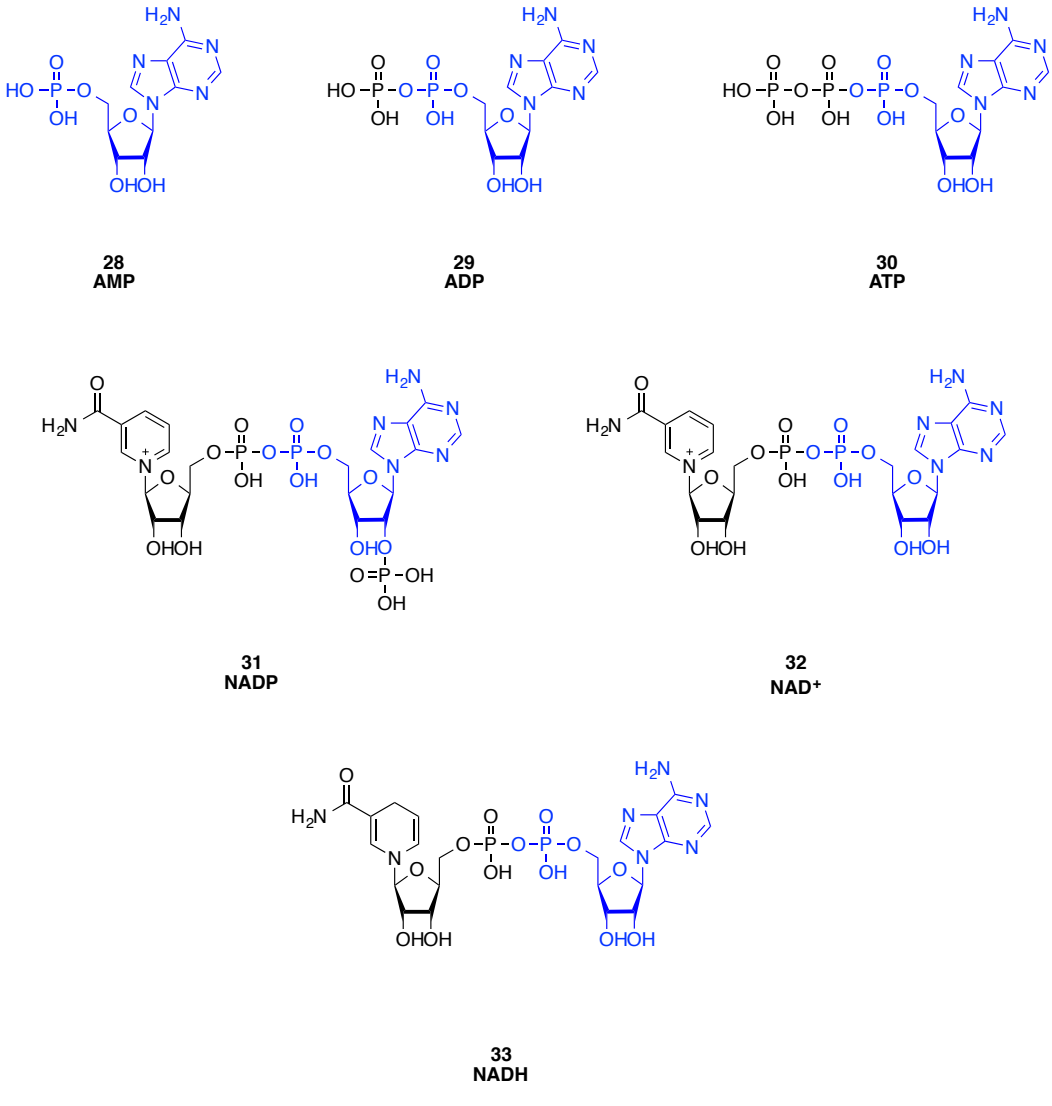


Figure 2.6 Structures of the nucleotides **AMP**, **ADP**, **ATP**, **NADP**, **NAD$^+$** and **NADH**.

2.4 Introduction to differential scanning fluorimetry (DSF)

Differential scanning fluorimetry (DSF)[10] is a technique that assesses a ligand's ability to bind to and stabilise a soluble protein. DSF uses a fluorescent dye called SYPRO® Orange to monitor the unfolding of a thermally treated protein (Figure 2.7). SYPRO® Orange is highly fluorescent in hydrophobic environments, and is quenched by aqueous ones. As the temperature of a protein is raised, it unfolds exposing hydrophobic regions that SYPRO® Orange binds to. The fluorescence intensity of SYPRO® Orange is consequently increased by the resulting hydrophobic interactions. Plotting the fluorescence intensity against temperature generates a sigmoidal curve that can be fitted using the Boltzmann equation to determine the inflection point of the transition curve (T_m). If a ligand is binding to and stabilising a protein, the protein should unfold at a higher temperature. The difference in the T_m in the presence and absence of a ligand (ΔT_m), therefore, provides a qualitative assessment of a ligand's ability to bind to and stabilise a protein. Generally, the magnitude of ΔT_m corresponds to the extent that the ligand stabilises the protein, with larger values corresponding to greater stabilisation.[10] It is worth noting that comparison of the magnitude of ΔT_m between ligands with very different physical properties can be misleading.

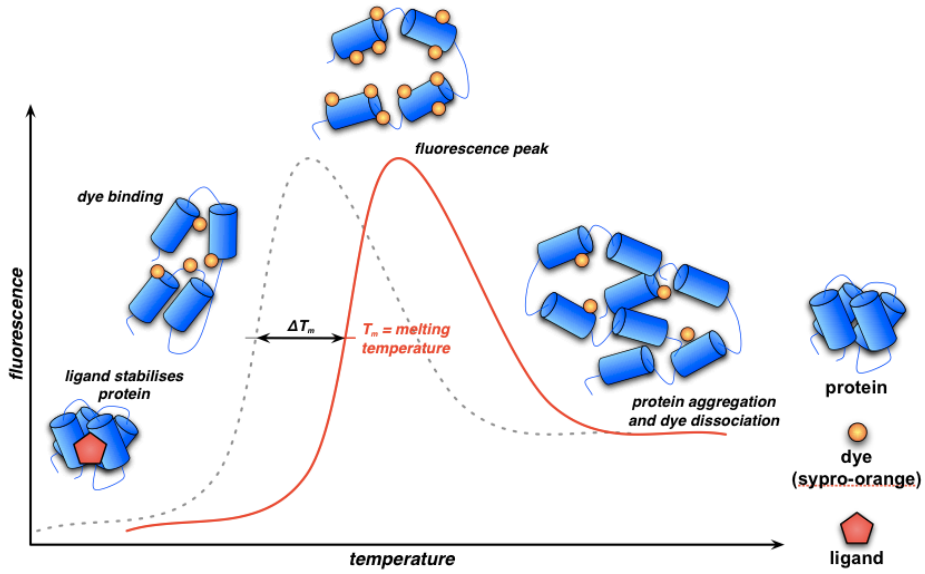


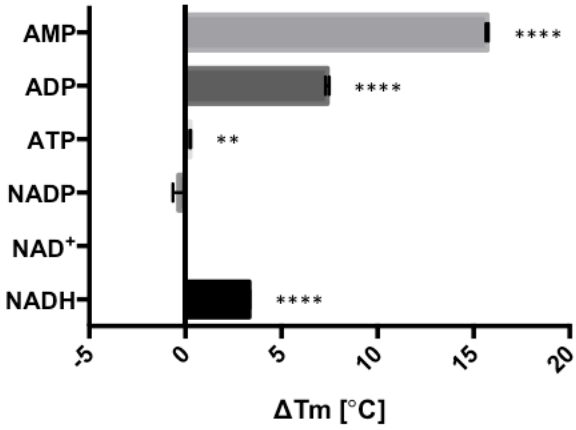
Figure 2.7 A visual representation of the theory behind differential scanning fluorimetry (DSF). Image provided by Prof. Stuart Conway.

2.5 DSF analysis of the stabilising effects of AMP, ADP, ATP, NADP, NAD^+ and NADH on *Sd*KefCTD

Although **AMP** had been modeled into the electron density found in the nucleotide-binding pocket, it was possible that a different adenine derived nucleotide was actually occupying it, with the remaining portion of the molecule not being resolved by X-ray crystallography. The possible candidates were determined to be **ADP**, **ATP**, **NADP**, **NAD^+** and **NADH**. It was therefore decided to test **AMP** and these nucleotides by DSF to see which of them stabilise *Sd*KefCTD. Those observed to stabilise *Sd*KefCTD might be able to occupy the nucleotide-binding pocket.

The DSF experiment identified **AMP**, **ADP** and **NADH** as being able to stabilise *Sd*KefCTD, with thermal shifts of ΔT_m = 16 °C, 7 °C and 3 °C respectively (Figure 2.8). The very high thermal shift of **AMP** supports the decision to model this nucleotide into the electron density occupying the nucleotide-binding pocket. **ATP**, **NADP** and **NAD^+** were ruled out from further investigations due to the lack of stabilisation that these nucleotides showed in the DSF experiment.

A



B

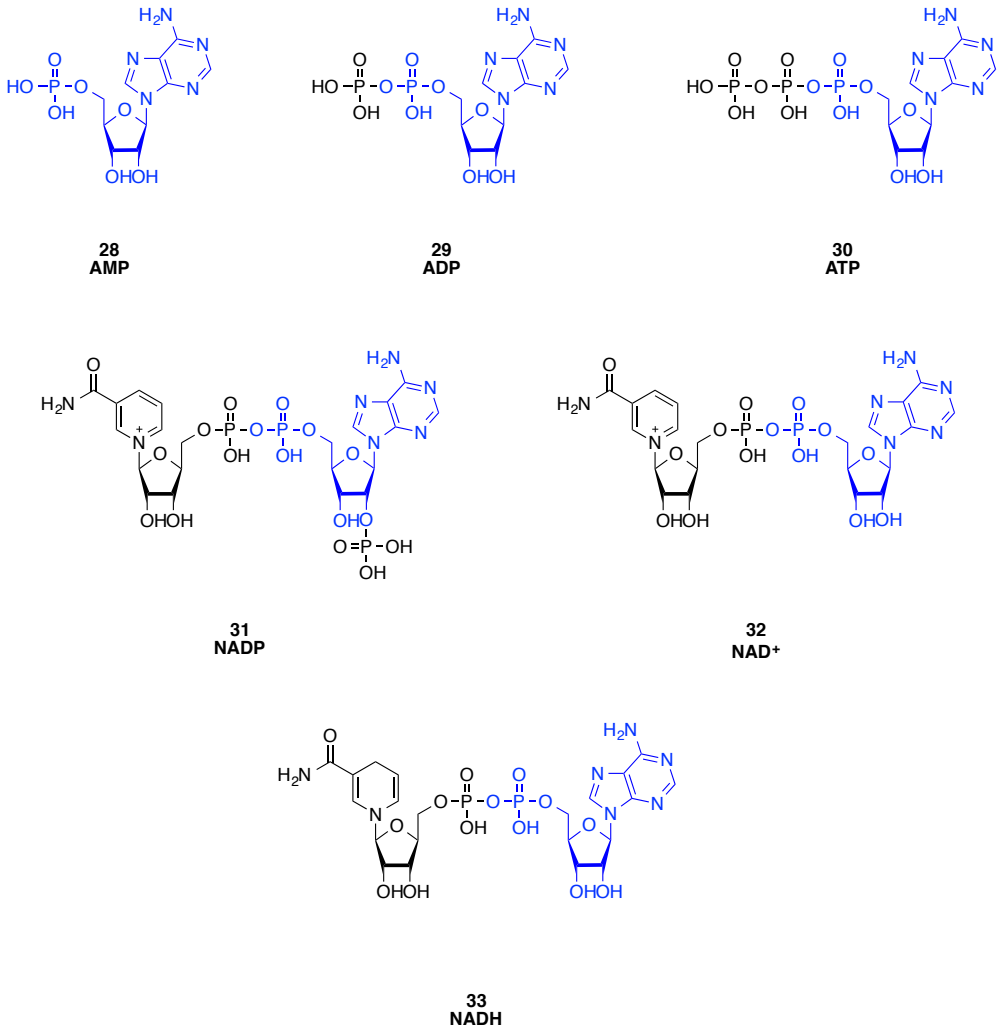
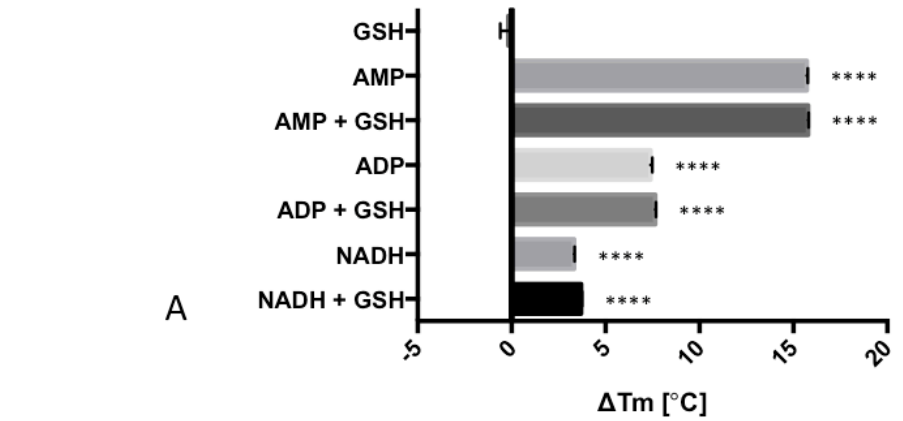


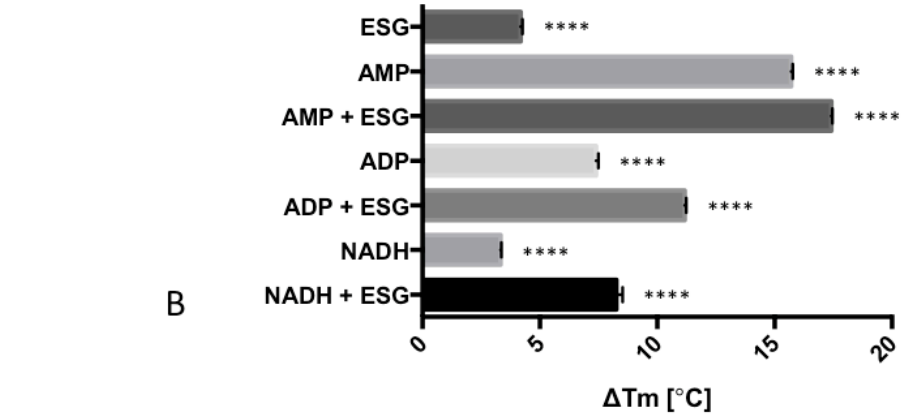
Figure 2.8 DSF analysis of the stabilising effects of the nucleotides **AMP**, **ADP**, **ATP**, **NADP**, **NAD$^+$** and **NADH** on *Sd*KefCTD (12 µM). (A) Shows changes in melting temperature in the presence of the indicated nucleotides (final concentrations of 1 mM) relative to that of protein + buffer. Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) The structures of the nucleotides tested.

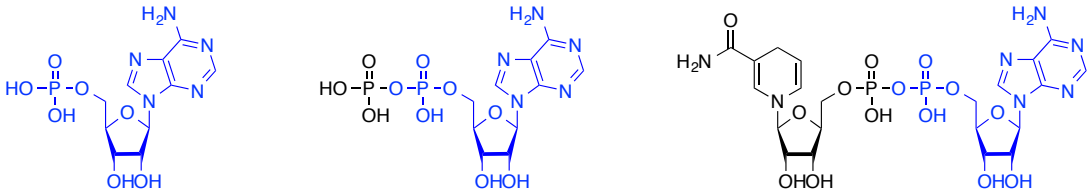
2.6 DSF analysis of the stabilising effects of AMP, ADP and NADH on *Sd*KefCTD in the presence of glutathione and the glutathione S-conjugate ESG

Homodimeric *Sd*KefCTD has two nucleotide- and two glutathione-binding sites. These sites are in close proximity, such that the X-ray crystal structure of *Sd*KefCTD predicts that the residue R416 interacts with the phosphate group of **AMP** and the backbone carbonyl oxygen of I505. The residue I505 is located in the glutamate region of the glutathione-binding pocket and so R416 appears to act as a link between the nucleotide- and glutathione-binding sites. To investigate whether, the binding of the nucleotides and glutathione or glutathione S-conjugate **ESG** were competitive, DSF experiments were performed to compare the stabilising effects of the nucleotides **AMP**, **ADP** and **NADH** on their own, to when they were in the presence of glutathione (Figure 2.9 A) or **ESG** (Figure 2.9 B).

The ΔT_m of glutathione alone was not statistically significant, and too small to draw conclusions on the competitive nature of the nucleotide and peptide binding (Figure 2.9 A). The ΔT_m of **ESG** alone, however, was approximately 4 °C, allowing the relationship between the nucleotide and peptide binding sites to be probed. The ΔT_m obtained for **ESG** and each of the nucleotides alone were found to be additive when **ESG** and each of the nucleotides were tested alongside one another (Figure 2.9 B). The additive nature of the stabilisation provided by **ESG** and each of the nucleotides suggests that the nucleotide- and peptide-binding sites are independent of one another. An additive effect would not be expected if **ESG** was unable to bind when a nucleotide was bound and *vice versa*. The implication of these results is that the nucleotides and peptides are able to bind to *Sd*KefCTD simultaneously.





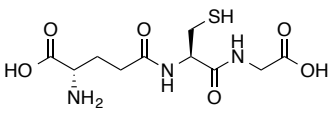


28
AMP

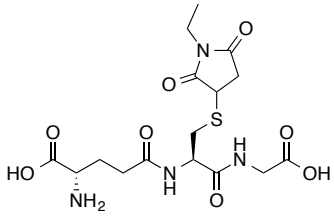
29
ADP

33
NADH

C



11
GSH



18
ESG

Figure 2.9 DSF analysis of the stabilising effects of the nucleotides **AMP**, **ADP** and **NADH** on *Sd*KefCTD (12 µM) in the presence of the regulatory peptides (A) Glutathione (**GSH**) and (B) **ESG**. Graphs show changes in melting temperature in the presence of the indicated ligands (final concentrations of 1 mM) relative to that of protein + buffer. Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (C) The structures of the compounds tested.

2.7 Establishing the identity of the nucleotide natively bound to *Sd*KefCTD using ^{1}H CPMG NMR

To establish the identity of the nucleotide that was natively bound to *Sd*KefCTD, Carr-Purcell-Meiboom-Gill (CPMG)[11-13] edited ^{1}H NMR experiments were performed in collaboration with Amjad Khan in the Claridge Group, University of Oxford. CPMG edited ^{1}H NMR experiments are typically implemented to attenuate broad protein resonances in a ^{1}H NMR spectrum so that it is possible to differentiate between a bound and free ligand.[11-13] When a ligand is bound to a protein, its NMR resonances broaden and lower in intensity due to it tumbling slower in the bound state, and consequently having a faster transverse relaxation rate (R_2). CPMG edited ^{1}H NMR experiments are usually performed by titrating protein into a ligand sample and observing a broadening of the ligand signals. It was speculated that this phenomenon could be reversed, such that if a nucleotide were natively bound to *Sd*KefCTD then no resonances would be observed for the nucleotide in the non-denatured spectrum of *Sd*KefCTD. However, upon denaturing *Sd*KefCTD, the nucleotide would be released into solution and the resonances corresponding to it should appear in the CPMG edited ^{1}H NMR spectrum. CPMG edited ^{1}H NMR experiments were therefore used to characterise the bound nucleotide by obtaining spectra of both non-denatured and denatured *Sd*KefCTD. Comparison of the new peaks that arose in the denatured spectrum to reference spectra of the most likely candidates identified in the DSF studies (**AMP**, **ADP** and **NADH**) revealed the identity of the bound nucleotide.

CPMG edited ^{1}H NMR experiments were thus performed on both non-denatured and denatured *Sd*KefCTD. The non-denatured protein spectrum (Figure 2.10 C) showed no resonances corresponding to **AMP**, **ADP** or **NADH**, suggesting that any nucleotide present was bound to the protein. In the denatured *Sd*KefCTD spectrum, however, sharp signals were observed corresponding to a free ligand (Figure 2.10 B). Comparison of the denatured spectrum of

*Sd*KefCTD to the reference spectra of **AMP**, **ADP** and **NADH** suggested that the free ligand in solution was **AMP** (Figure 2.10 A).

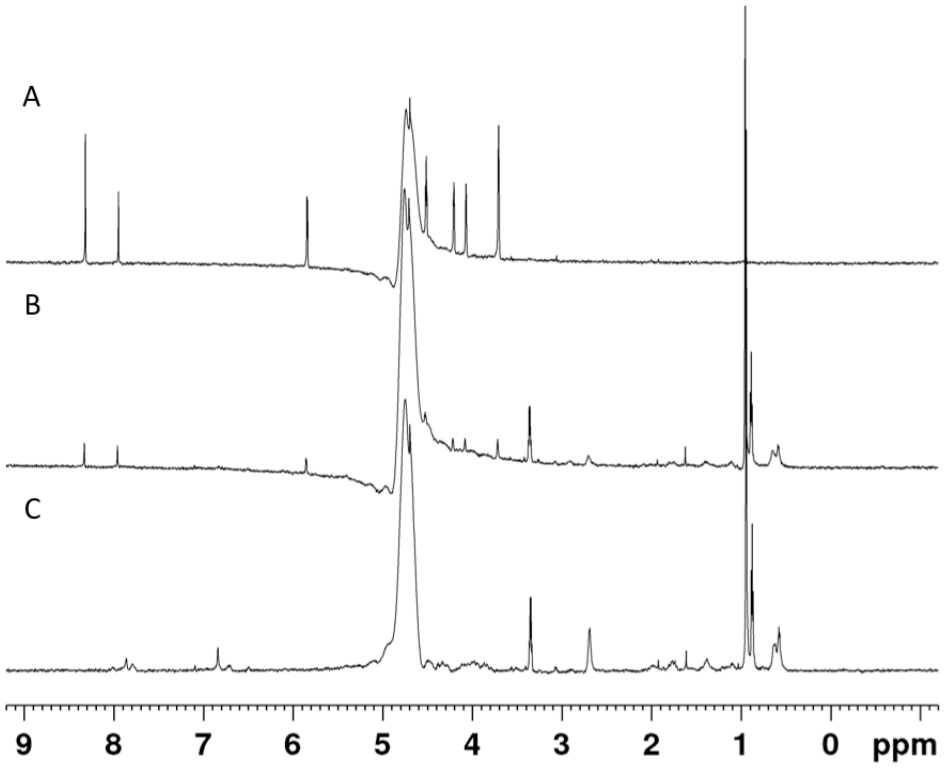


Figure 2.10 Overlay of (A) CPMG edited ^{1}H NMR reference spectrum of **AMP** (330 µM); (B) CPMG edited ^{1}H NMR spectrum of *Sd*KefCTD protein (330 µM) after denaturation at 80 °C for 3 h; (C) CPMG edited ^{1}H NMR spectrum of *Sd*KefCTD protein (330 µM) before denaturation.

ADP and **NADH** were exposed to the same conditions used to denature *Sd*KefCTD (heating for 3 h at 80 °C) to confirm that **AMP** was not produced by the hydrolysis of **ADP** or **NADH**. The results of these experiments are shown in Figure 2.11 and Figure 2.12. Although new peaks were observed after heat treatment, **AMP** did not appear to be formed in either spectrum. Furthermore, overlaying the heat-treated **ADP** and **NADH** spectra with that of the denatured protein suggests that neither of them is the nucleotide natively bound to *Sd*KefCTD.

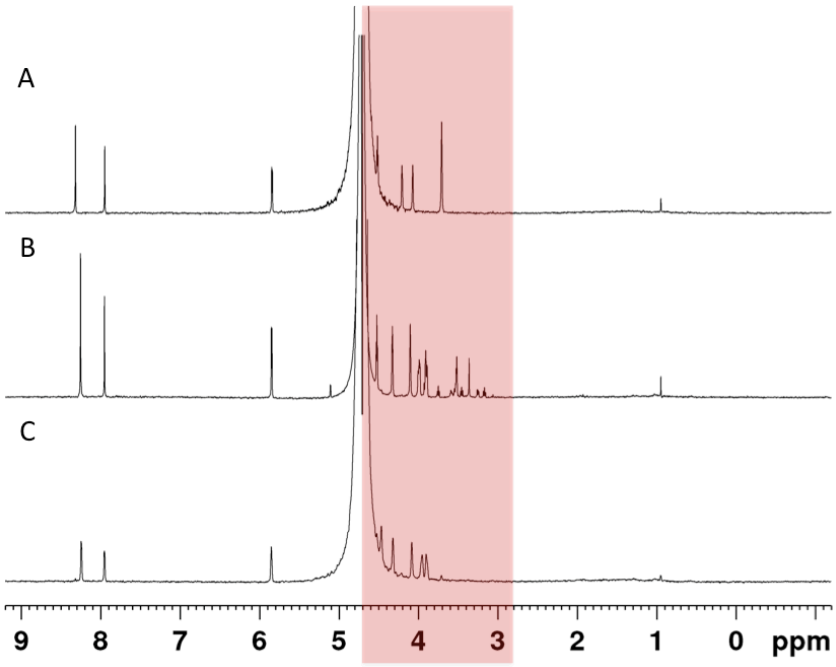


Figure 2.11 Overlay of (A) CPMG edited ^{1}H NMR reference spectrum of **AMP** (330 µM); (B) CPMG edited ^{1}H NMR reference spectrum of **ADP** (330 µM) after heating to 80 °C for 3 h; (C) CPMG edited ^{1}H NMR reference spectrum of **ADP** (330 µM). The red transparent box highlights the region of the spectrum where the most significant differences between the CPMG edited ^{1}H NMR spectra of **AMP** and **ADP** (before and after heating) occur.

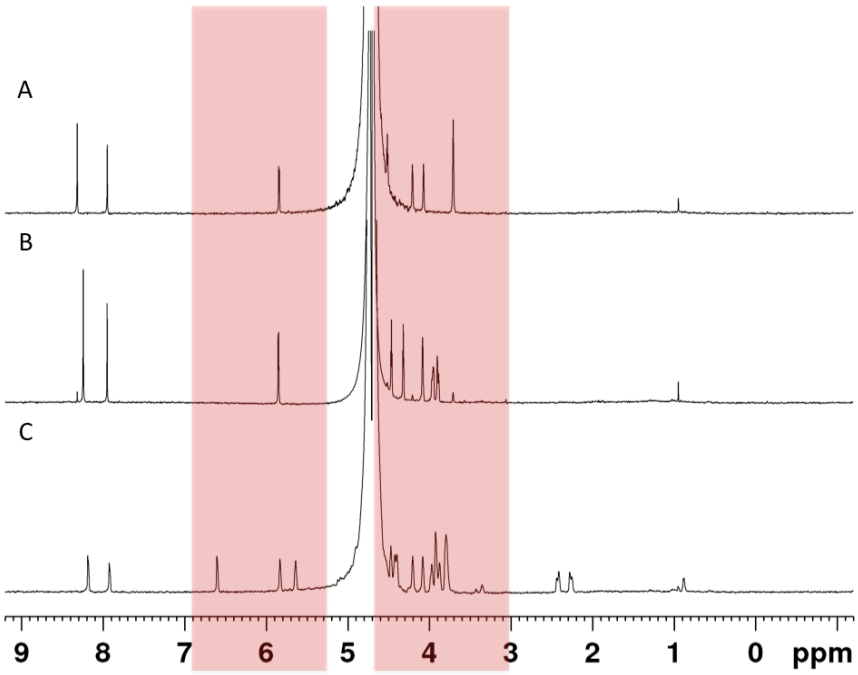


Figure 2.12 Overlay of (A) CPMG edited ^{1}H NMR reference spectrum of **AMP** (330 µM); (B) CPMG edited ^{1}H NMR reference spectrum **NADH** (330 µM) after heating to 80 °C for 3 h; (C) CPMG edited ^{1}H NMR reference spectrum of **NADH** (330 µM). The red transparent boxes highlight the regions of the spectrum where the most significant differences between the CPMG edited ^{1}H NMR spectra of **AMP** and **NADH** (before and after heating) occur.

AMP was also subjected to the same heat treatment as *Sd*KefCTD and was found to be stable under these conditions (Figure 2.13 A & B). This was further confirmed by overlaying the **AMP** spectra with an adenosine reference spectrum (Figure 2.13 C).

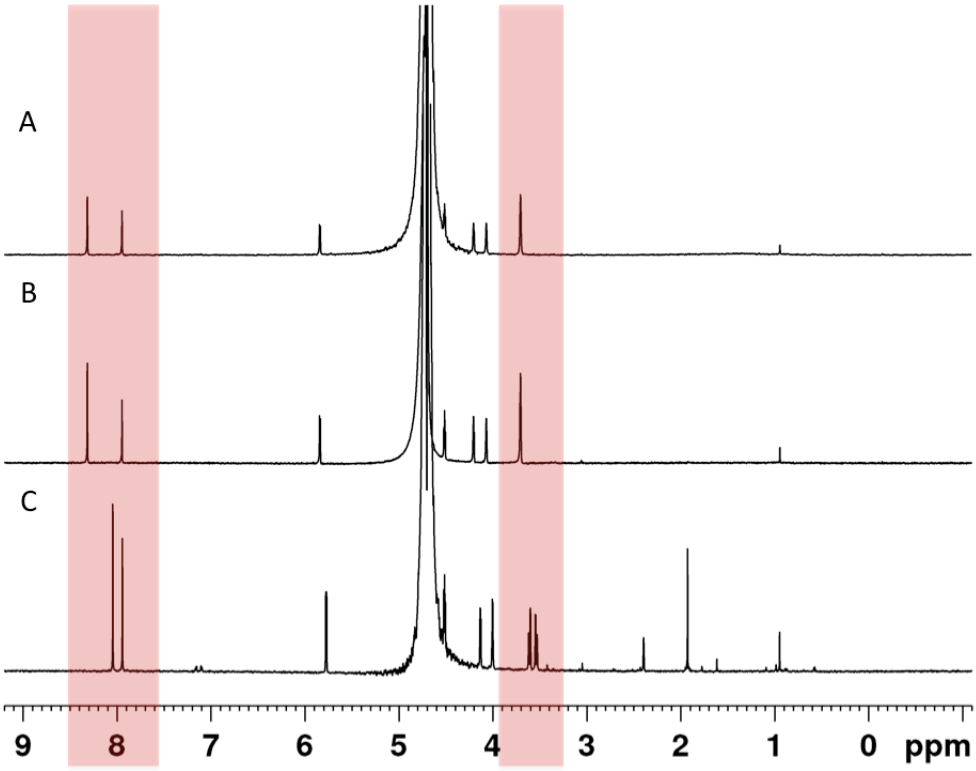


Figure 2.13 Overlay of (A) CPMG edited ^{1}H NMR reference spectrum of **AMP** (330 µM); (B) CPMG edited ^{1}H NMR reference spectrum of **AMP** (330 µM) after heating to 80 °C for 3 h; (C) CPMG edited ^{1}H NMR reference spectrum of adenosine (330 µM). The red transparent boxes highlight the regions of the spectrum where the most significant differences between the CPMG edited ^{1}H NMR spectra of **AMP** and adenosine occur.

2.8 Verifying the identity of the nucleotide using analytical HPLC

To further verify that the identity of the natively bound nucleotide was **AMP**, purified *Sd*KefCTD was denatured by heating at 95 °C for 5 min, the supernatant was filtered through a 30 kDa cut-off concentrator and the filtrate was subjected to analysis by HPLC. The resulting spectrum was compared to analytical HPLC traces of **AMP** (Figure 2.14 A), **ADP** (Figure 2.14 B) and **NADH** (Figure 2.14 C). An overlaid trace of **AMP**, **ADP** and **NADH** (Figure 2.14 D) shows that these

nucleotides have distinct enough retention times, 6.3 min, 2.3 min and 6.7 min, respectively, that it would be possible deduce the identity of the nucleotide using analytical HPLC. Figure 2.14 E shows the analytical HPLC trace of the filtrate from the denatured sample of *Sd*KefCTD that contains a peak with a retention time of 6.5 min. Collecting the peak obtained in Figure 2.14 E and subjecting it to low resolution mass spectrometric analysis gave a peak in the negative mass spectrum of 346 corresponding to the mass of **AMP** minus a proton. Furthermore, spiking a sample of the filtrate of denatured *Sd*KefCTD with the same concentration of **AMP** gave a single peak, suggesting that the peak corresponds to **AMP** (Figure 2.14 F).

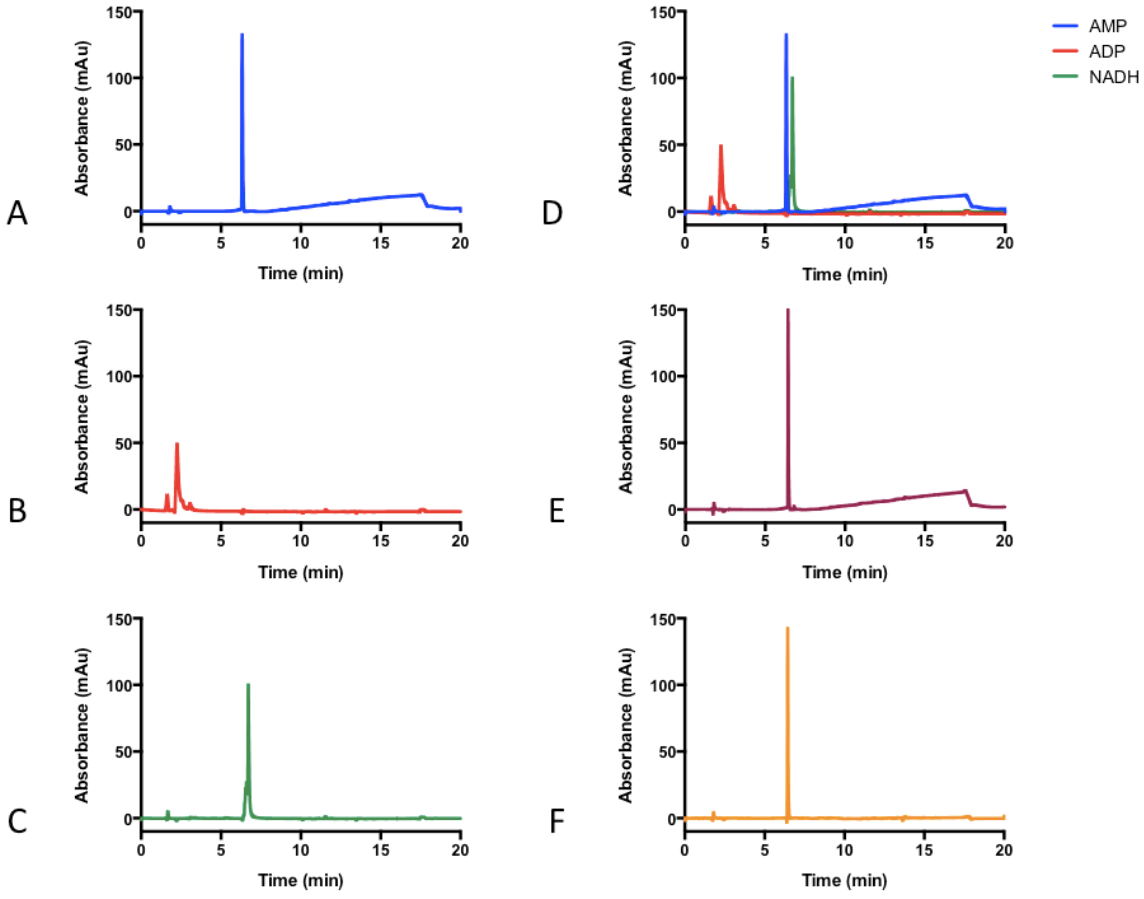


Figure 2.14 Analytical HPLC traces of (A) **AMP** (50 µM); (B) **ADP** (50 µM); (C) **NADH** (50 µM); (D) **AMP** (50 µM), **ADP** (50 µM) and **NADH** (50 µM) overlaid; (E) The supernatant of wild-type *Sd*KefCTD (50 µM) denatured by heating to 95 °C for 5 min; (F) A spiking experiment, where denatured wild-type *Sd*KefCTD (25 µM) and a pure sample of **AMP** (25 µM) were mixed and injected together.

2.9 Investigating the stoichiometry of the nucleotide-binding pockets using nanoelectrospray mass spectrometry

With strong evidence supporting the view that the identity of the bound nucleotide was **AMP**, it was decided to investigate whether there were two **AMP** molecules per *Sd*KefCTD homodimer, using native mass spectrometry. The rationale behind this line of investigation was that both of the nucleotide-binding pockets in the X-ray crystal structure of *Sd*KefCTD were occupied by additional electron density.

Shane Chandler in the Benesch Group, University of Oxford, performed nanoelectrospray mass spectrometry measurements to confirm the presence of the homodimer and determine the stoichiometry of **AMP** (347 Da) bound to *Sd*KefCTD. The nanoelectrospray mass spectrometry measurements were performed under conditions that preserved noncovalent interactions in the gas phase.[14] The mass spectra obtained found that two **AMP** molecules were bound to an *Sd*KefCTD homodimer (53,233 Da) for the charge states +13, +14 and +15 (Figure 2.15). Step-wise loss of bound **AMP** as a neutral species from *Sd*KefCTD was then achieved by collisionally activating the complex in the collision cell. This resulted in the presence of three species for each of the homodimer charge states that correspond to the loss of zero, one and two **AMP** molecules from the *Sd*KefCTD complex. The use of mass spectrometry, under native conditions, therefore suggested that both of the nucleotide-binding pockets in the *Sd*KefCTD homodimer are fully occupied with one **AMP** molecule each.

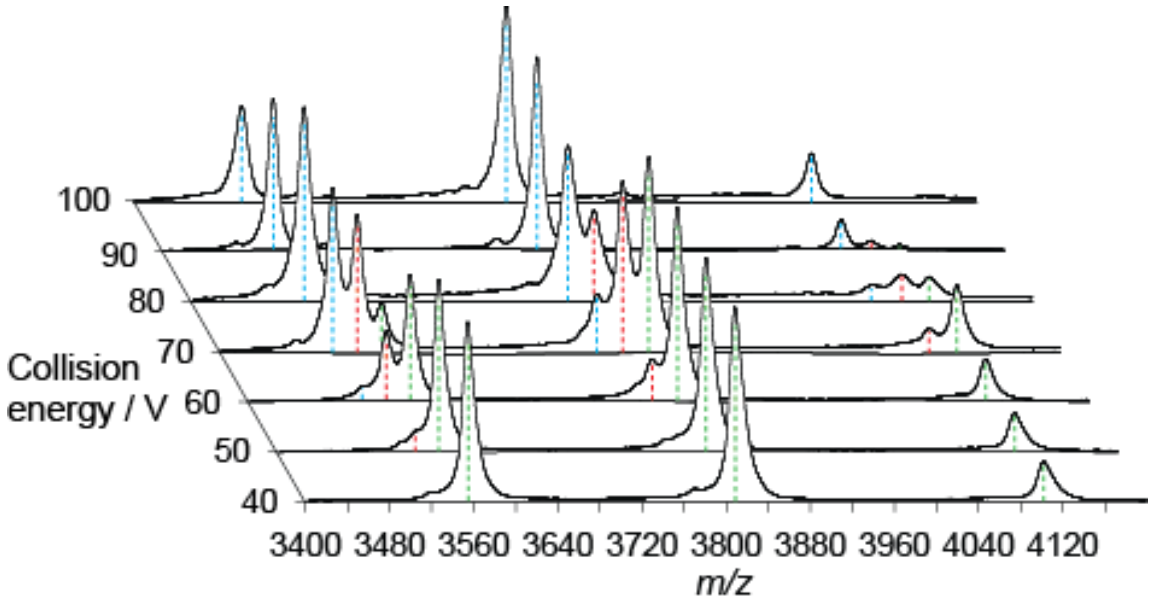


Figure 2.15 Native mass spectrometry investigations to detect the presence of two **AMP** molecules per homodimer of *Sd*KefCTD. Three species associated to the homodimer charge states are observed upon increasing the collision energy. These charge states correspond to the singly (red) and doubly (green) bound **AMP** and the apo (blue) form of the *Sd*KefCTD homodimer.

It is worth highlighting here that the DSF results do not initially appear to corroborate the observation that the nucleotide-binding pockets are fully occupied with **AMP**, because it would not be expected that additional **AMP** could bind to and stabilise *Sd*KefCTD. The stabilisation observed by DSF can, however, be explained if you consider that as the protein unfolds, its affinity for **AMP** is lowered. Thus, an increase in temperature might also cause the off-rate of the bound **AMP** to increase, vacating the nucleotide-binding sites. In this situation, having an excess of surrounding **AMP** molecules would be beneficial, as they could reoccupy the vacant site to help stabilise *Sd*KefCTD.

2.10 Understanding the affinity and role of AMP in *Sd*KefCTD

In order to rationalise the high ΔT_m and presumed high affinity that **AMP** shows for *Sd*KefCTD, and understand why it is present, the binding mode of **AMP** in the X-ray crystal structure of *Sd*KefCTD was analysed (Figure 2.16).

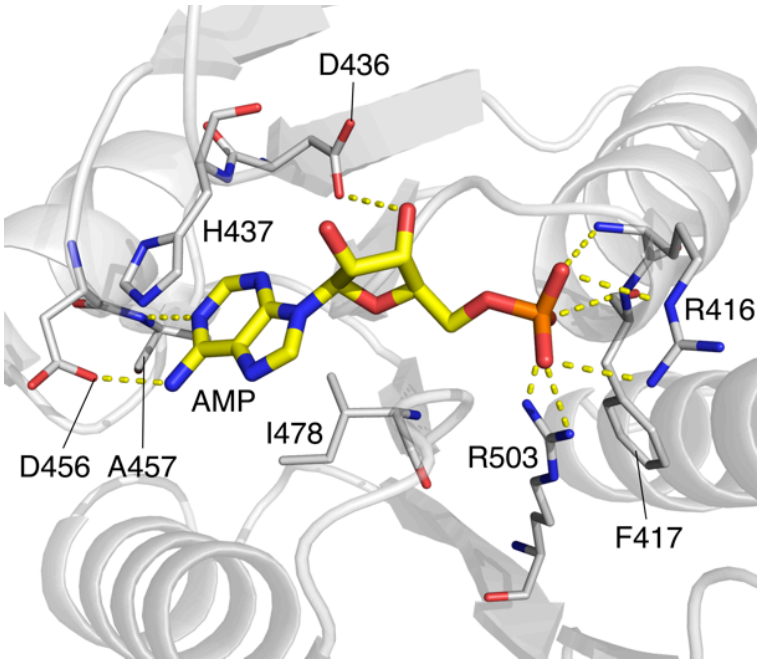


Figure 2.16 The binding mode of the native **AMP** molecule (shown as sticks, carbon = yellow) identified in the X-ray crystal structure of *Sd*KefCTD (grey cartoon). Predicted polar interactions are shown as yellow dashed lines. Image constructed using PyMOL.[9]

The adenine ring of **AMP** is located in a pocket encompassed by I478 and H437. H437 and the purine ring are 3.9-4.0 Å apart, a distance over which π-stacking between the two aromatic rings is possible. The adenine ring also appears to form hydrogen bonds to *Sd*KefCTD *via* the aniline-like amine to D456 and the adjacent pyridine-like nitrogen to the backbone NH of A457. Further hydrogen bonds are proposed to form between the hydroxyl groups on the ribose ring of **AMP** with the residue D436. The final interactions are through the phosphate group, which forms salt bridges to R416 and R503, and hydrogen bonds to the backbone NH of R416 and F417.

It was proposed to probe the significance of these interactions by biologically mutating *Sd*KefCTD and chemically mutating **AMP**.

2.11 Biological mutations of the residues H437 and D436 in full-length *Sd*Kef and *Sd*KefCTD

Silvia Ekkerman in the Miller Group, University of Aberdeen, performed a mutational analysis of the residues H437 and D436 in *Sd*KefCTD and full-length *Sd*Kef protein.

2.12 Biological mutations of the residue H437 in *Sd*KefCTD and full-length *Sd*Kef

To investigate whether π-stacking interactions between the aromatic rings of H437 and the purine ring of **AMP**, contribute to the high affinity that **AMP** shows for *Sd*KefCTD, H437 was mutated to H437A and H437N. These mutations were found to have similar expression levels to wild-type *Sd*KefCTD and full-length *Sd*Kef. Furthermore, *Sd*Kef (H437A) and *Sd*Kef (H437N) were both active in a K^+ efflux assay, with H437N exhibiting close to wild-type activity and *Sd*Kef (H437A) exhibiting approximately half wild-type activity. This mutational analysis suggests that if present, π-stacking interactions between H437 and the purine ring of **AMP** make minimal contributions to the affinity.

2.13 Biological mutations of the residue D436 in *Sd*KefCTD and full-length *Sd*Kef

To investigate whether hydrogen bonds between D436 and the hydroxyl groups on the ribose ring of **AMP**, contribute to the high affinity that **AMP** shows for *Sd*KefCTD, D436 was mutated to D436E, D436N and D436A. These mutations were selected for the following reasons: D436E should form similar interactions with **AMP**; D436N is not charged and so should form weaker interactions with **AMP**; D436A would be unable to hydrogen bond to **AMP**. D436E demonstrated

similar expression levels to wild-type *Sd*KefCTD and full-length *Sd*Kef. *Sd*Kef (D436E) was also active in a K^+ efflux assay, exhibiting close to wild-type activity. D436N and D436A demonstrated significantly reduced expression levels to wild-type *Sd*KefCTD and full-length *Sd*Kef, with no protein detected for the D436N mutant of *Sd*KefCTD. *Sd*Kef (D436N) and *Sd*Kef (D436A) also showed severely reduced activity in a K^+ efflux assay. This mutational analysis suggests that hydrogen bonding interactions between D436 and the hydroxyl groups on the ribose ring of **AMP** contribute to the high affinity of **AMP** for *Sd*KefCTD.

2.14 Comparison of the stabilising effects of AMP and adenosine on *Sd*KefCTD by DSF

Given that the most significant interactions were predicted to come through the phosphate group of **AMP**, the effect of removing the phosphate on **AMP**'s ability to stabilise *Sd*KefCTD was investigated by DSF. The compound resulting from the chemical mutation of removing the phosphate group from **AMP** is adenosine (Figure 2.17 B). Adenosine was found to have poor solubility and so the DSF assay had to be run in **DMSO**. A control was therefore put in place that involved running **AMP** in **DMSO**. The presence of **DMSO** was found to have no effect on **AMP**'s ability to stabilise *Sd*KefCTD (Figure 2.17 A). Furthermore, adenosine was not found to stabilise *Sd*KefCTD, supporting the hypothesis that the majority of **AMP**'s affinity for *Sd*KefCTD is derived from interactions made by the phosphate group (Figure 2.17).

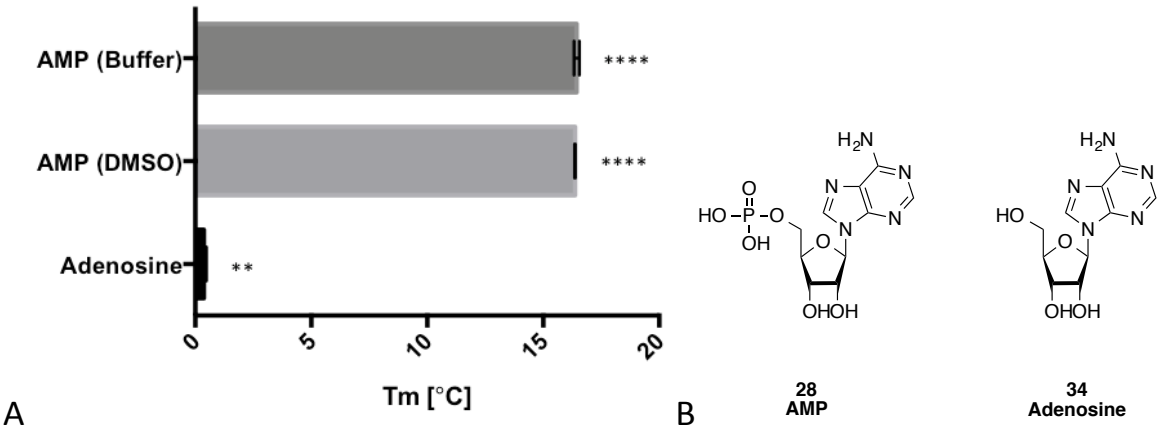


Figure 2.17 DSF analysis of the stabilising effects of **AMP** and adenosine on *Sd*KefCTD (12 μM). (A) Shows changes in melting temperature in the presence of the indicated ligands (final concentrations of 1 mM) relative to that of protein + buffer or protein + **DMSO**. Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) The structures of **AMP** and adenosine, **34**.

2.15 Mutational analysis of R416E

The interactions of the **AMP** phosphate group with *Sd*KefCTD were further investigated by mutational analysis of the residue R416. Silvia Ekkerman in the Miller Group, University of Aberdeen, performed the mutational work. The *Sd*KefCTD mutant R416E was also subsequently expressed and purified by Dr. Anthony Chan in the Conway Group.

R416 is proposed to form a salt bridge with the phosphate group of **AMP** in the X-ray crystal structure of *Sd*KefCTD (Figure 2.16). To investigate the contribution of this salt bridge to **AMP**'s affinity for *Sd*KefCTD, R416 was mutated to R416A, R416M and R416E. *Sd*Kef (R416A) and *Sd*Kef (R416M) demonstrated significantly reduced expression levels compared to wild-type full-length *Sd*Kef and *Sd*KefCTD (R416E) demonstrated significantly reduced expression levels compared to wild-type *Sd*KefCTD. Testing the mutants in the K^+ efflux assay found that *Sd*Kef (R416A) retained some activity, *Sd*Kef (R416M) had significantly reduced activity and *Sd*Kef (R416E) showed no

activity. The level of activity in the K^+ efflux assay shown by each mutant correlates with the expected disruption that each mutation would make to **AMP** binding. Although both of the mutations R416A and R416M remove the salt bridge to the **AMP** phosphate group, R416A should still be able to accommodate the **AMP** in the binding site, whereas the more sterically bulky R416M might not. R416E is the most disruptive mutation, as the **AMP** phosphate group would have unfavourable interactions with the negatively charged ɣ-carboxylate of glutamate.

2.16 DSF analysis of the stabilising effects of AMP on the *Sd*KefCTD R416E mutant

The effect of removing the salt bridge between the **AMP** phosphate group and the residue R416 on **AMP**'s ability to bind to and stabilise *Sd*KefCTD was investigated using DSF (Figure 2.18). This was achieved by comparing the ΔT_m values of wild-type *Sd*KefCTD and the *Sd*KefCTD mutant R416E in the presence of excess **AMP**. It was observed that ΔT_m = 16 °C for wild-type *Sd*KefCTD in the presence of **AMP**, whereas no stabilisation was observed in the presence of the *Sd*KefCTD mutant R416E. This suggests that **AMP** is unable to bind when the mutation R416E is present, highlighting the importance of the salt bridge that R416 forms with the **AMP** phosphate group.

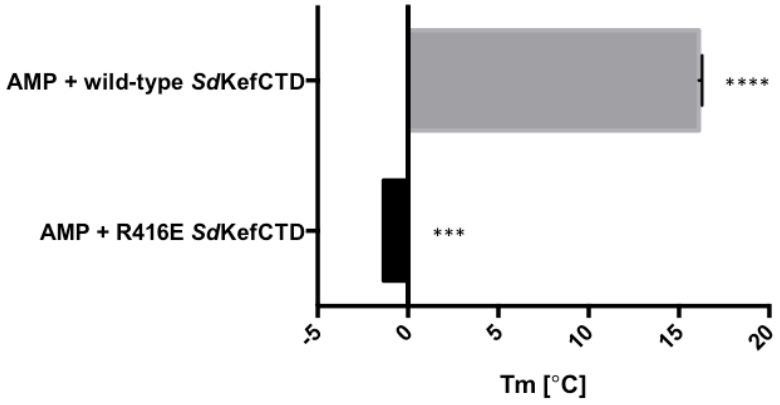


Figure 2.18 DSF analysis of the stabilising effects of the nucleotide **AMP** in the presence of wild-type *Sd*KefCTD (8.7 µM) and the mutant R416E of *Sd*KefCTD (8.7 µM). The graph shows changes in melting temperature in the presence of the indicated ligands (final concentrations of 1 mM) relative to that of protein + buffer. Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$).

It is, however, worth noting that the transition curves for the *Sd*KefCTD mutant R416E were not ideal (Figure 2.19), as highlighted by comparing them to the curves obtained for wild-type *Sd*KefCTD (Figure 2.20).

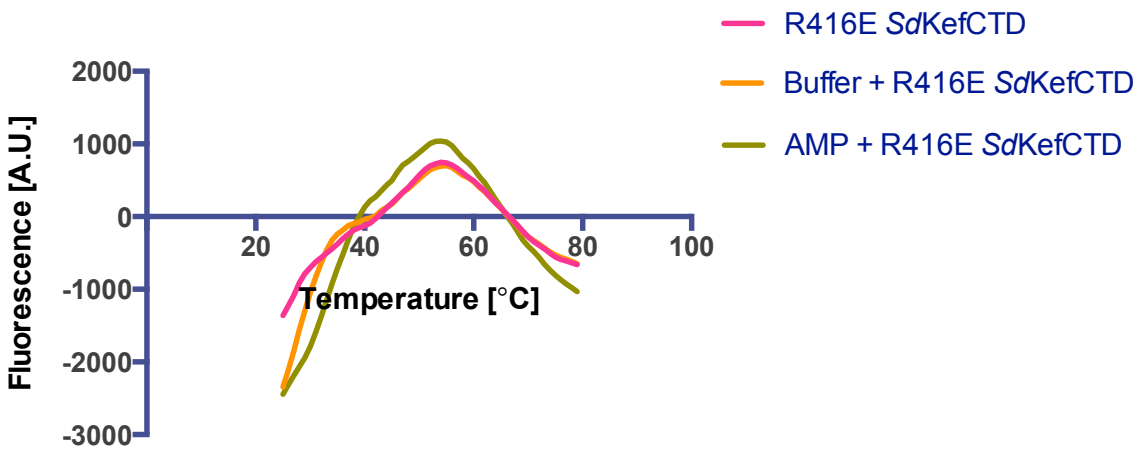


Figure 2.19 DSF melting curves of R416E *Sd*KefCTD (8.7 µM) in the absence and presence of **AMP** (1 mM). Curves show an average of triplicate runs.

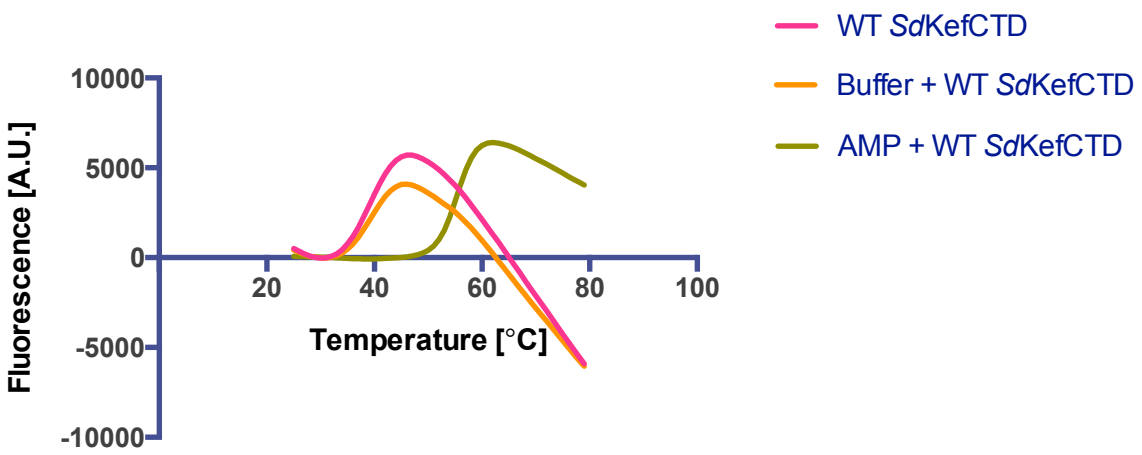


Figure 2.20 DSF melting curves of wild-type *Sd*KefCTD (8.7 µM) in the absence and presence of **AMP** (1 mM). Curves show an average of triplicate runs.

As the ΔT_m values are calculated using the Boltzmann equation highlighted in Equation 2.1 and there was not a clearly defined minimum intensity (LL) for the *Sd*KefCTD mutant R416E, the shifts were calculated by assigning LL as the intensity of the curve at 45 °C. Caution therefore

needs to be taken when interpreting these results, but it is clear that R416 is important for the stability of the protein and the binding of **AMP**.

$$y = LL + \frac{(UL - LL)}{1 + e^{(\frac{T_m - x}{a})}}$$

Equation 2.1 Boltzmann equation: LL and UL are the values of minimum and maximum intensities, respectively, and *a* denotes the slope of the curve within Tm.

2.17 Establishing whether AMP is retained during the purification of the mutant R416E by ^{1}H CPMG NMR

Having established that **AMP** does not have a stabilising effect on the R416E mutant of *Sd*KefCTD, the mutant was subjected to CPMG edited ^{1}H NMR experiments to determine whether **AMP** was retained during its purification. CPMG edited ^{1}H NMR experiments were performed in collaboration with Amjad Khan in the Claridge Group, University of Oxford, on the R416E mutant of *Sd*KefCTD that was expressed and purified by Dr. Anthony Chan.

CPMG edited ^{1}H NMR experiments were thus performed on both the non-denatured and denatured R416E mutant of *Sd*KefCTD. The non-denatured protein spectrum (Figure 2.21 C) showed no resonances corresponding to **AMP**, suggesting that if **AMP** was present, it was bound to the protein. The spectrum of the denatured R416E mutant of *Sd*KefCTD (Figure 2.21 B) also had no signals corresponding to the **AMP** reference spectrum (Figure 2.21 A). This result suggests that *Sd*KefCTD no longer has a high enough affinity for **AMP** to retain it during the expression and purification process when R416 is mutated to R416E.

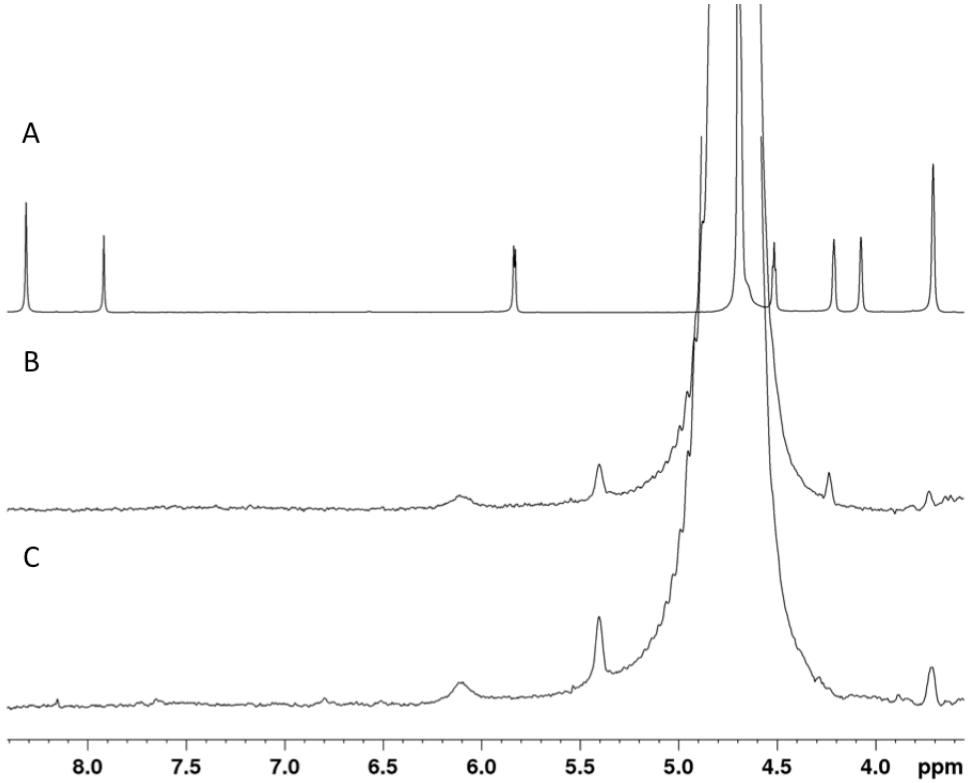


Figure 2.21 Overlay of (A) CPMG edited ^{1}H NMR reference spectrum of **AMP**; (B) CPMG edited ^{1}H NMR spectrum of mutant R416E *Sd*KefCTD protein (5.2 µM) after denaturation at 80 °C for 3 h; (C) CPMG edited ^{1}H NMR spectrum of mutant R416E *Sd*KefCTD protein (5.2 µM) before denaturation.

As R416E showed poor levels of expression, it was only possible to test it by CPMG edited ^{1}H NMR experiments at a final concentration of 5.2 µM. CPMG edited ^{1}H NMR experiments were thus performed on both non-denatured and denatured wild-type *Sd*KefCTD at the same final concentration of 5.2 µM. The reason for performing this experiment was to see whether the lack of signals corresponding to **AMP** when R416E was denatured could be attributed to the low concentration of the protein.

The non-denatured protein spectrum of wild-type *Sd*KefCTD at a final concentration of 5.2 µM showed no resonances corresponding to **AMP** (Figure 2.22 D). In the denatured spectrum of wild-type *Sd*KefCTD, however, sharp signals were observed corresponding to **AMP** (Figure 2.22 C). The denatured sample was spiked with 15 µM of **AMP** (Figure 2.22 B) to further

validate that the new peaks corresponded to **AMP** (Figure 2.22 A). Spiking the denatured sample with **AMP** enhanced the intensity of the new peaks, supporting the assignment of them to **AMP** (Figure 2.22 B). It was thus concluded that if **AMP** was present in R416E to the same extent that it is present in wild-type *Sd*KefCTD, it should have been observable when testing the protein at a final concentration of 5.2 µM.

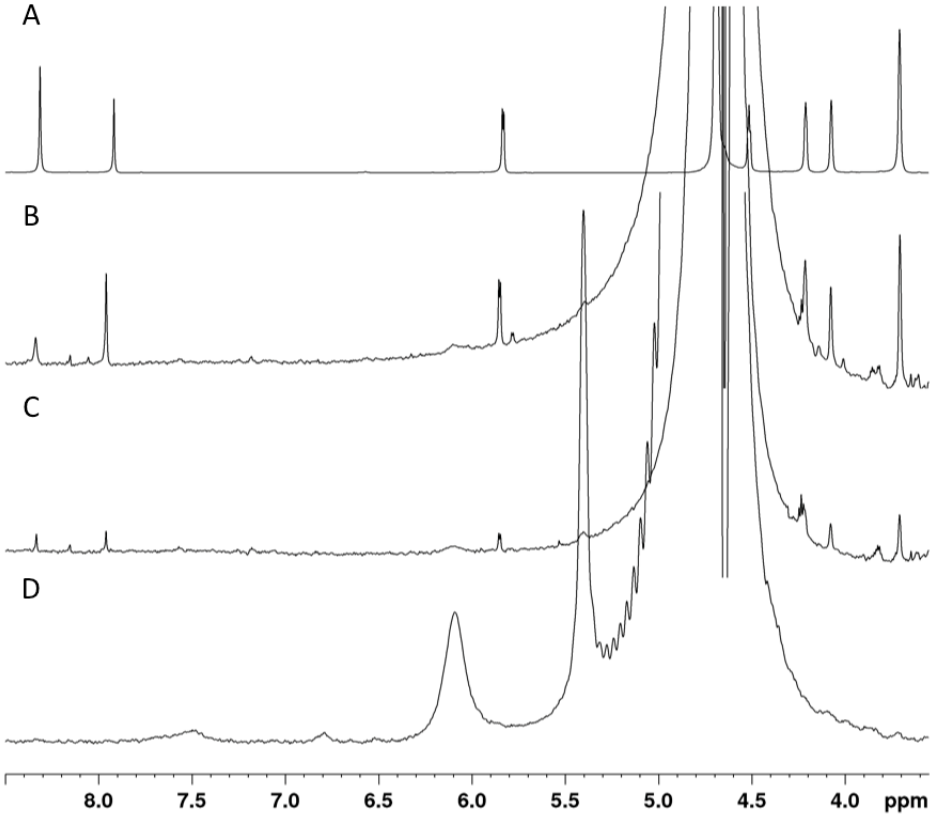


Figure 2.22 Overlay of (A) CPMG edited ^{1}H NMR reference spectrum of **AMP**; (B) CPMG edited ^{1}H NMR spectrum of wild-type *Sd*KefCTD protein (5.2 µM**)** after denaturation at 80 °C for 3 h spiked with **AMP** (15 µM**);** (C) CPMG edited ^{1}H NMR spectrum of wild-type *Sd*KefCTD protein (5.2 µM**)** after denaturation at 80 °C for 3 h; (D) CPMG edited ^{1}H NMR spectrum of wild-type *Sd*KefCTD protein (5.2 µM**)** before denaturation.

2.18 Establishing whether AMP is retained during the purification of the *Sd*KefCTD mutant R416E by analytical HPLC

To further verify that **AMP** was not present in the purified *Sd*KefCTD mutant R416E, the supernatant of the denatured *Sd*KefCTD mutant R416E was filtered through a 30 kDa cut-off concentrator and the filtrate was subjected to analysis by HPLC. The resulting spectrum was compared to an analytical HPLC trace of **AMP** that contained a peak with a retention time of 6.5 min (Figure 2.23 A). Dr. Anthony Chan in the Conway Group performed the expression and purification of the *Sd*KefCTD mutant R416E and the subsequent HPLC analysis.

Figure 2.23 B shows the analytical HPLC trace of the filtrate from the denatured sample of the *Sd*KefCTD mutant R416E that contains a peak with a retention time of 6.4 min. Spiking a sample of the filtrate of the denatured *Sd*KefCTD mutant R416E with the same concentration of **AMP** gave a single peak, suggesting that the peak could correspond to **AMP** (Figure 2.23 C). However, the peak in Figure 2.23 B is severely diminished in intensity when compared to the reference spectrum of **AMP** (the intensity of the peak in Figure 2.23 B is 34% the magnitude of the peak in Figure 2.23 A). If the *Sd*KefCTD mutant R416E does retain **AMP**, it clearly does not retain **AMP** as well as wild-type SdKefCTD, as demonstrated in the HPLC analysis in Figure 2.14.

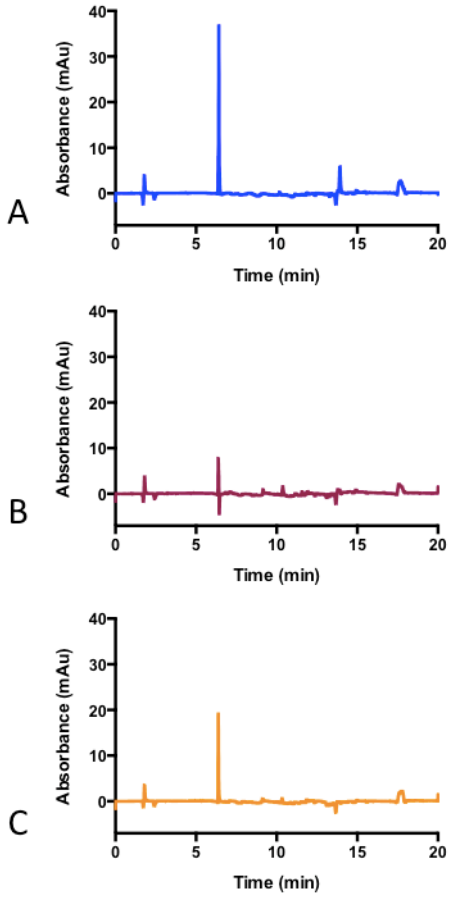


Figure 2.23 Analytical HPLC traces of (A) **AMP** (8.6 µM); (B) The supernatant of the *Sd*KefCTD mutant R416E (8.6 µM) denatured by heating to 95 °C for 5 min; (C) A spiking experiment, where the denatured *Sd*KefCTD mutant R416E (8.6 µM) and a pure sample of **AMP** (8.6 µM) were mixed and injected together.

2.19 Conclusions

A 2.9 Å resolution X-ray crystal structure of *Sd*KefCTD was resolved by our collaborator Dr. Christos Pliotas in the Naismith group, University of St Andrews. This X-ray crystal structure was found to have additional electron density in the nucleotide-binding pocket that was attributed to the nucleotide **AMP**. Evidence was gathered to support the modeling of **AMP** into the nucleotide-binding pocket notably: DSF analysis; CPMG edited ^{1}H NMR; analytical HPLC. The combination of these analytical techniques suggested that the bound nucleotide was **AMP**. Native mass spectrometry performed by Shane Chandler in the Benesch Group, University of Oxford, deduced that there are two **AMP** molecules bound per homodimer of *Sd*KefCTD.

Mutational analysis performed by Silvia Ekkerman in the Miller Group, University of Aberdeen, suggested that the key interactions of **AMP** with *Sd*KefCTD were hydrogen bonds between the hydroxyl groups on the ribose ring and D436, and a salt bridge formed between the phosphate group and R416. The importance of the phosphate group to **AMP**'s ability to stabilise *Sd*KefCTD was further investigated through DSF analysis of adenosine, which showed that removing the phosphate group resulted a loss in stabilisation of *Sd*KefCTD. The mutant R416E of *Sd*KefCTD was subjected to DSF, CPMG edited ^{1}H NMR and HPLC analysis to determine whether **AMP** was able to stabilise the R416E mutant and whether **AMP** was retained in the mutant during purification. It was found that **AMP** was no longer able to stabilise SdKefCTD and that AMP was not retained in the nucleotide-binding site when R416 was mutated to R416E.

Overall, the results suggest that **AMP** is likely to play a key role in stabilising *Sd*KefCTD, with the majority of the stability coming through the phosphate group. Furthermore DSF analysis of the ΔT_m of *Sd*KefCTD in the presence of both **AMP** and the glutathione S-conjugate **ESG** suggests that the nucleotide and peptide are able to bind simultaneously and independently of one another. The independence of the nucleotide- and glutathione-binding site is in accord with **AMP** having a structural or stabilising role in *Sd*KefCTD.

2.20 Chapter 2 References:

[1] J. Healy, S. Ekkerman, C. Pliotas, M. Richard, W. Bartlett, S. C. Grayer, G. M. Morris, S. Miller, I. R. Booth, S. J. Conway, et al., *Biochemistry* **2014**, *53*, 1982–1992.

[2] S. Miller, R. M. Douglas, P. Carter, I. R. Booth, *J. Biol. Chem.* **1997**, *272*, 24942–24947.

[3] T. P. Roosild, S. Castronovo, S. Miller, C. Li, T. Rasmussen, W. Bartlett, B. Gunasekera, S. Choe, I. R. Booth, *Structure* **2009**, *17*, 893–903.

[4] T. P. Roosild, S. Castronovo, J. Healy, S. Miller, C. Pliotas, T. Rasmussen, W. Bartlett, S. J. Conway, I. R. Booth, *PNAS* **2010**, *107*, 19784–19789.

[5] S. Miller, L. S. Ness, C. M. Wood, B. C. Fox, I. R. Booth, *J. Bacteriol.* **2000**, *182*, 6536–6540.

[6] Y. Jiang, A. Lee, J. Chen, M. Cadene, B. T. Chait, R. MacKinnon, *Nature* **2002**, *417*, 515–522.

[7] R. A. Albright, J.-L. V. Ibar, C. U. Kim, S. M. Gruner, J. H. Morais-Cabral, *Cell* **2006**, *126*, 1147–1159.

[8] T. P. Roosild, S. Miller, I. R. Booth, S. Choe, *Cell* **2002**, *109*, 781–791.

[9] The PyMOL Molecular Graphics System, Version 1.7.4 Schrodinger, LLC, **2010**.

[10] F. H. Niesen, H. Berglund, M. Vedadi, *Nat Protoc* **2007**, DOI 10.1038/nprot.2007.321.

[11] H. Y. Carr, E. M. Purcell, *Physical Review* **1954**.

[12] S. Meiboom, D. Gill, *Review of scientific instruments* **2004**.

[13] P. J. Hajduk, E. T. Olejniczak, S. W. Fesik, *J. Am. Chem. Soc.* **1997**, 119 12257–12261.

[14] F. D. L. Kondrat, W. B. Struwe, J. L. P. Benesch, *Methods Mol. Biol.* **2015**, *1261*, 349–371.

Chapter 3:

Lead identification and *in vitro* testing with a competition fluorescence assay and STD NMR

3. Lead identification and *in vitro* testing with a competition fluorescence assay and STD NMR

3.1 Introduction and aims

To identify novel, non-peptidic, small molecule probes of Kef, for use in the validation of Kef as an antibiotic target, cycles of computational screening, chemical synthesis, and biological evaluation were used. Three approaches were employed: 1) structure-based virtual screening; 2) ligand-based virtual screening; 3) chemical synthesis of inhibitors reported to bind to other glutathione-binding proteins. The first approach involved *in silico* virtual screening of industrial sponsor InhibOx's in-house database, Scopius-CSpace 6, which consists of 6.6 million commercially available drug-like compounds, against a homology model of the C-terminal domain of Kef from *Shewanella denitrificans* (*Sd*KefCTD). The *in silico* virtual screening was carried out using the protein-ligand docking programs AutoDock Vina[1] (henceforth referred to as Vina) and GOLD.[2] The second approach involved ligand-based virtual screening, using industrial collaborator InhibOx's proprietary program ElectroShape[3] to identify mimics of glutathione by comparing a virtual library of commercially available and 'drug-like' compounds to glutathione based on combinations of shape and charge. Ligand-based virtual screening uses the principle that ligands are often observed to bind with similar shapes to a protein. Therefore, mimicking the shape of a biologically active ligand may provide a starting point for developing drug-like compounds.[4-11] ElectroShape[3] is a program which comprises the Ultra-Fast Shape Recognition (USR) approach[12,13] and, Chiral Shape Recognition (CSR)[14], with the additional capability of incorporating electrostatic information as a fourth dimension into the search. USR works by computing a set of shape descriptors for a given query ligand, against which a library of molecules with pre-computed shape descriptors can be compared. CSR is an extension of USR that enables chirality to be accounted for by distinguishing enantiomers. The advantage of using

ElectroShape over USR and CSR is that it allows the electrostatic contributions and chirality of the binding modes to be considered, allowing for a more complete model than USR or CSR.

Both of the *in silico* structure-based approaches implemented individual docking of compounds using the programs AutoDock 4[15], Vina[1] and GOLD[2] to aid with the optimisation of hits identified from *in vitro* screening. AutoDock 4[15] is an automated docking program used for predicting the free energy of binding of ligands to their respective macromolecular targets, as well as, the conformation they adopt once bound.[16] It employs a Lamarckian genetic algorithm as its main approach to conformational and binding mode searching, and a semi-empirical free energy forcefield for predicting the free energy of binding.[15,17] Vina[1] is a molecular docking and virtual screening program produced by the same lab that developed AutoDock 4[15]. Vina[1] implements the Iterated Local Search global optimiser algorithm, which involves a succession of steps, each comprising a mutation and local optimisation using the Broyden-Fletcher-Goldfarb-Shanno method.[1] GOLD[2] (Genetic Optimisation for Ligand Docking) also implements a genetic algorithm to explore ligand conformational flexibility and binding modes.[2] The GOLD[2] scoring function of choice, for this study, is ChemPLP,[18] which has been demonstrated to be the most effective scoring function for pose prediction and virtual screening.[19]

In contrast, the third approach did not use *in silico* screening but instead involved the synthesis of small molecules that have been reported to inhibit other glutathione binding proteins. The rationale behind this was to see if their activity would translate to Kef, and thus provide starting points for the development of non-peptidic ligands of Kef. A small molecule inhibitor of *S*-nitrosoglutathione reductase (GSNOR), a receptor for the glutathione S-conjugate *S*-nitrosoglutathione (GSNO) was therefore synthesised and tested.[20]

3.2 Competition fluorescence assay and saturation transfer difference NMR

3.2.1 Competition fluorescence assay

Two biophysical techniques were employed in this chapter to validate the *in silico* leads. The first of these was a published competition fluorescence assay, developed by Dr. Jess Healy, which uses a fluorescent probe (a glutathione S-conjugate denominated **DNGSH**) that binds in the glutathione binding site.[21] This probe exhibits solvatochromic properties, such that transitions from hydrophobic to hydrophilic environments cause a decrease in the intensity of emission (Figure 3.1). These properties can be exploited to identify compounds that compete with the probe for the glutathione-binding site. A decrease in the fluorescence intensity of the probe upon introduction of a competing ligand was taken as an indication of binding, due to the probe being displaced from the hydrophobic site of the protein surface to the hydrophilic environment of the buffer. The extent of displacement can be represented by the ratio F_B/F_L, where F_B is the fluorescence intensity of the probe in the absence of a competing ligand and F_L is its intensity in the presence of a competing ligand (Figure 3.1). By subtracting one from this ratio ($F_B/F_L - 1$), a qualitative indication of the strength of binding of a competing ligand can be obtained, with ($F_B/F_L - 1$) = 0 suggesting no binding and non-zero values of ($F_B/F_L - 1$) giving an indication of the strength of binding, with larger values implying tighter binding.

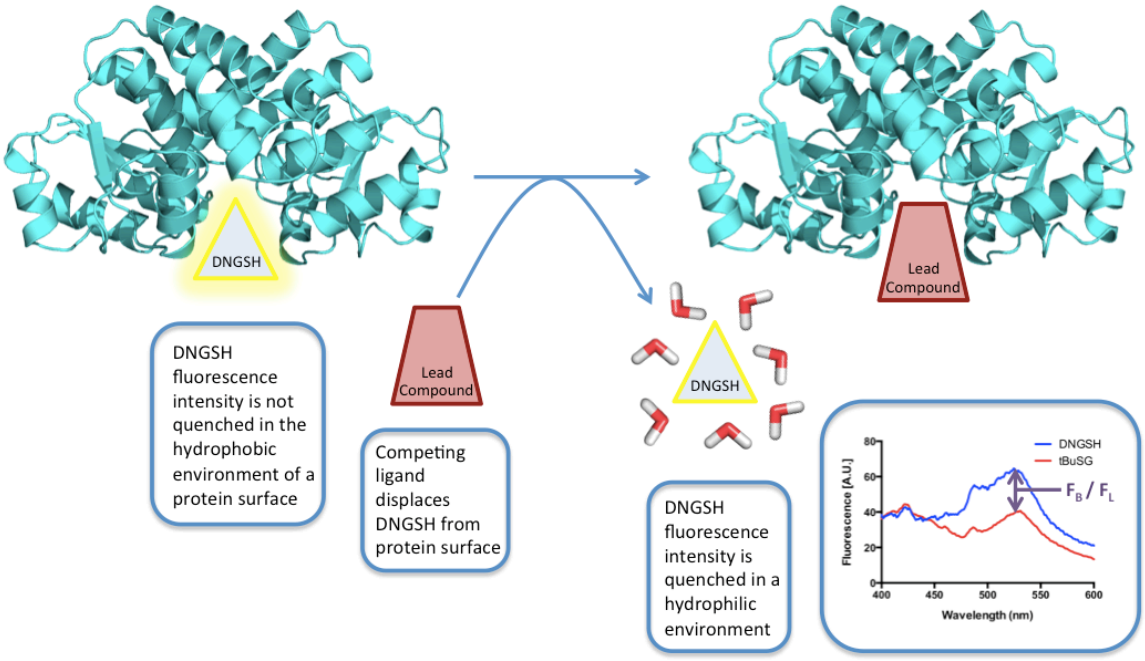
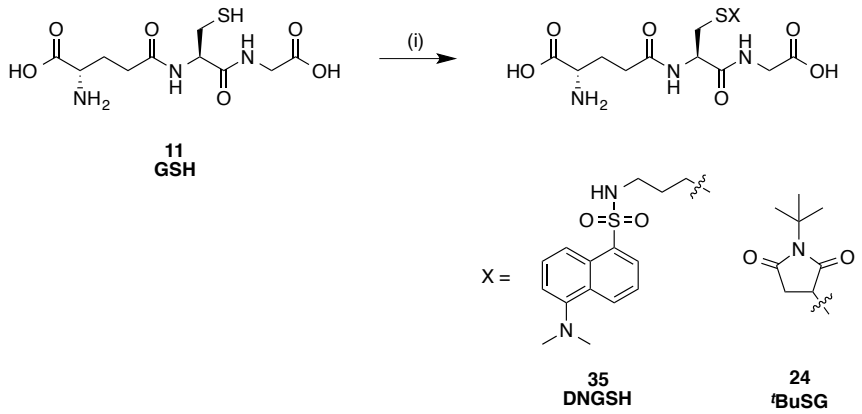


Figure 3.1 Cartoon representation of the competition fluorescence assay. The graph shows the emission spectra for *Sd*KefCTD (6 μM) and **DNGSH** (5 μM) on their own (blue) and in the presence of the glutathione S-conjugate **^{t}BuSG** (1 mM; red) at 340 nm excitation.

3.2.1.1 Synthesis of fluorescent probe DNGSH and positive control ^{t}BuSG

The fluorescent probe, *S*-((5-(dimethylamino)naphthalene-1-yl)sulfonylaminopropyl) glutathione (**DNGSH**; **35**) was synthesised using a photochemical thiol-ene reaction[7] to couple glutathione with *N*-allyl-5-(dimethylamino)naphthalene-1-sulfonamide. A glutathione S-conjugate of known affinity was also synthesised to act as a positive control in the competition fluorescence assay. The glutathione S-conjugate selected was *S*-(*N*-*tert*-butylsuccinimido) glutathione (**^{t}BuSG**; **24**), which has a reported K_D of 400 ± 200 nM for *Sd*KefCTD (measured by fluorescence emission spectra)[21] and was synthesised by conjugate addition of glutathione to *N*-*tert*-butylmaleimide (Scheme 3.1).



Scheme 3.1 Synthesis of *S*-((5-(dimethylamino)naphthalene-1-yl)sulfonylaminopropyl) glutathione, **35**, and *S*-(*N*-*tert*-butylsuccinimido) glutathione, **24**. *Reagents and conditions*: (i) For **35**: *N*-Allyl-5-(dimethylamino)naphthalene-1-sulfonamide, TCEP·HCl, 2,2-dimethoxyphenyl acetophenone, THF/H_2O, RT, hv, 5 h, 11%; For **24**: *N*-*tert*-butylmaleimide, NaOH, RT, 20 min, 86%.

3.2.2 Saturation transfer difference NMR

All saturation transfer difference NMR studies were initially performed by Dr. Jess Healy and then subsequently by Dr. Sonia Diab.

The second technique used to validate the *in silico* leads was saturation transfer difference NMR (STD NMR),[22,23] which exploits spin diffusion and the nuclear Overhauser effect to identify ligands that bind to a protein of interest. When a protein is selectively magnetically saturated, this saturation is only transferred to a ligand *via* the aforementioned effects if it binds. Saturation of bound ligands results in a decrease in their signal intensity, and so subtraction of a spectrum taken when the protein is selectively saturated from a reference spectrum, when there is no saturation, leaves only the signals for ligands that bind. Further application of these principles potentially allows the intensity of the signals to be used to determine both the mode and magnitude of binding.[22,23]

3.3 *In silico* screening - protein-ligand docking

Before commencing protein-ligand docking, a homologue of Kef was selected to act as a suitable model system. The most commonly studied homologue, *Escherichia coli* KefC (*Ec*KefC), was not an ideal candidate as it requires an additional soluble ancillary protein, KefF, for full activity, which has been found to complicate purification and biophysical screening (Figure 3.2 A).[24,25] The homologue Kef from *Shewanella denitrificans* (*Sd*Kef), however, is not dependent upon an ancillary protein for full activity, and so was identified as an alternative for the biophysical studies of the Kef system.

At the time the protein-ligand docking studies were performed, there was not a crystal structure of Kef from *Shewanella denitrificans*, so industrial supervisor Dr. Garrett Morris constructed a homology model of the C-terminal domain of Kef from *Shewanella denitrificans* (Hm1*Sd*KefCTD) on which the docking studies could be performed (Figure 3.2 B). A homology model is a prediction of the three-dimensional structure of a protein based on the concept and observation that the tertiary structures of homologous proteins are more conserved during evolution than their primary structures.[26] The tertiary structures of a protein can therefore be used to predict the tertiary structures of homologous proteins with similar primary structures to it.[26] Hm1*Sd*KefCTD was constructed using the X-ray crystal structure obtained by Roosild *et al.* of the *E. coli* KefC C-terminal domain (*Ec*KefCCTD) in complex with glutathione (PDB code: 3L9W; Figure 3.2 A)[27] as a template for the *Sd*KefCTD sequence from GenBank (ABE53663.1).

It is worth noting at this point that as Hm1*Sd*KefCTD was constructed using the conformation adopted by *Ec*KefCCTD when glutathione is binding as a template (PDB code: 3L9W),[27] the *in silico* virtual screening was likely to be biased towards identifying inhibitors of *Sd*Kef, rather than activators. Furthermore it was not known at the time of the protein-ligand docking that

glutathione has an affinity of 900 µM (fluorescence emission spectra).[21] This meant that the *in silico* virtual screening performed using Hm1*Sd*KefCTD was likely to be limited to the identification of low affinity leads.

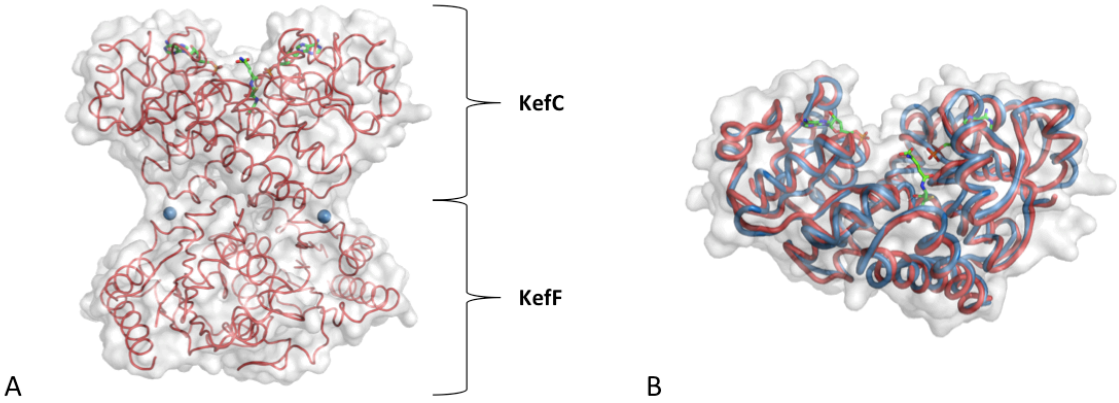


Figure 3.2 Cartoon loop and surface representations of (A) The X-ray crystal structure of *Ec*KefCCTD in complex with KefF (PDB code: 3L9W); (B) Homology model of *S. denitrificans* C-terminal domain Kef (Hm1*Sd*KefCTD; blue), constructed by industrial supervisor Dr. Garrett Morris, superimposed onto *Ec*KefCCTD (PDB code: 3L9W, red). The *Ec*KefCCTD molecular surface is depicted in grey. Images were constructed using PyMOL.[28]

3.4 Water placement in homology models

With a homology model in hand, water placement into the model was carefully considered, as the omission of structural water molecules during protein-ligand docking can result in *in silico* binding mode predictions that are drastically different to reality.[29] Three approaches were taken to water placement: 1) copy water positions over from the X-ray crystal structure 3L9W,[27] and perform an energy minimisation, deleting waters that clash with residues; 2) use the program WaterDock[30] to predict the binding sites of waters and their probabilities of being displaced (Figure 3.3 A); 3) omit waters. Models containing each of these approaches are given the abbreviations Hm1*Sd*KefCTDW1, Hm1*Sd*KefCTDW2 and Hm1*Sd*KefCTDW3 respectively. Industrial supervisor Dr. Garrett Morris constructed Hm1*Sd*KefCTDW1, and I constructed Hm1*Sd*KefCTDW2 and Hm1*Sd*KefCTDW3.

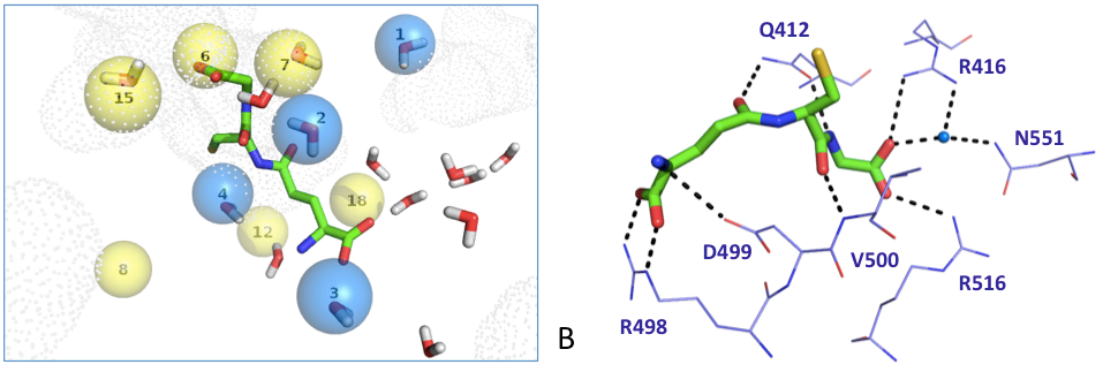


Figure 3.3 (A) Water binding positions predicted by WaterDock 2.1 overlaid with water molecules taken from 3L9W. Water molecules predicted by WaterDock 2.1 are shown as spheres. Blue spheres depict water molecules with *P(D)* < 0.2 and yellow spheres depict water molecules with 0.2 < *P(D)* < 0.4, where *P(D)* is the probability that a water molecule is displaced. Water molecules with a *P(D)* < 0.2 were determined to be conserved and included in the modified model, Hm1*Sd*KefCTDW2, and water molecules with a *P(D)* > 0.2 were excluded. Water molecules that were copied over from 3L9W into homology model Hm1*Sd*KefCTDW1 are shown as sticks for comparison. Glutathione is overlaid as sticks (carbon = green). (B) The key residues of *Ec*KefCCTD that interact with the peptidic core of glutathione and glutathione S-conjugates. Glutathione is shown as sticks (carbon = green) and the mediating water is shown as a blue sphere. Images were constructed using PyMOL.[28]

Hm1*Sd*KefCTDW2 was constructed using WaterDock 2.1, an updated version of Ross *et al.*'s WaterDock[30], which uses Vina[1] to predict the binding sites of waters. Having predicted the location of waters in Hm1*Sd*KefCTD, Ross *et. al.*'s Water Classifier[30] was used to calculate the probability of a ligand displacing each water molecule (Figure 3.3 A). The threshold for conserving a water molecule in the model (shown as blue spheres in Figure 3.3 A) was arbitrarily assigned to be *P(D)* < 0.2, where *P(D)* is the probability of displacement. Analysis of the predicted water molecules suggested that none of them would act as hydrogen bonding mediators between residues and the glutathione Gly-carboxylate, as observed in the *Ec*KefCCTD X-ray crystal structure (PDB code: 3L9W; Figure 3.3 B).[27] Water 6 in Figure 3.3 A is near to where the mediating water in the *Ec*KefCCTD X-ray crystal structure (PDB code: 3L9W)[27] would be (Figure 3.3 B), however, the water classifier did not predict it to be conserved, and it sterically clashes with the glutathione Gly-carboxylate when the bound conformation of glutathione is overlaid

into the model. Hm1*Sd*KefCTDW2 was therefore constructed containing waters with *P(D)* < 0.2. Waters classified within 0.2 < *P(D)* < 0.4 were omitted during docking but subsequently overlaid to identify ligands that could potentially displace them.

3.5 Validation of the docking methods

Work commenced with a validation step to determine whether AutoDock 4,[15] Vina[1] and GOLD[2] were suitable docking programs for use with the Kef system. It is standard practice to perform this validation step by docking a database of known actives and inactives into the binding site in question and using a receiver operating characteristic (ROC) curve to compare the number of true-positives identified by the docking algorithm to the number of false-positives or "decoys".[31] However, to the best of our knowledge, at the time the docking studies were performed, there were no published decoys for glutathione-binding sites so an alternative approach was pursued. Validation was instead performed by docking glutathione back into the glutathione-binding site of X-ray crystal structure 3L9W[27] to determine whether the docking methods could replicate the crystallographically-observed conformation adopted by glutathione in 3L9W[27] (Figure 3.4 A). The docking was performed under identical conditions to those used to prepare the crystal structure; most notably the residues of 3L9W[27] and glutathione were protonated at pH 5.5. GOLD[2] managed to emulate the conformation of glutathione adopted in the X-ray crystal structure (Figure 3.4 B), with no alteration of parameters required. AutoDock 4[15] and Vina,[1] on the other hand, required the optimisation of two of parameters in order to replicate the conformation adopted by glutathione in 3L9W.[27] For AutoDock 4,[15] this involved increasing the number of energy evaluations (evals) in the genetic algorithm from 250,000 to 2,500,000, and the number of genetic algorithm runs from 10 to 100 (Figure 3.4 C). For Vina,[1] an increase in the exhaustiveness from 8 to 20, and the number of modes from 9 to 20 was required (Figure 3.4 D).

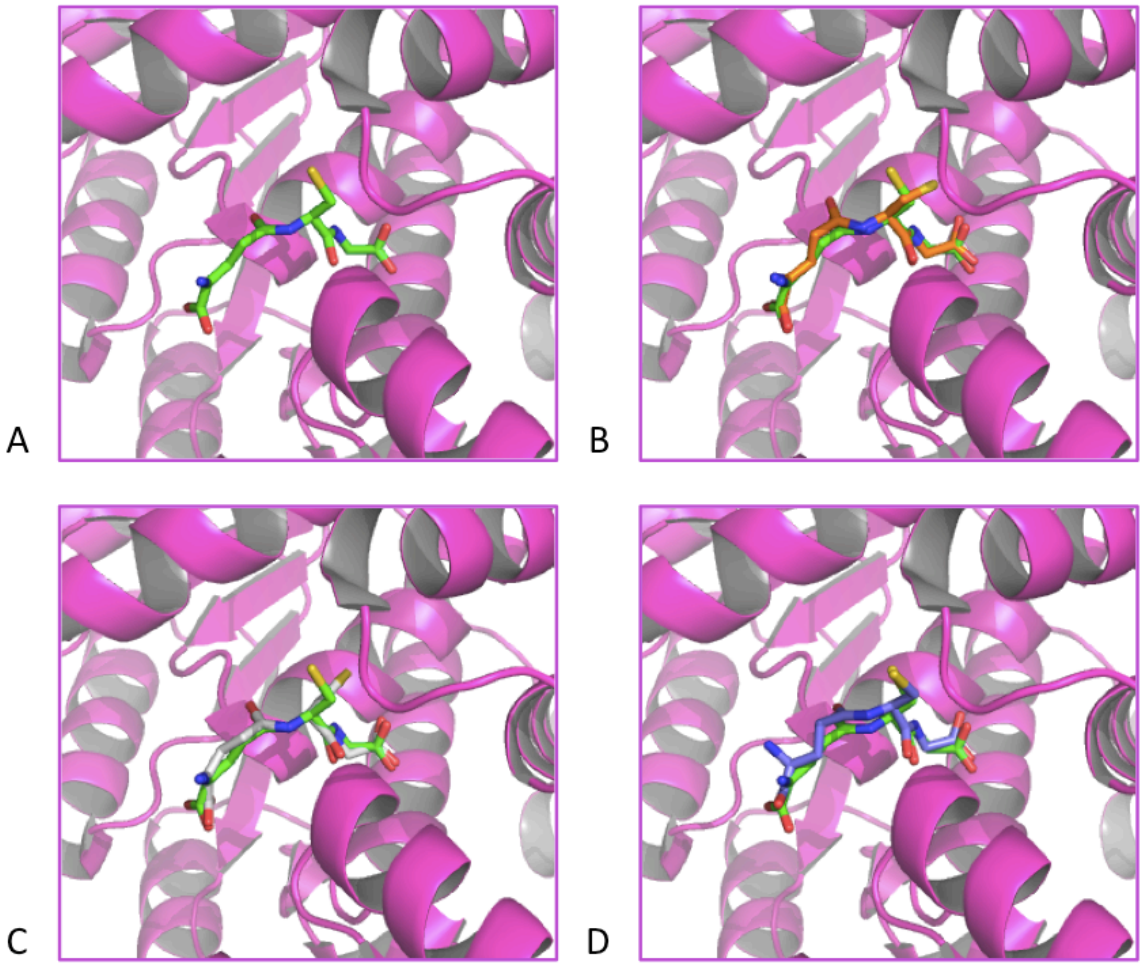


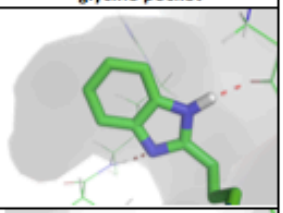
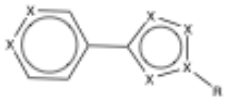
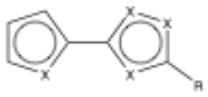
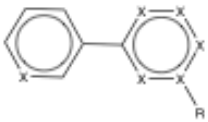
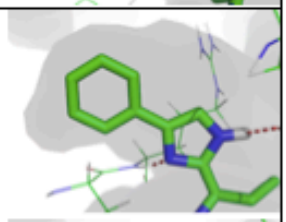
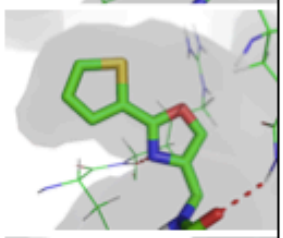
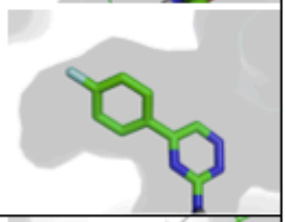
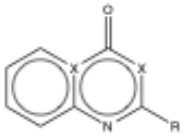
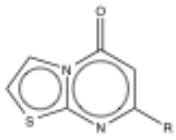
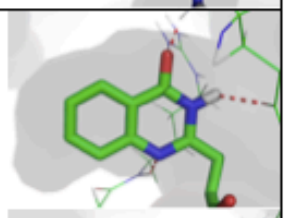
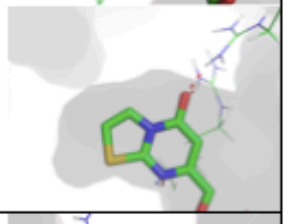
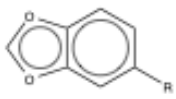
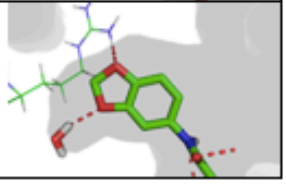
Figure 3.4 Validation of AutoDock 4, Vina and GOLD for use with the glutathione-binding site of Kef systems. Conformations of glutathione (A) adopted in *Ec*KefCCTD 3L9W (carbon = green; experimentally observed binding mode); (B) generated by GOLD (carbon = orange); (C) generated by AutoDock 4 (carbon = white); (D) generated by Vina (carbon = purple); (B), (C) and (D) have the conformation adopted by glutathione in 3L9W (carbon = green) overlaid for ease of comparison. Images were constructed using PyMOL.[28]

3.6 High throughput protein-ligand docking

As we were satisfied that the docking programs could be applied to Kef systems, and their parameters had been suitably optimised, high throughput protein-ligand docking was performed with Vina[1] and GOLD[2] on the homology models Hm1*Sd*KefCTD using a 96-core cluster. GOLD[2] was used to screen Hm1*Sd*KefCTD with all three variations of the water placements highlighted in Section 3.4 (Hm1*Sd*KefCTDW1-3), and Vina[1] was used to screen Hm1*Sd*KefCTDW2. The screening was performed using 661 diversity subsets of InhibOx's in-house database Scopius-Cspace 6, containing a total of 661,000 drug-like and commercially available compounds. It should be noted that in order to speed up the screening process for Vina,[1] it was required that the exhaustiveness was reduced to 2 in this instance.

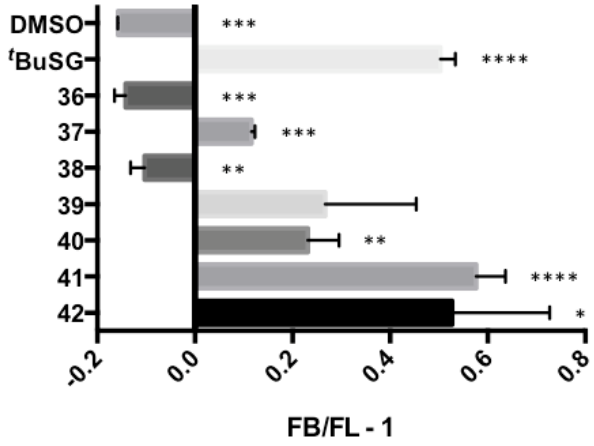
The docked compounds were ranked by their docking scores (ChemPLP[18] for GOLD;[2] affinity for Vina[1]) and the top 100 hits in the screen using Hm1*Sd*KefCTDW1 were compared to the top 100 hits in the screens using Hm1*Sd*KefCTDW2 and Hm1*Sd*KefCTDW3. It was found that the glutathione-Gly binding pocket was occupied by much more lipophilic scaffolds in the screens using Hm1*Sd*KefCTDW2 and Hm1*Sd*KefCTDW3. A manual curation of these lipophilic scaffolds found that they could be sorted into groups based on their structures, as highlighted in Table 3.1. Subsequent overlaying of the WaterDock 2.1 waters with a probability of displacement of $0.2 < P(D) < 0.4$ found that the majority of these lipophilic scaffolds were predicted to displace water 7 in Figure 3.3 A. Displacement of an ordered water molecule into the bulk solvent can result in a favourable entropic gain due to an increase in the water molecule's translational and orientational degrees of freedom.[30] Therefore, if water 7 is a fair prediction of an ordered water molecule in *Sd*KefCTD, ligands predicted to displace it might demonstrate a reasonable affinity for *Sd*KefCTD through entropic gains.

Table 3.1 Frequently observed substituents in the glutathione-Gly pocket from the high-throughput protein-ligand docking. Displayed are the chemical structures of substituents in the glutathione-Gly pocket along with the frequency of occurrence of the top 100 hits of each virtual screen (shown as a percentage of the top 100 hits). The virtual screens were performed by GOLD and Vina using Hm1*Sd*KefCTD with WaterDock 2.1 waters incorporated into it (Hm1*Sd*KefCTDW2) or with no waters (Hm1*Sd*KefCTDW3).

General Chemical Structure of Substituent	GOLD screen Hm1*Sd*KefCTDW2 %	Vina screen Hm1*Sd*KefCTDW2 %	GOLD screen Hm1*Sd*KefCTDW3 %	Predicted binding mode of an example substituent in GSH-glycine pocket
 X = C, N, O or S	34	9	16	
   X = C, N, O or S	5	23	18	  
  X = C or N	0	0	12	 
	0	11	0	

Ten scaffolds were selected for purchase, with a representative from each of the groups highlighted in Table 3.1, and were tested with assistance from Dr. Jess Healy using the competition fluorescence assay (Figure 3.5).

A



B

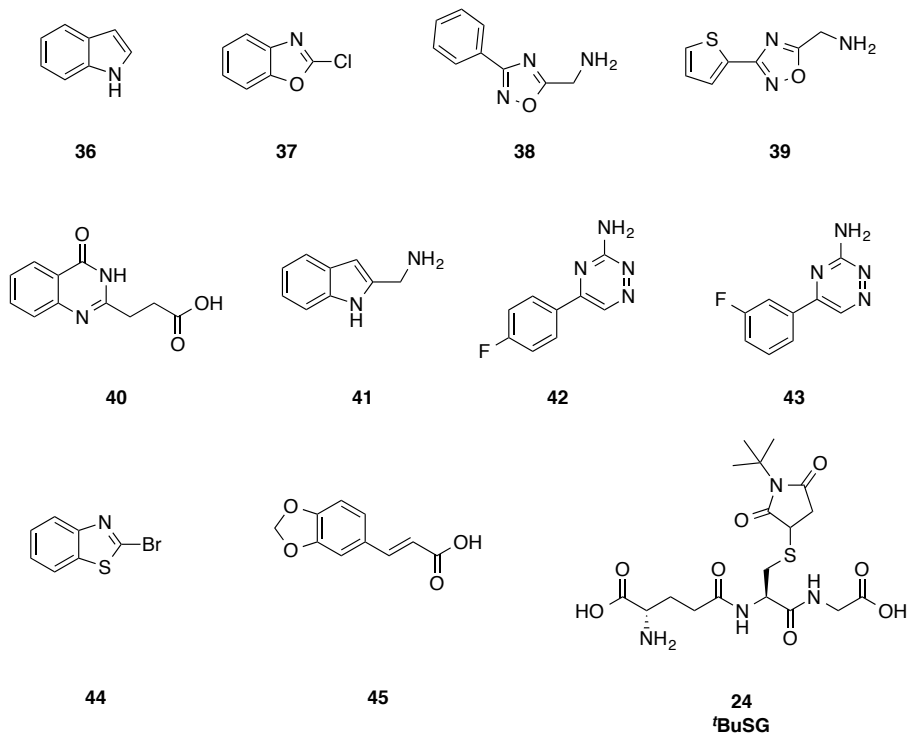


Figure 3.5 Lead compounds identified using high throughput protein-ligand docking. (A) A graph showing (F_B/F_L – 1) values of each compound (at a concentration of 5 mM, with the exception of **42**, which was tested at 1 mM due to low solubility) relative to the positive control **^{t}BuSG** (5 mM) and negative control, **DMSO**, at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3, with the exception of **37**, where n = 2). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) Structures of the compounds tested.

Due to low solubility, it was not possible to test **43**, **44** or **45**, however, four promising leads were identified from this screen: **39**; **40**; **41**; and **42**. Compounds **40**, **41** and **42** were progressed to STD NMR studies, and these compounds were verified to bind using this assay (Figure 3.6). Considering their low molecular weights and the magnitude of their (F_B/F_L – 1) values relative to **^{t}BuSG**, these compounds appeared to show great promise for the development of non-peptidic probes, and indeed represent good examples of the first stage of fragment-based inhibitor design.

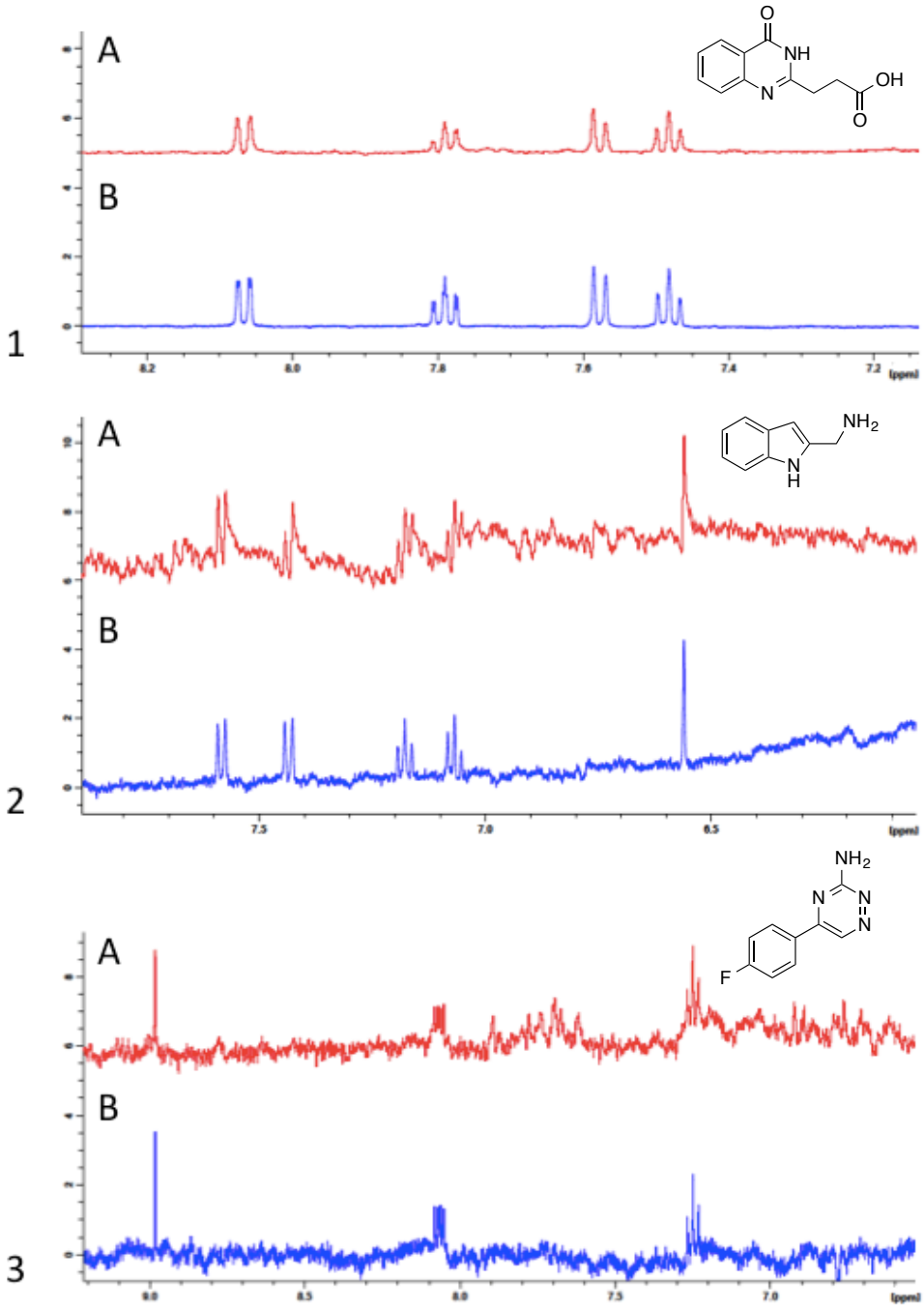


Figure 3.6 STD NMR spectrum for compounds: (1) **40**; (2) **41**; (3) **42**. In each case, (A) is the difference spectrum (spectrum before saturation minus the saturated spectrum); (B) is the spectrum before saturation.

3.7 Synthesis and derivatisation of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine

5-(4-Fluorophenyl)-1,2,4-triazin-3-amine **42** was focused on first, due to its synthetic tractability and the potential that it showed for Structure-Activity-Relationship (SAR) studies. Individual docking of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, into Hm1*Sd*KefCTDW1 using Vina[1] predicted that nitrogen 1 of the 1,2,4-triazine ring could be forming a polar contact with the carbonyl of the key residue Q419 (Q419 is the equivalent residue to Q412 in *Ec*KefCCTD), nitrogen 4 of the 1,2,4-triazine ring may form an electrostatic interaction with a water molecule, with the aniline hydrogen bonding to Q441 (Figure 3.7 A). Comparison of this binding mode to that of glutathione (Figure 3.7 B) suggested that there was scope for synthetic optimisation through altering the 5-aryl substituent to pick up some of the interactions predicted to be made by the glutathione Gly-carboxylate with R423, R507 and R523 (R423 and R523 are the equivalent residues to R416 and R516 in *Ec*KefCCTD). Furthermore, there was the possibility of picking up interactions in the glutathione-Glu region of the binding site through building off the aniline of the 1,2,4-triazine ring of compound **42** (Figure 3.7 C).

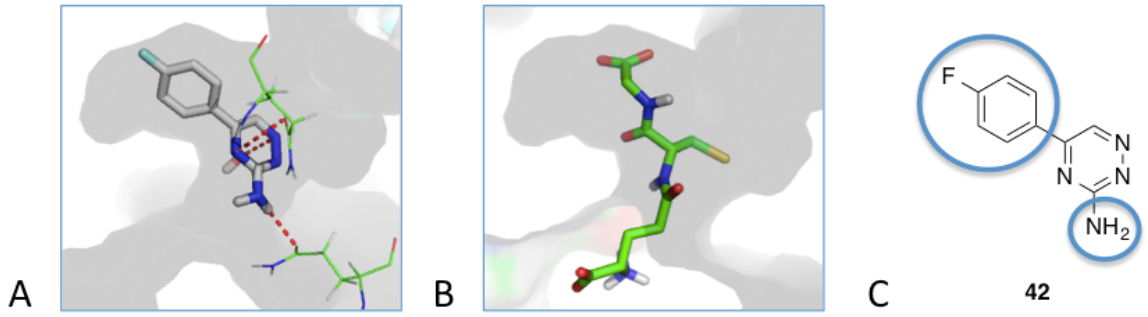
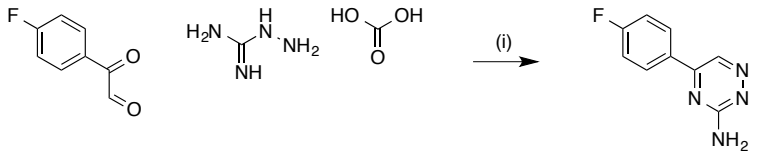


Figure 3.7 Options for optimisation of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**. (A) Demonstrates the predicted binding mode of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, that has been docked in the homology model Hm1*Sd*KefCTDW1 using Vina; (B) Demonstrates the binding mode of glutathione in the homology model Hm1*Sd*KefCTDW1; (C) Highlights positions for SAR development of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**. Images were constructed using PyMOL.[28]

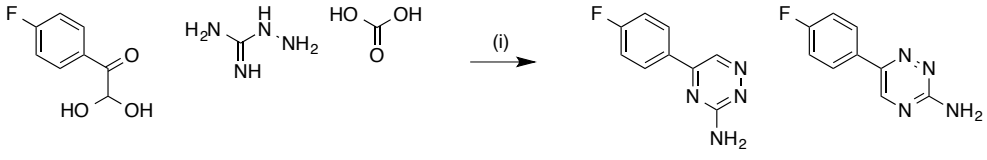
Before synthesising analogues, a synthetic route to 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, was established. A literature search highlighted a one step synthesis of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, through a condensation reaction of 4-fluorophenylglyoxal, **46**, with aminoguanidine bicarbonate, **47**, in aqueous ethanol to give 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, in 70% yield (Scheme 3.2).[32]



46 **47** **42**

Scheme 3.2 Reported synthesis of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine **42**. *Reagents and conditions*: (i) Aqueous ethanol, 25 °C, 22 h, 50 °C, 1 h, 70%.[32]

4-Fluorophenylglyoxal was only found to be commercially available as a hydrate, **48**, which altered the regioselectivity of the reaction significantly. The hydrated aldehyde was no longer the most electrophilic site, causing the more nucleophilic amino nitrogen of aminoguanidine to attack the ketone instead, favouring the formation of the undesired regioisomer 6-(4-fluorophenyl)-1,2,4-triazin-3-amine, **49**, over 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, (Scheme 3.3).



48 **47** **42** **49**

Scheme 3.3 Attempted synthesis of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, using 4-fluorophenylglyoxal hydrate, **48**. *Reagents and conditions*: (i) Aqueous ethanol, 25 °C, 22 h, 50 °C, 1 h, **49** isolated in 36% yield.

The regioisomers were isolated, and distinguished using ^{1}H-^{13}C HMBC NMR, as the proton H_A in Figure 3.8 can only couple to one carbon atom in 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**,

(Figure 3.8 A), whereas it can couple to two carbon atoms in 6-(4-fluorophenyl)-1,2,4-triazin-3-amine, **49**, (Figure 3.8 B). A crude ^{1}H NMR of the reaction mixture found that the ratio of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, to 6-(4-fluorophenyl)-1,2,4-triazin-3-amine, **49**, was 1:3.

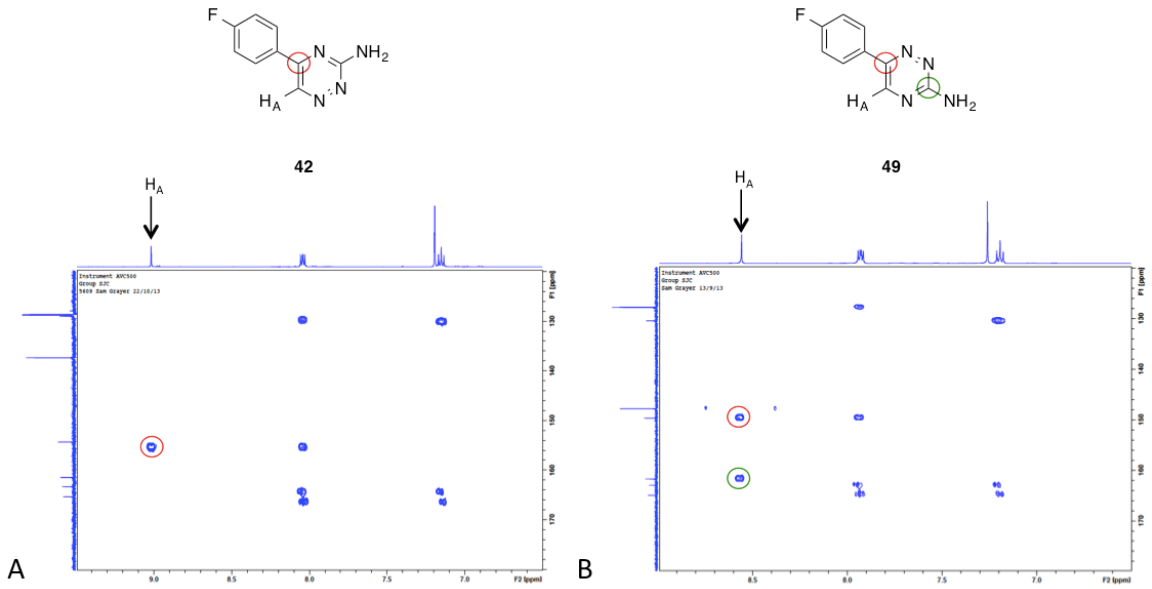
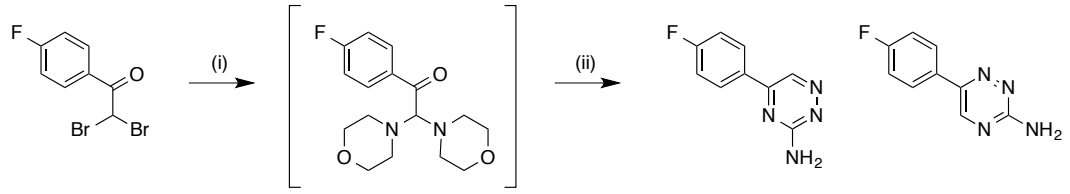


Figure 3.8 ^{1}H-^{13}C HMBC spectra for (A) 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, and (B) 6-(4-fluorophenyl)-1,2,4-triazin-3-amine, **49**. The cross peaks corresponding to the carbons that couple to the proton H_A are circled on each spectra. Each circle is colour coded according to the carbon that the cross peaks correspond to.

A further literature search identified a paper whose authors, Limanto *et al.*, had experienced the same problem with regioselectivity when synthesising 5-substituted-3-amino-1,2,4-triazines from aminoguanidine and commercially available glyoxal hydrates.[33] Limanto *et al.* engineered an elegant solution, where α,α-dibromoketones were reacted with morpholine to form ketoaminals that could subsequently be reacted with aminoguanidine in the presence of acetic acid in methanol to form 5-substituted-3-amino-1,2,4-triazines. Treatment of the ketoaminal (for example **51** in Scheme 3.4) with acid forms an imine, enhancing the difference in reactivity between the two electrophilic sites, favouring the addition of the more nucleophilic amino

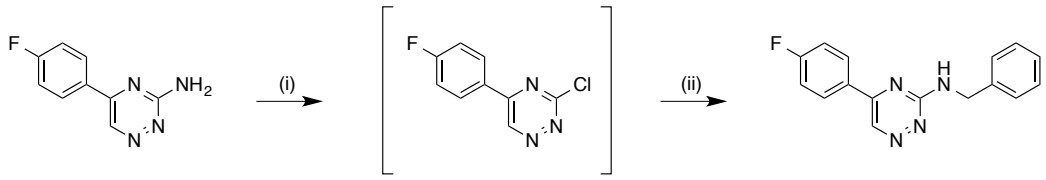
nitrogen of aminoguanidine into the imine, and the less nucleophilic guanidine nitrogen into the ketone. This approach was applied to the synthesis of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, and a crude ^{1}H NMR of the resulting reaction mixture indicated that the regioselectivity was improved, with a ratio of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, to 6-(4-fluorophenyl)-1,2,4-triazin-3-amine, **49**, of 4:1 (Scheme 3.4).



50 **51** **42** **49**

Scheme 3.4 Synthesis of 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, using 2,2-dibromo-4'-fluoroacetophenone, **50**. *Reagents and conditions*: (i) Morpholine, THF, 50 °C, 24 h; (ii) Aminoguanidine bicarbonate, glacial acetic acid, RT, 2 h, 80 °C, 14 h.

Having successfully synthesised 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, the synthetic options available for probing the SAR of the aniline were explored. It was proposed that one way to introduce a broad range of substituents at the aniline position was through S_NAr substitutions of alkyl- and aryl- amines into 3-chloro-5-(4-fluorophenyl)-1,2,4-triazine, **52**. Synthesis of the proposed building block 3-chloro-5-(4-fluorophenyl)-1,2,4-triazine, **52**, was achieved by performing a Sandmeyer reaction on 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, using conditions reported to work on 5-substituted-3-amino-1,2,4-triazines.[34] As a proof of concept for this approach to derivatisation, benzylamine was coupled to 3-chloro-5-(4-fluorophenyl)-1,2,4-triazine, **52**, to form *N*-benzyl-5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **53**, (Scheme 3.5).



42 **52** **53**

Scheme 3.5 Synthesis of *N*-benzyl-5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **53**, *via* an S_NAr reaction of benzylamine with 3-chloro-5-(4-fluorophenyl)-1,2,4-triazine, **52**. *Reagents and conditions*: (i) ^tBuONO, $CuCl_2$, MeCN, CH_2Cl_2, 65 °C, 3 h; (ii) Benzylamine, DIPEA, MeCN, 85 °C, 15.5 h, 29% (over two steps).

Unfortunately, the poor solubility of *N*-benzyl-5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **53**, meant that it was not possible to test this compound using either the competition fluorescence assay or STD NMR. Furthermore, as will be highlighted in Section 3.12, 5-(4-fluorophenyl)-1,2,4-triazin-3-amine, **42**, was identified to be a false positive and so further work on this series was terminated.

3.8 Ligand-based virtual screening and halogen bonding

Prior to starting my DPhil, InhibOx performed an ElectroShape search to identify a preliminary set of non-peptidic ligands that mimicked the shape of glutathione. My industrial supervisor at InhibOx, Dr. Garrett Morris, performed this search by screening 5,389,545 commercially available compounds from InhibOx's in-house database, Scopius-CSpace using ElectroShape, to find molecules with similar shapes, chirality and electrostatic properties to the conformation adopted by glutathione in the X-ray crystal structure of *Ec*KefCCTD (PDB code: 3L9W).[27] After ranking and filtering the top 5000 ElectroShape hits, using COver (a proprietary InhibOx tool that compares the shapes of molecules using Gaussians superimposed on their atoms) and CCG MOE, 222 molecules were selected as potential leads. Dr. Jess Healy and Niall Igoe bought or synthesised some of these compounds and tested them using the competition fluorescence assay. Figure 3.9 highlights a lead compound that they identified, **54**, in which a sulfonamide bioisostere is

predicted to mimic the carboxylic acid of Gly, and an unorthodox iodo substituent is predicted to mimic the thiol of Cys.

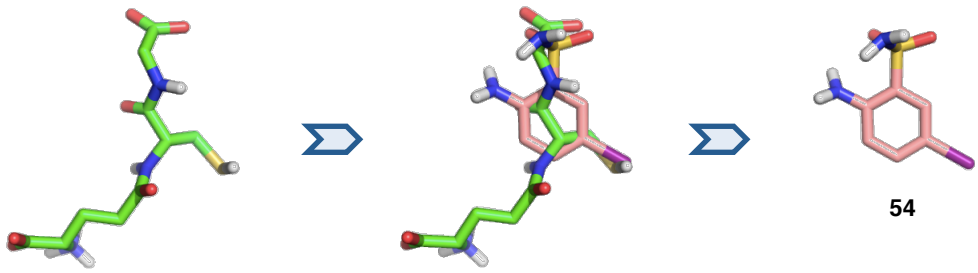


Figure 3.9 Example of a lead compound identified by the ElectroShape screen. Compound **54** mimics the carboxylic acid of glutathione's Gly with a sulfonamide bioisostere and the Cys-thiol with an iodo substituent (shown in purple). Images were constructed using PyMOL.[28]

In addition to this compound, the regioisomers **55** and **56a** were synthesised and screened by Dr. Jess Healy. All three of these were found to bind in the competition fluorescence assay, however, compound **56a** appeared to displace the fluorescent probe significantly more than **54** and **55**, suggesting a much higher affinity (Figure 3.10). Compound **56a** was, therefore, progressed and tested using STD NMR, which also suggested that **56a** was binding to *Sd*KefCTD (Figure 3.11).

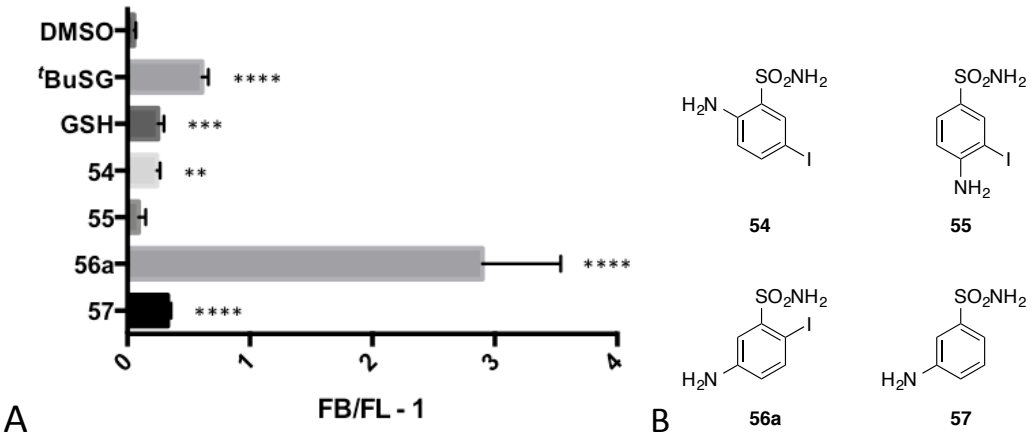


Figure 3.10 Analysis of the regioisomers of **54** binding to *Sd*KefCTD using the competition fluorescence assay. (A) Shows (F_B/F_L – 1) values of each compound (5 mM) relative to the positive control **^{t}BuSG** (5 mM) and negative control **DMSO** at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3, with the exception of **54**, where n = 2). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) Structures of the compounds tested.

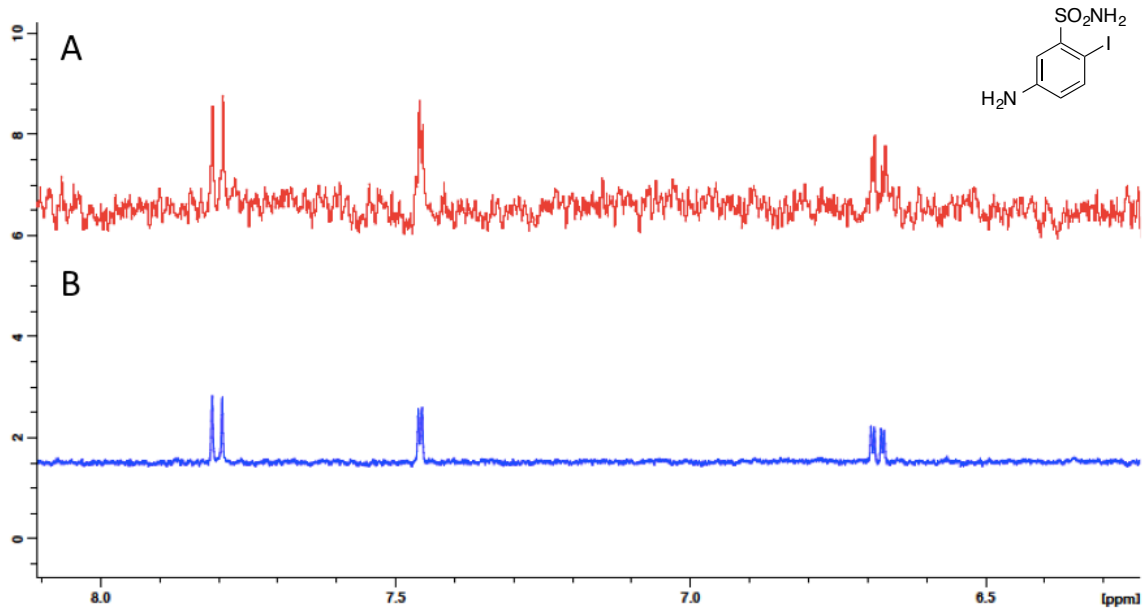


Figure 3.11 STD NMR spectrum for compound **56a**. (A) is the difference spectrum (spectrum before saturation minus the saturated spectrum); (B) is the spectrum before saturation.

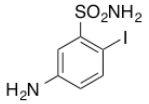
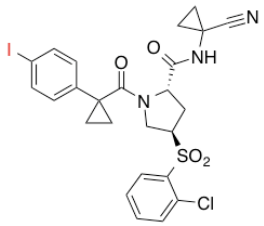
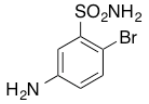
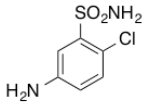
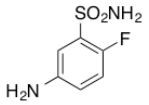
Comparison of the (F_B/F_L − 1) value of **56a** to 3-aminobenzene-1-sulfonamide, **57**, indicated a significant drop in binding affinity, suggesting that the iodo substituent was making a contribution to binding (Figure 3.10). It was hypothesised that a halogen bond to the protein

could explain the role of the iodo substituent. A halogen bond is a non-covalent interaction between a σ-hole, a region of positive electrostatic potential on a covalently bonded halogen, and a nucleophilic site.[35] The strength of a halogen bond is determined by the size of the σ-hole, which increases as the polarisability of the halogen increases. The significance of investigating the proposed hypothesis was that halogen bonding is a relatively new concept in the field of medicinal chemistry and examples of probes that computational chemists can use to develop models of halogen bonding to proteins are rare.

3.9 Halogen bonding series

Associate Professor Mark Taylor's group in the Department of Chemistry, University of Toronto, calculated the partial positive electrostatic potential surface (ESP) of **56a** at the iodo atom to evaluate the potential for halogen bonging, because the ESP can be linearly related to the free energy of halogen bonding in solution in simple systems (Table 3.2). The ESP of **56a** (14.30 kcal/mol) was found to be in line with the ESP of an inhibitor of hCatL, **58**, (18.23 kcal/mol), which is reported to act as a halogen bond donor,[36] providing us with some confidence in pursuing the halogen bonding hypothesis experimentally.

Table 3.2 Calculated partial positive electrostatic potential surface (ESP) of halogens (highlighted in red where more than one halogen is present). [a] Calculated with DFT (B3LYP/6-31+G(d,p) [on all atoms except Br and I] – LANL2DZdp [on Br and I]). [b] Calculated with DFT (B3LYP/6-31+G(d,p)[on all atoms except Br and I] – LANL2DZ [on Br and I]).

Compound ID	Structure	ESP at Halogen (kcal/mol)[a]
56a		14.30
58		18.23 [b]
56b		9.36
56c		4.76
56d		-11.62

If a halogen bond was contributing to the potency of **56a**, then replacing the iodo substituent with each of the other halogens would be expected to affect binding in accordance with the trend for halogen bond strengths: I > Br > Cl > F. This theory was confirmed computationally, by Mark Taylor's group who calculated the ESP at each halogen atom (Table 3.2; Figure 3.12). A series of 5-amino-2-halobenzene-1-sulfonamides, **56b**-**d** were therefore synthesised and tested using the competition fluorescence assay.

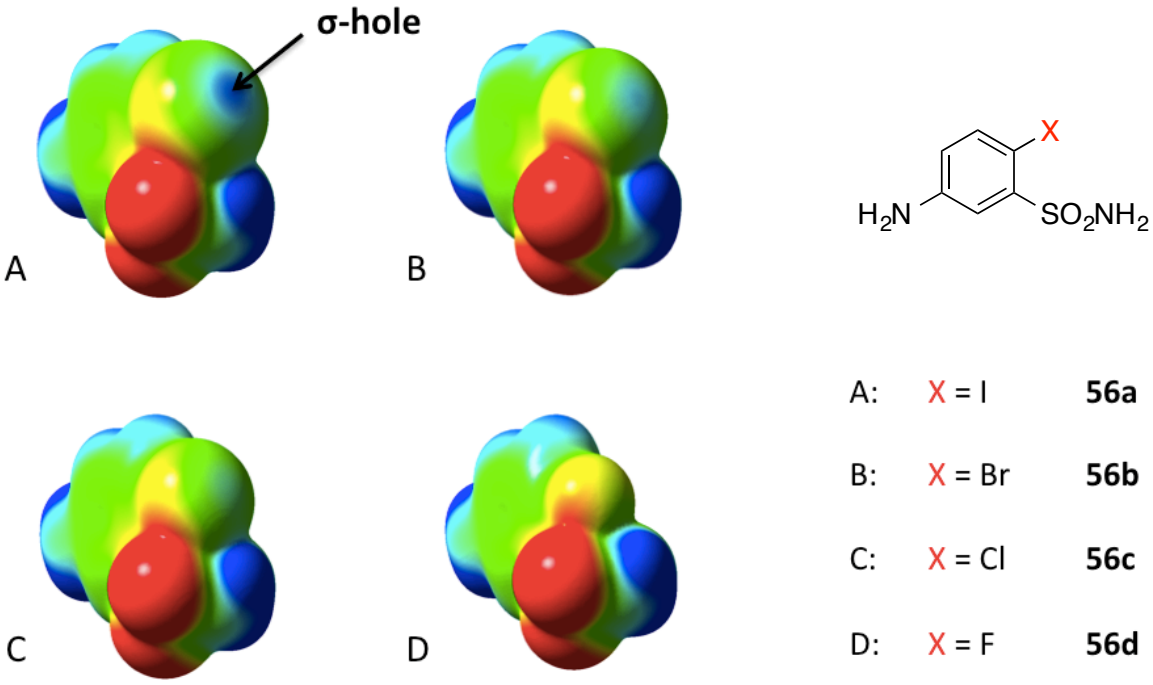
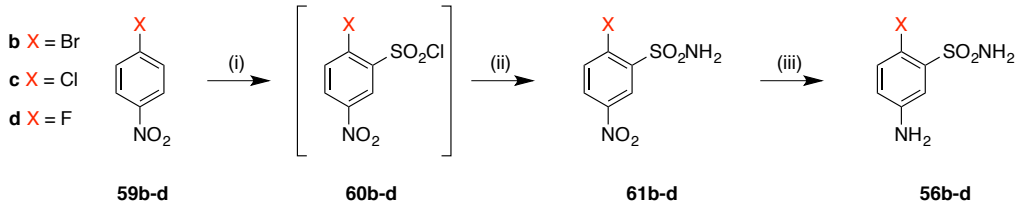


Figure 3.12 Demonstrates the electrostatic surface potential, where blue is positive and red is negative, of compound **56a** (A) compared to each of the other halogen analogues, **56b**-**d**, (B)-(D). Images of electrostatic potential surfaces, made by Mark Taylor's group, with density = 0.001 and limits of ±0.045 Hartrees (±28.24 kcal/mol).

3.10 Synthesis of the halogen bonding series

A literature search, suggested that the simplest route to each of the 5-amino-2-halobenzene-1-sulfonamides was to commence with the 1-halo-4-nitrobenzenes **59b**-**d** and use the respective *ortho*- and *meta*- directing abilities of the halo- and nitro- groups to form the 2-halo-5-nitrobenzene-1-sulfonyl chlorides **60b**-**d**. Treatment of the 2-halo-5-nitrobenzene-1-sulfonyl chlorides **60b**-**d** with ammonium hydroxide then yields the 2-halo-5-nitrobenzene-1-sulfonamides **61b**-**d**, which could undergo a subsequent reduction to form the 5-amino-2-halobenzene-1-sulfonamides **56b**-**d**. Each one of these steps was straightforward, with the desired regiochemistry being confirmed by ^{1}H-^{13}C HMBC, however, the final step required the conditions for the reduction of the nitro- group to be modified for X = Br and Cl (**61b** and **61c**) from Pd/C, H_2 to milder Zn/HCl. This change was necessary due to the Ar-X bond being

reduced by the former to give the undesired 3-aminobenzene-1-sulfonamide, **57**, with 5-amino-2-chlorobenzene-1-sulfonamide, **56c**, only being isolated in 9% yield using the former approach, despite literature precedence for the formation of **56c** in 83% yield from 2-chloro-5-nitrobenzene-1-sulfonamide, **61c**, in the presence of Pd/C, H_2.[37]



Scheme 3.6 Synthesis of the halogen bonding series. *Reagents and conditions*: (i) Chlorosulfonic acid, 110-120 °C, 16-20 h; (ii) Ammonium hydroxide, ethyl acetate (X = Br, F; **60b** and **60d**) or THF (X = Cl; **60c**), RT, 4-16 h, 30-48% (over two steps); (iii) Pd/C, H_2, MeOH, RT, 16 h, 91% (X = F; **61d**); Zn, THF/HCl, RT, 4-7 h, 79-89% (X = Br, Cl; **61b** and **61c**).

3.11 Testing the halogen bonding series by the competition fluorescence assay

Once synthesised, the 5-amino-2-halobenzene-1-sulfonamides **56b-d** were tested using the competition fluorescence assay (Figure 3.13). In addition to this series, 5-amino-2-methylbenzene-1-sulfonamide, **56e**, was screened, given the similar van der Waal radii of the iodo and methyl group, to rule out the possibility that the contribution of the iodo substituent was purely hydrophobic. Remarkably, it was found that variation of the halogen substituent had the expected effect on potency, in accordance with the hypothesised halogen bonding contribution. Furthermore, 5-amino-2-methylbenzene-1-sulfonamide, **56e**, bound worse than even the fluoro variant, **56d**, suggesting firstly that hydrophobicity alone was not a significant contribution to binding and secondly that the halogens are not altering activity by affecting the basicity of the aniline lone pair.

Having investigated the significance of the iodo-substituent to binding, it was decided to further probe the SAR of this series by exploring the contribution of the aniline. This was approached from two angles: the first involved acetylation of the aniline, **62**, and the second involved removal of the aniline, **63**. Both of these changes resulted in a significant loss in binding, which brought the halogen bonding theory into question because, assuming that they were all binding in the same place, it was unclear as to why the affinity would be completely lost through removal of either the aniline or the iodo substituents.

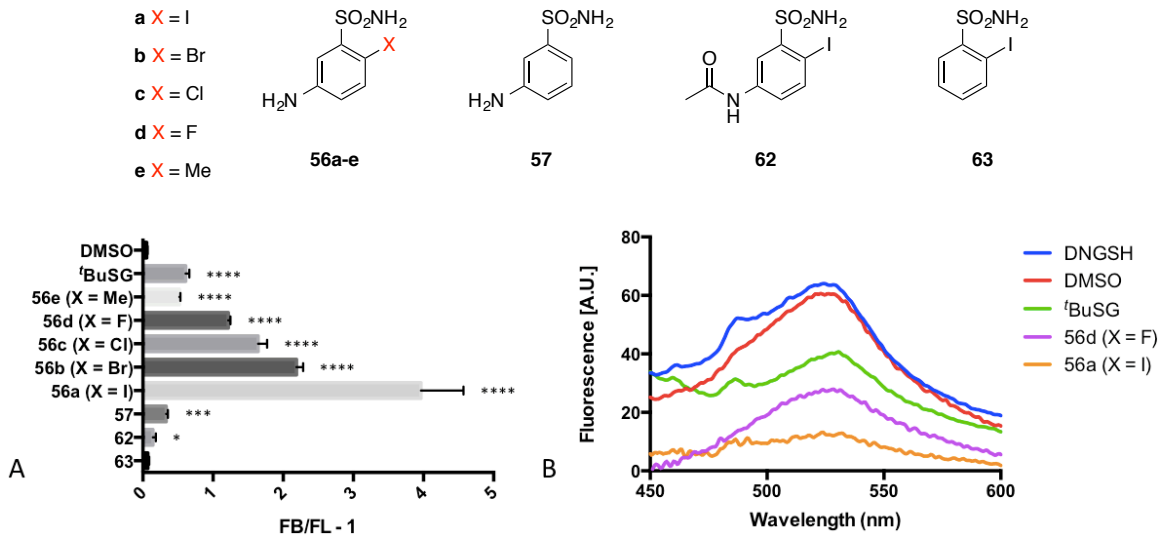


Figure 3.13 Analysis of the halogen-series binding to *Sd*KefCTD using the competition fluorescence assay. (A) ($F_B/F_L - 1$) values of each of these compounds (5 mM) relative to the positive control **^{t}BuSG** (5 mM) and negative control, **DMSO**, at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$; (B) The emission spectra for *Sd*KefCTD (6 µM) and **DNGSH** (5 µM) on their own (**DNGSH**) and in the presence of **^{t}BuSG**, **56a** and **56d** (5 mM) minus the emission spectra of the compounds (5 mM) on their own at 340 nm excitation.

The negative control of **DMSO** was determined to not be sufficient and the possibility that the 5-amino-2-halobenzene-1-sulfonamides, **56a**-**d**, could be false positives was explored through an additional negative control. The negative control that was introduced into the experimental design was to run the assay in the absence of protein. Unfortunately, it was found that the

(F_B/F_L – 1) values were the same in the presence and absence of protein (Figure 3.14), suggesting that the decrease in intensity was not a result of the **DNGSH** probe being displaced from the protein surface but due to the dosed compounds interfering with it.

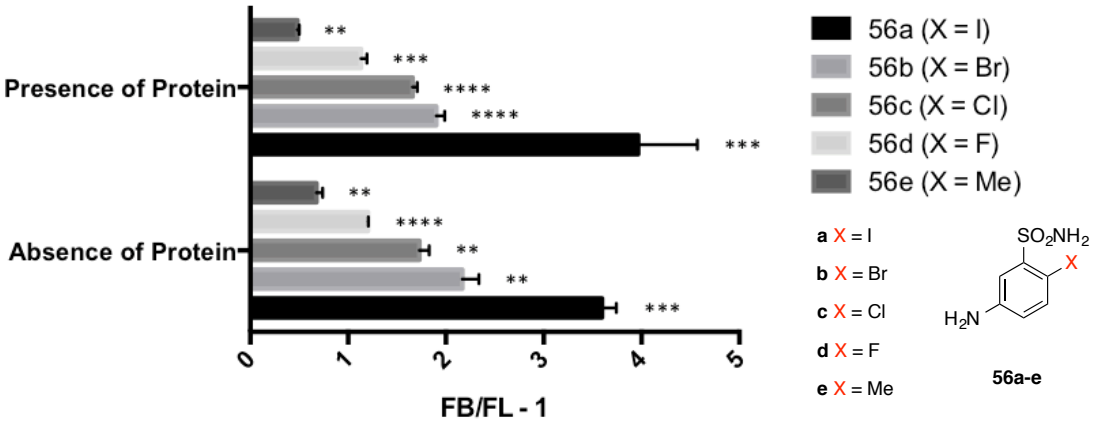


Figure 3.14 Comparison of the (F_B/F_L – 1) values obtained when running the competition fluorescence assay in the presence and absence of *Sd*KefCTD for the halogen-series. The graph shows (F_B/F_L – 1) values of each of these compounds (5 mM) at 525.5 nm emission in the presence of **DNGSH** (5 µM) and the presence and absence of *Sd*KefCTD (6 µM). Error bars indicate one standard deviation of uncertainty (n = 2, with the exception of **56a** in the presence of protein where n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$).

3.12 False positives arising from the competition fluorescence assay

In light of the false positives arising from the halogen bonding series, and suspicious about the high hit rate from the protein-ligand docking, it was decided to screen a selection of the hits in the absence of protein to identify any additional false positives. The compounds that were selected for retesting were **39**, **41** and **42** due to compound availability and the interest that was shown in them for optimisation stages. As can be observed in Figure 3.15, with the exception of **^{t}BuSG**, which did not show a positive result with any statistical significance in the absence of protein, all of the compounds had a statistically significant (F_B/F_L – 1) value in both the presence and absence of protein, suggesting that they are false positives.

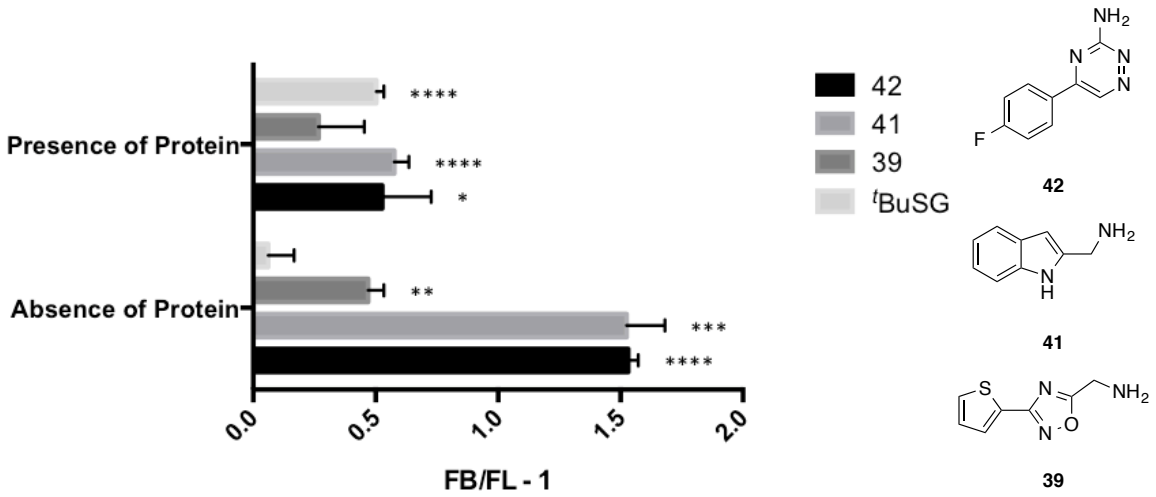


Figure 3.15 Comparison of the (F_B/F_L – 1) values obtained when running the competition fluorescence assay in the presence and absence of *Sd*KefCTD for leads of interest identified in the high throughput protein-ligand docking. The graph shows (F_B/F_L – 1) of each of these compounds (5 mM) at 525.5 nm emission in the presence of **DNGSH** (5 μM) and the presence and absence of *Sd*KefCTD (6 μM). Error bars indicate one standard deviation of uncertainty (n = 2, for the absence of protein, and n = 3 for the presence of protein). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$).

In summary, a number of compounds had been identified in the ligand-based virtual screening and high-throughput protein-ligand docking that showed potential to act as non-peptidic ligands of *Sd*KefCTD. When tested by both the competition fluorescence assay and STD NMR several of the compounds appeared to bind to *Sd*KefCTD. A series of compounds was synthesised based on the hypothesis that a halogen-bonding interaction was playing a key role in the affinity of one of the hits. When tested using the competition fluorescence assay, this series produced convincing SAR that was in line with the halogen bonding theory. However, upon introducing more rigorous controls, it was apparent that the competition assay results were misleading and the hits identified were false positives. As this was the first time that the competition fluorescence assay had been applied to non-peptidic ligands, this set back was not foreseen and caution was taken in further use of the competition fluorescence assay and STD NMR to detect binding.

3.13 Ligand-based virtual screening – sulfonamide series

In addition to the protein-ligand docking, AutoDock 4[15], Vina[1] and GOLD[2] were used to optimise the hits identified by Dr. Jess Healy and Niall Igoe in the ElectroShape screen: **54** (Figure 3.9).

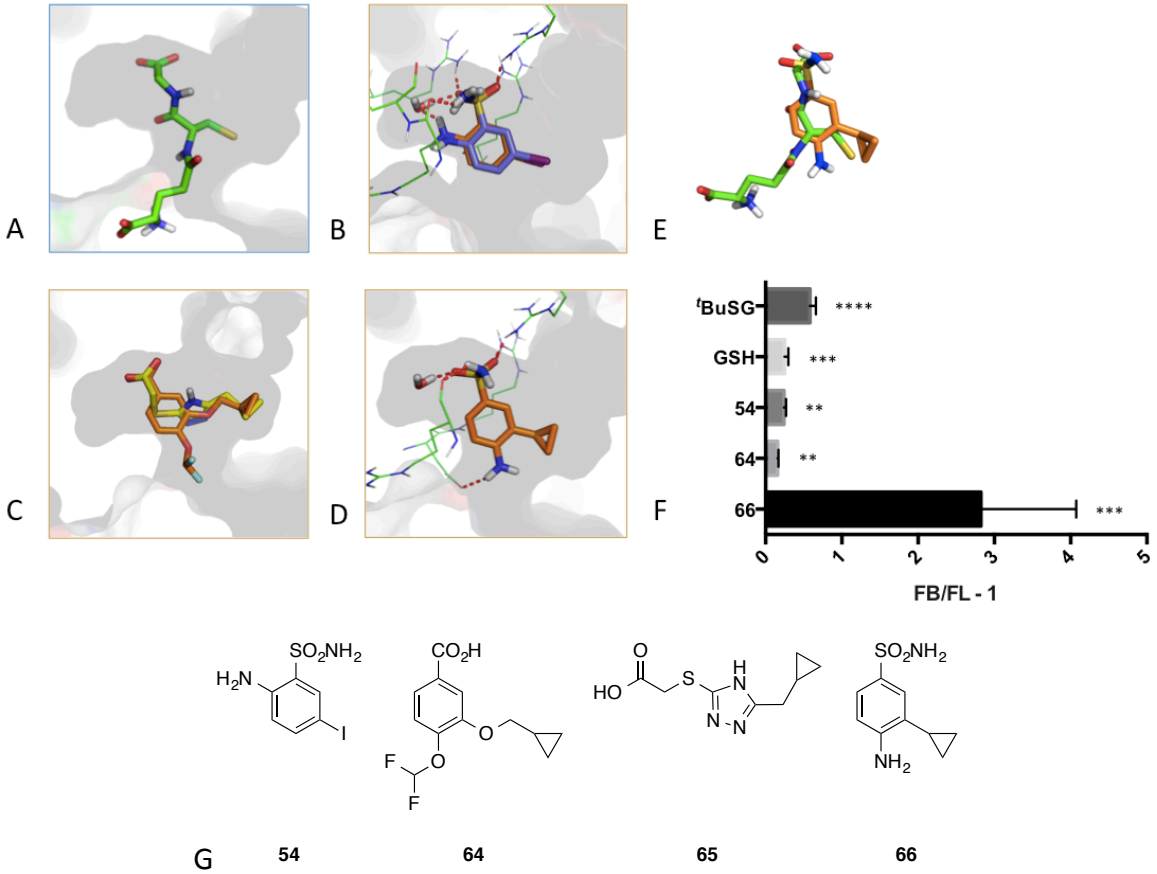


Figure 3.16 Optimisation of **54** using Vina and GOLD. Vina model carbons are depicted as purple and GOLD as orange (or yellow in the case of C). (A) The predicted bound conformation of glutathione in Hm1*Sd*KefCTDW1; (B) Docked conformations of **54**; (C) Two hits, **64** and **65** from the GOLD screen with Hm1*Sd*KefCTDW1 which have a cyclopropyl group occupying the glutathione thiol pocket; (D) The docked conformation of **66**; (E) Compound **66** overlaid with the predicted bound conformation of glutathione; (F) ($F_B/F_L - 1$) values of each compound (5 mM) relative to the positive control, **^{t}BuSG**, (5 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3, with the exception of **54**, where n = 2). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (G) The structures of the compounds tested. Images were constructed using PyMOL.[28]

54 was docked into Hm1*Sd*KefCTDW1 using each of these programs and a predicted binding mode that was common to all three programs was identified. This binding mode was in line with the ElectroShape prediction of Figure 3.9, where the sulfonamide acts as a bioisostere of glutathione's Gly-carboxylate group, interacting with R423, R507, R523 and the carbonyl of the R416-F417 amide bond of Hm1*Sd*KefCTDW1 (R423 and R523 are the equivalent residues to R416 and R516 in *Ec*KefCCTD). Furthermore, the iodo-substituent is oriented in the hydrophobic pocket that encompasses the thiol of glutathione (Figure 3.16 A & B). In an attempt to optimise **54** it was decided to modify the iodo-substituent to another hydrophobic group that could occupy the thiol-pocket. Inspiration for this change came from two compounds identified in the GOLD[2] high throughput protein-ligand docking with Hm1*Sd*KefCTDW1, **64** and **65**, (Figure 3.16 C & G), both of which had a cyclopropyl group occupying the thiol-pocket. Further analysis of the docked structure predicted that the aniline was not forming any key electrostatic interactions and so could potentially be synthetically repositioned. It was decided to try moving the aniline to the 4 position of the benzene ring as this would provide the opportunity for further optimisation through probing the Glu-region of the glutathione-site (Figure 3.16 E). The resulting compound after these modifications was **66**, which was docked into the homology model Hm1*Sd*KefCTDW1 using AutoDock 4[15], Vina[1] and GOLD[2] and predicted to bind in the proposed mode by GOLD[2], with the sulfonamide forming electrostatic interactions with R423, R507 and the carbonyl of R416-F417 amide bond (Figure 3.16 D). **66** was synthesised and tested alongside **64** by Dr. Jess Healy using the competition fluorescence assay. Although **64** did not appear to be any more potent than glutathione (**GSH**), **66** appeared to bind better than **^{t}BuSG**. A series of compounds was thus developed based on **66**.

It was decided to replace the cyclopropyl group with a number of lipophilic substituents, which were docked using AutoDock 4[15], Vina[1] and GOLD[2] to determine the optimal mimic of the glutathione thiol. As a further optimisation, the aniline was also acetylated to potentially mimic

the interactions of the Cys-Glu amide bond, as demonstrated by compound **67a** overlaid with glutathione (Figure 3.17 A-D). A docked conformation of compound **67a** demonstrates how the sulfonamide is inferred to mimic the predicted interactions of the Glu-Gly carboxylate with the residues R423, R507, R523 and the acetyl group picks up a polar interaction with Q441 (Figure 3.17 E).

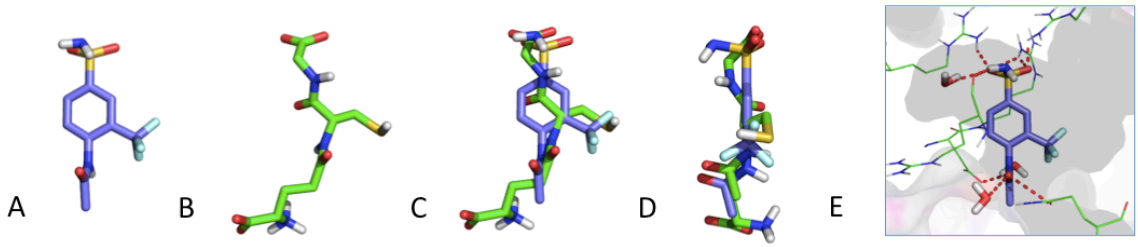
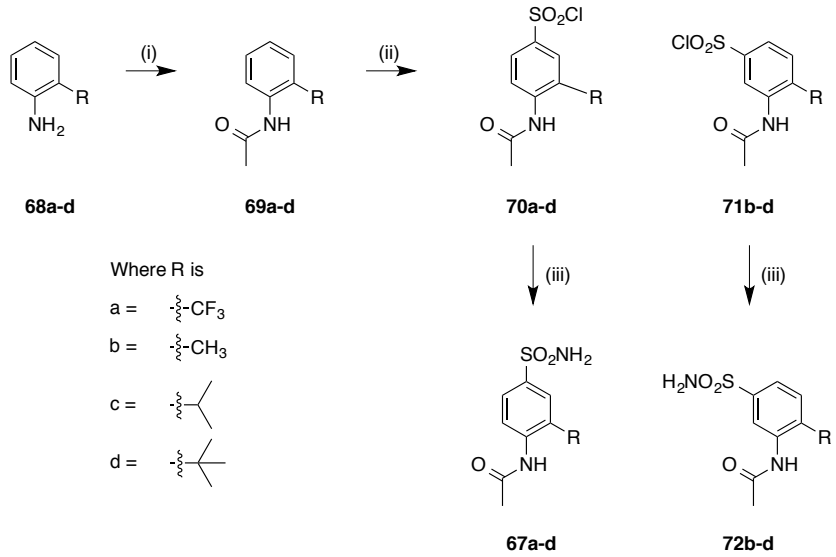


Figure 3.17 Optimisation of the sulfonamide series. (A) Structure of compound **67a**; (B) Structure of **GSH**; (C) Front view of **GSH** overlaid with **67a**; (D) Side view of **GSH** overlaid with **67a**; (E) Predicted binding mode of compound **67a** from docking into Hm1*Sd*KefCTDW1 with Vina. Images were constructed using PyMOL.[28]

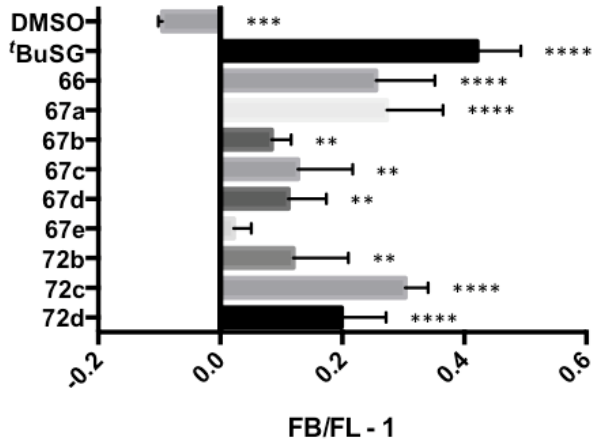
The analogues that were predicted to bind best were synthesised and tested (Scheme 3.7). The 2-alkylanilines of the proposed analogues **68a-d** were all commercially available and so were initially acetylated to form the *para*-directing arylacetylamide to favour formation of the 4-acetamido-3-alkylbenzene-1-sulfonyl chlorides **70a**-**d** in the chlorosulfonylation step. However, the chlorosulfonylation was not as straightforward as anticipated, with competition from the *para*-directing alkyl groups favouring the formation of the 3-acetamido-4-alkylbenzene-1-sulfonyl chlorides **71b**-**d**. The only exception to this observation was when R = CF_3, **67a**, due to the electron withdrawing capabilities of this substituent favouring chlorosulfonylation at the position *meta*- to the CF_3. Analysis of the crude ^{1}H NMR spectra of the sulfonamides **67/72c-d** resulting from the reaction of the mixture of sulfonyl chlorides with ammonium hydroxide found that the ratio of the desired regioisomer *N*-(2-alkyl-4-sulfamoylphenyl)acetamides, **67b**-**d**, to *N*-(2-alkyl-5-sulfamoylphenyl)acetamides, **67b**-**d**, was 1:1.8. Separation of the regioisomers was achieved through successive crystallisations in aqueous ethanol (methanol in the case of

compound **67c**). The more abundant regioisomer *N*-(2-alkyl-5-sulfamoylphenyl)acetamides **72b**-**d** crystallised out in the first round of crystallisation, allowing purification of the desired regioisomer *N*-(2-methyl-4-sulfamoylphenyl)acetamides **67b**-**d** in a second round of crystallisation.

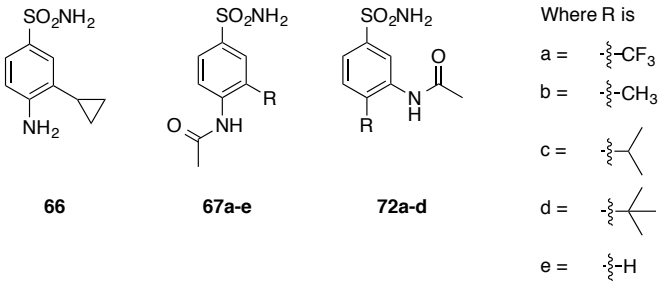


Scheme 3.7 Synthesis of the sulfonamide series. *Reagents and conditions*: (i) NEt_3, CH_2Cl_2 or $CHCl_3$, acetyl chloride, RT, 3-5.5 h, 61-100%; (ii) Chlorosulfonic acid, 60 °C 0.5-21 h; 2-13% (iii) For **67a-d**: Ammonium hydroxide, RT, 2.5-16 h, 13-33% (over two steps); For **72b-d**: Ammonium hydroxide, RT, 1-16 h, 20-33% (over two steps).

Given that both regioisomers had been synthesised and purified, it was decided to test both in the competition fluorescence assay alongside **66** (Figure 3.18). It was found that **66** was not as potent as in the original screen performed by Dr. Jess Healy. Interestingly, **67e**, where R = H, appeared not to bind with any statistical significance, suggesting that the lipophilic groups in the 2 position could be contributing to affinity. **67a** and **72c**, which gave the largest ($F_B/F_L - 1$) values, and two of the medium binders **67b** and **72b** were progressed to STD NMR to confirm binding. It was found that **67a, 67b** and **72b** bound by STD NMR (Figure 3.19) but **72c** did not.



A



B

Figure 3.18 Analysis of the sulfonamide-series binding to *Sd*KefCTD using the competition fluorescence assay. (A) Shows (F_B/F_L – 1) values of each of these compounds (5 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) Structures of the compounds tested.

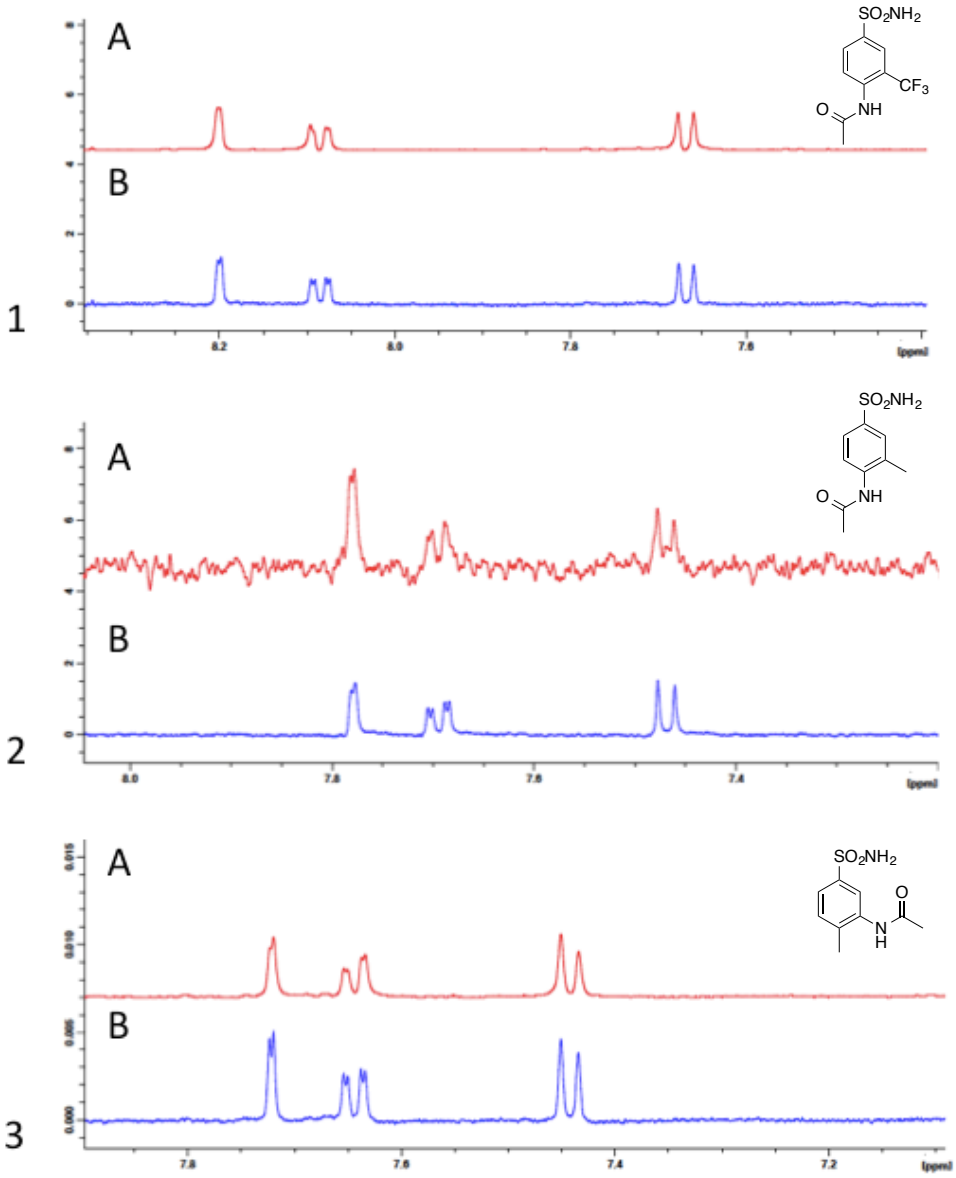
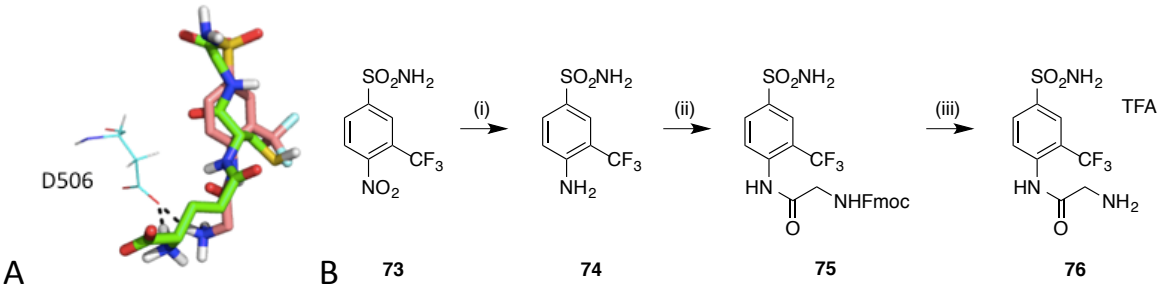


Figure 3.19 STD NMR spectrum for compounds: (1) **67a**, (2) **67b**, (3) **72b**. In each case, (A) is the difference spectrum (spectrum before saturation minus the saturated spectrum); (B) is the spectrum before saturation.

Given that **67a** was the most potent compound in the competition fluorescence assay that was found to bind by STD NMR, it was decided to pursue the idea of gaining affinity through building off the aniline into the glutathione-Glu region. A virtual library of compounds was constructed, by coupling 1,017,567 commercially available carboxylic acids in InhibOx's in-house database to the aniline of **74** using InhibOx's proprietary program ReactiOx. This library was subsequently docked into Hm1*Sd*KefCTDW3 using GOLD[2]. In order to lock the sulfonamide of **74** into its

predicted binding mode, where the sulfonamide forms hydrogen bonds with the residues R423, R507, R523 (R423 and R523 are the equivalent residues to R416 and R516 in *Ec*KefCCTD), *protein_h_bond* constraints were imposed so that GOLD[2] favoured solutions involving hydrogen bonds between the docked ligand and these residues. The *protein_h_bond* constraints consisted of: the <*constraint weight*> or penalty for not forming the specified hydrogen bonds, which was set to 10.0000, and the <*min. geometry weight*> that determines the quality of a hydrogen bonding interaction, which was kept at its default value of 0.005.

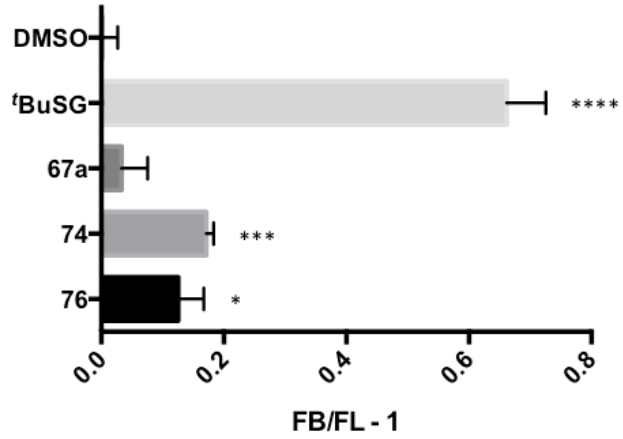
Manual curation of the top 100 hits, based on ligand efficiency (the binding energy of the ligand per atom),[38] of the docked library of carboxylic acids coupled to **74**, identified an interesting compound, where the coupled carboxylic acid was glycine (compound **76**; Figure 3.8 A). The reason for the interest in compound **76** was that it was predicted to mimic the interactions that glutathione picks up with D506, which is predicted to form a hydrogen bonding interaction with the Glu of glutathione (Scheme 3.8 A). Compound **76** was therefore selected for further investigation, due to it acting as a good proof of concept for investigating whether affinity could be gained by building into the Glu-region of the glutathione-binding site. As the previous route to **67a** was low yielding, due to the electron deficiency of the ring system, an alternative synthesis was pursued using commercially available 4-nitro-3-(trifluoromethyl)benzene-1-sulfonamide, **73**, as a starting point (Scheme 3.8 B). Once **74** was in hand, Fmoc-protected glycine was converted to its acid chloride and coupled at the aniline position of **74**, before a final deprotection step was performed to yield **76**.



Scheme 3.8 *In silico* optimisation of compound **67a**. (A) Glutathione overlaid with **76**, highlighting the potential to mimic glutathione further by picking up additional interactions with the residue D506. (B) Synthesis of **76**. *Reagents and conditions*: (i) Pd/C, H_2, MeOH, RT, 5 h, 99%, (ii) Fmoc-glycyl chloride, MeCN, NEt_3, RT, 15.5 h, 10% (over two steps); (iii) Piperidine, DMF, RT, 18 h, 10%. Image (A) was constructed using PyMOL.[28]

A modest increase in potency was observed by the competition fluorescence assay when comparing **76** to **67a** (Figure 3.20). However, **67a** no longer appeared to bind with statistical significance and no binding was observed by STD NMR analysis for **76**.

A



B

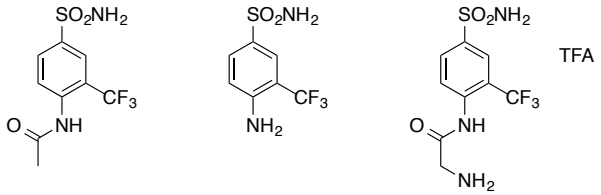
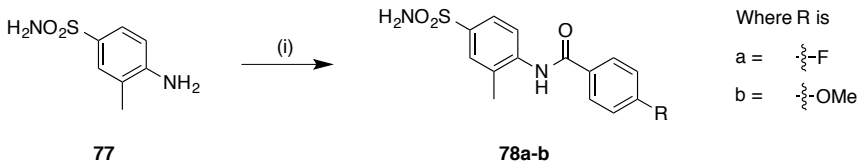


Figure 3.20 Analysis of analogues of **SG56A** binding to *Sd*KefCTD using the competition fluorescence assay. (A) Shows (F_B/F_L – 1) values of each of these compounds (5 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 μM) and **DNGSH** (5 μM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) Structures of the compounds tested.

Before suspending the investigation of this series of compounds, it was decided to investigate whether affinity could instead be gained by building off the aniline with more hydrophobic groups such 4-methoxyphenyl and 4-fluorophenyl. Instead of using 4-amino-3-(trifluoromethyl)benzene-1-sulfonamide, **74**, as the building block, 4-amino-3-methylbenzene-1-sulfonamide, **77**, was used due to difficulties that were encountered with coupling to the aniline of 4-amino-3-(trifluoromethyl)benzene-1-sulfonamide, **74**, which was presumably due to the electron deficiency of the ring system. 4-Amino-3-methylbenzene-1-sulfonamide, **77**, was commercially available and so was coupled to 4-fluorobenzoyl chloride and 4-methoxybenzoyl chloride to afford **78a** and **78b** with yields of 45% and 48% respectively (Scheme 3.9). Unfortunately due to solubility issues encountered with

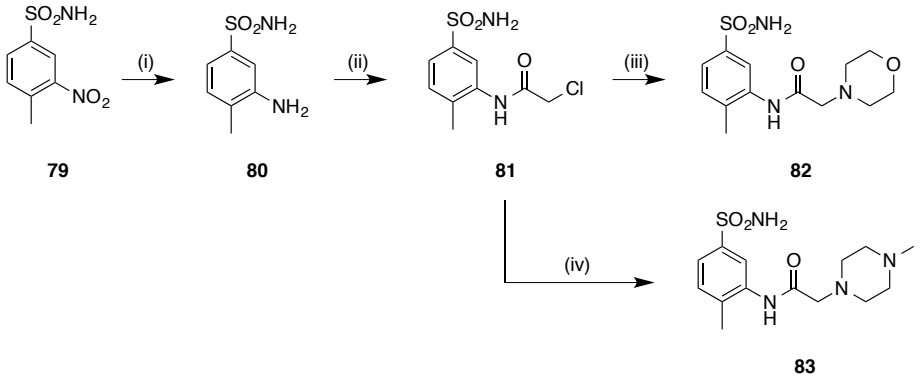
78a and **78b**, it was not possible to test them using either the competition fluorescence assay or STD NMR.



77 **78a-b**

Scheme 3.9 Synthesis of 4-fluoro-*N*-(2-methyl-4-sulfamoylphenyl)benzamide, **78a**, and 4-methoxy-*N*-(2-methyl-4-sulfamoylphenyl)benzamide, **78b**. *Reagents and conditions*: (i) For **78a**: NEt_3, THF, 4-fluorobenzoyl chloride, RT, 22 h, 45%; For **78b**: or 4-methoxybenzoyl chloride, RT, 22 h, 48%.

The SAR of **72b**, was also explored by building off the aniline position in parallel to the investigations of **78a** and **78b**. Given the low solubility encountered from building off the aniline of **77**, it was decided to append groups that might aid with the solubility, such as morpholine and methylpiperazine. These groups were built onto **80** by acetylating the aniline with chloroacetyl chloride followed by an S_N2 reaction with morpholine and methyl piperazine to afford compounds **82** and **83** (Scheme 3.10). Solubility was found to improve in this series of compounds, and so they were tested using STD NMR.



79 **80** **81** **82** **83**

Scheme 3.10 Synthesis of *N*-(2-methyl-5-sulfamoylphenyl)-2-(morpholin-4-yl)acetamide, **82**, and *N*-(2-methyl-5-sulfamoylphenyl)-2-(4-methylpiperazin-1-yl)acetamide, **83**. *Reagents and conditions*: (i) Pd/C, H_2, MeOH, RT, 63.5 h, 99%; (ii) Chloroacetyl chloride, CH_2Cl_2, RT, 0.5 h, 36%; (iii) Morpholine in ethanol, MeOH, 80 °C, 6 h, 47%; (iv) 1-Methylpiperazine in ethanol, MeOH, 80 °C, 6 h, 51%.

Interestingly, unlike with compound **72b**, where each of the peaks were observed to have saturation transferred from the protein to the ligand protons with equal intensity (Figure 3.19 3A), one of the protons in **82** appeared to have saturation transferred to it to a greater extent (H^6; Figure 3.21 A). It was thought that this observation could be characteristic of a binding epitope effect, where the protons on the ligand that are closer to the protein in the binding mode have more saturation transferred to them.[22] Consultation with Prof. Tim Claridge, however, suggested that the results obtained by STD NMR were probe-related false positives, and he suggested running some of the compounds on a different spectrometer to see if they still bound. Compound **82** and its precursor **72b** were therefore retested using STD NMR on the AVIII 600 MHz spectrometer. Compound **82** was found not to bind, and only the methyl peaks of **72b** could be observed in the difference spectrum, but none of its aromatic peaks, as was previously observed. Furthermore, it was pointed out by Prof. Tim Claridge that in a number of the previous STD NMR spectra obtained, the peak for the protein in the difference spectrum was negative relative to the peaks corresponding to the ligands, which did not make sense. Given the suspiciously high hit rate from STD NMR, and the lack of confidence in the technique at the time, it was decided to suspend the development of compounds using this technique.

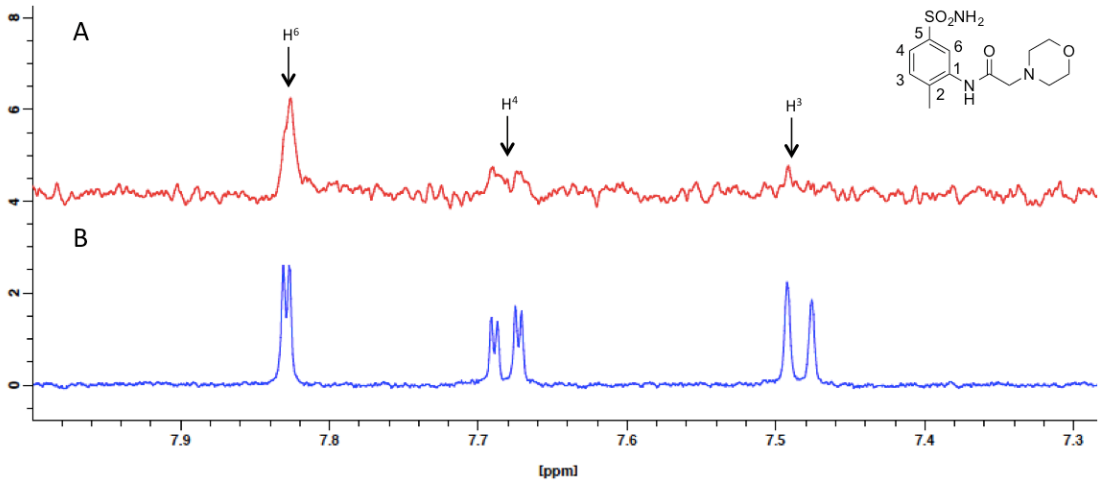


Figure 3.21 STD NMR spectrum for compound **82** (A) shows the difference spectrum (spectrum before saturation minus the saturated spectrum); (B) shows the spectrum before saturation.

In parallel to the above work, another route to optimising **66** (Figure 3.22 A) was explored, which involved combining **66** with compound **84**, a hit identified in the GOLD[2] high throughput protein-ligand docking with Hm1*Sd*KefCTDW1 (Figure 3.22 B).

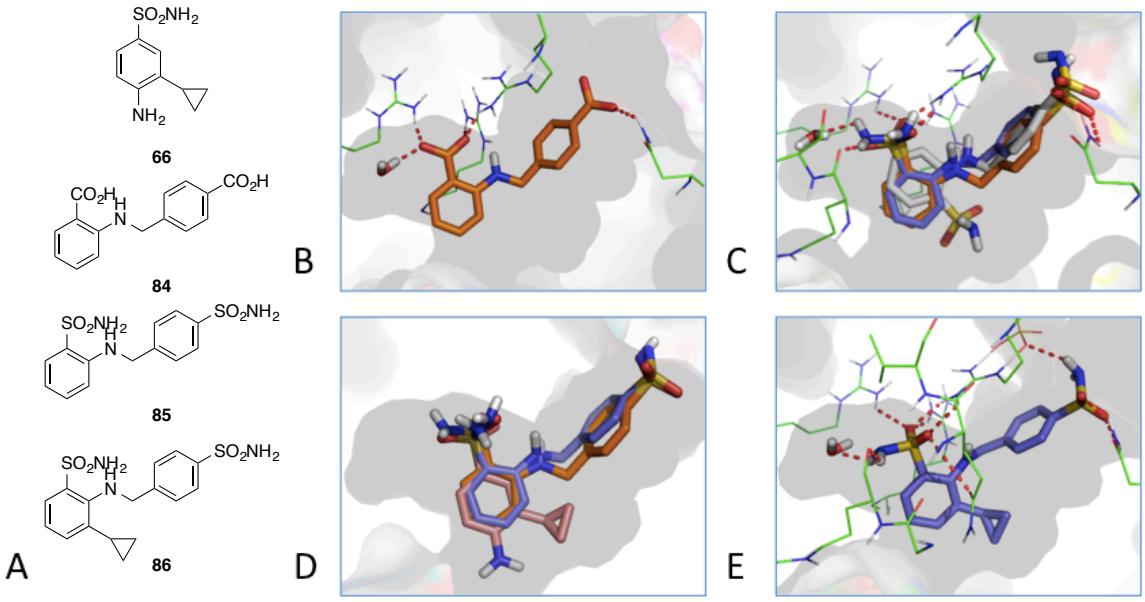
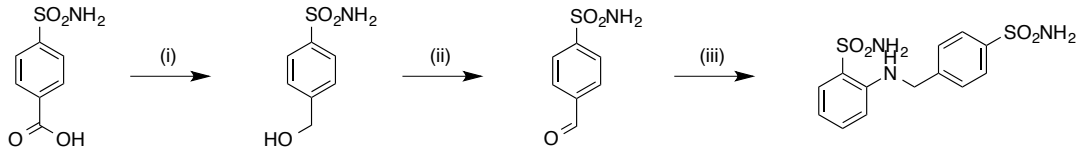


Figure 3.22 *In silico* optimisation of **66** using AutoDock 4, Vina and GOLD. AutoDock 4 model carbons are depicted as white, Vina as purple and GOLD as orange. (A) Structures of the compounds docked. (B) The docked conformation of **84** from the GOLD screen using Hm1*Sd*KefCTDW1; (C) The docked conformations of **85**; (D) The docked conformations of **66** and **85** overlaid; (E) The docked conformation of **86**. Images were constructed using PyMOL.[28]

This compound was of interest because it probed a novel region of the protein, which is not occupied by glutathione and so could help with building in selectivity over other glutathione proteins. Analysis of the binding mode of **84** found that a carboxylic acid was mimicking the interactions of the Gly of glutathione with the residues R423, R507, R523, with a 4-(aminomethyl)benzoic acid substituent occupying the novel pocket, picking up interactions through the carboxylic acid with the residue N450. Given that **66** used a sulfonamide to act as a bioisostere for a carboxylic acid, it was decided to replace the carboxylic acids of **84** and dock the resulting compound, **85**, using AutoDock 4,[15] Vina[1] and GOLD[2] (Figure 3.22 C). Vina[1] and GOLD[2] were able to replicate the binding mode of compound **84** predicted by the original GOLD

screen when docking in compound **85** (Figure 3.22 C). Furthermore, by overlaying these docked structures with the docked structure of **66**, it was found that placing a cyclopropyl group in the 3 position of compound **85** might help to pick up interactions with the glutathione-thiol binding pocket as well (Figure 3.22 D). This hypothesis was investigated through docking compound **86** with AutoDock 4[15], Vina[1] and GOLD.[2] Vina[1] found a solution that potentially supported it (Figure 3.22 E).

Before synthesising the analogue **86** it was decided to test **85** first to determine if it would bind. Compound **85** was synthesised by performing a reductive amination of 2-aminobenzene-1-sulfonamide with 4-formylbenzene-1-sulfonamide **89** (Scheme 3.11).



87 **88** **89** **85**

Scheme 3.11 Synthesis of **85**. *Reagents and conditions*: (i) Borane·THF, THF, RT, 3.5 h, 64%, (ii) Manganese dioxide, MeCN, RT, 48 h, 67%; (iii) 2-Aminobenzene-1-sulfonamide, acetic acid, $NaCNBH_3$, ethanol, 25 °C, 6 h, 55%.

Compound **85** was tested using competition fluorescence assay and STD NMR. Compound **85** was too insoluble to be tested in the competition fluorescence assay but was found to bind using STD NMR (Figure 3.23). However, given issues with STD NMR that are highlighted in Section 3.13 further work on this series of compounds was terminated.

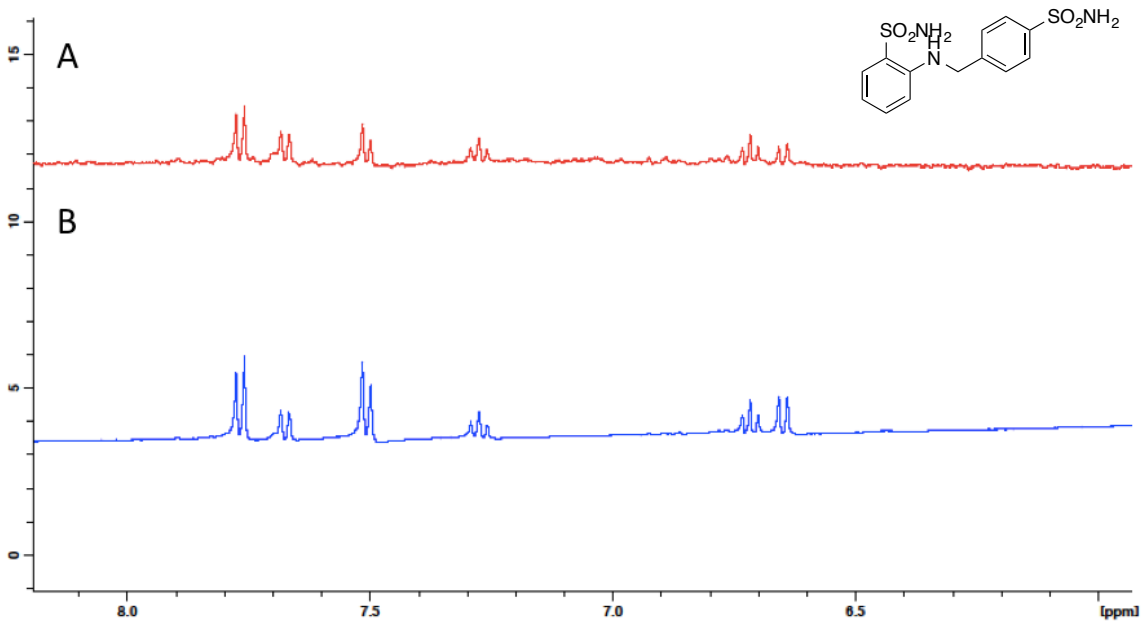
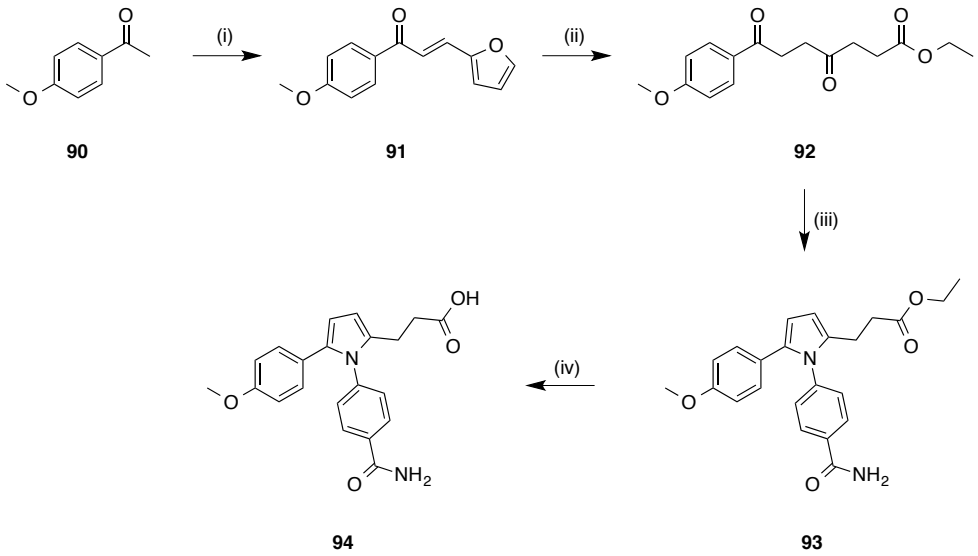


Figure 3.23 STD NMR spectrum for compound **85** (A) shows the difference spectrum (spectrum before saturation minus the saturated spectrum); (B) shows the spectrum before saturation.

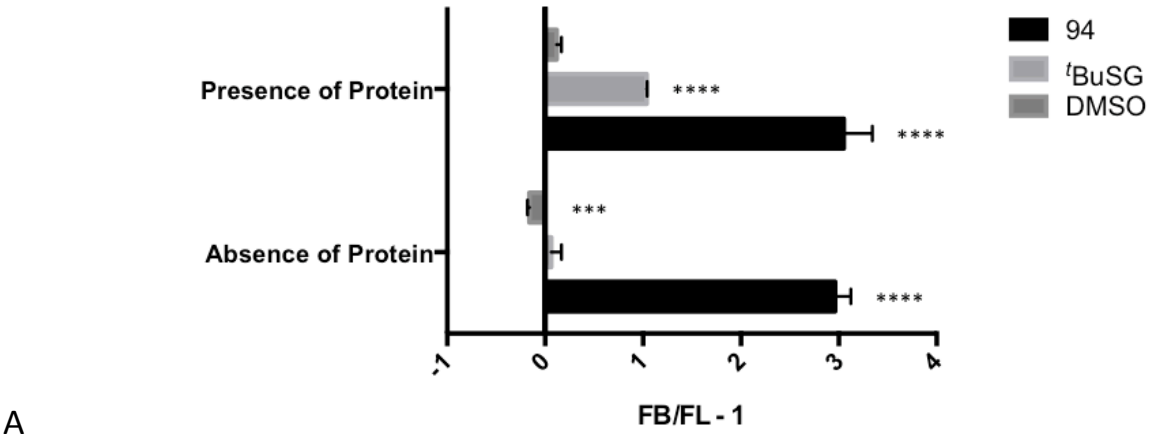
3.14 Synthesis and *in vitro* testing of a non-peptidic *S*-nitrosoglutathione reductase inhibitor

Another approach to identifying non-peptidic ligands that was employed in parallel to the *in silico* screening was to synthesise and test reported small non-peptidic ligands of other glutathione binding sites to see if their activity could be translated to Kef. A literature search identified small molecule inhibitors of *S*-nitrosoglutathione reductase (GSNOR), a member of the alcohol dehydrogenase family that regulates intracellular *S*-nitrosoglutathione (GSNO).[20] As an initial test, one of the small molecules, **94**, (Scheme 3.12) was selected for synthesis and screening using the competition fluorescence assay.

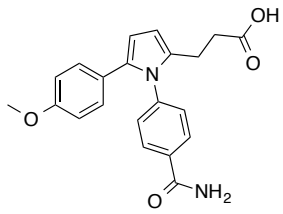


Scheme 3.12 Synthesis of **94**. *Reagents and conditions*: (i) 2-Furaldehyde, NaOMe, MeOH, RT, 23 h, 100%; (ii) Conc. HCl, ethanol, 80 °C, 16 h; (iii) 4-Aminobenzamide, 4-toluenesulfonic acid monohydrate, ethanol, 80 °C, 16 h, 9% (over two steps); (iv) aqueous 1 M NaOH, ethanol, RT, 1 h, 45 °C, 1 h, 87%.

As the problem of false positives arising from the competition fluorescence assay had become apparent from the halogen bonding series, the negative control of no protein was also performed when testing **94** (Figure 3.24). Unfortunately any apparent binding of **94** was observed in both the absence and presence of protein, further highlighting the need to develop an alternative, quantitative assay that was not as prone to false positives when testing non-peptidic ligands.



A



B

94

Figure 3.24 Analysis of **94** binding to *Sd*KefCTD using the competition fluorescence assay. (A) Shows $(F_B/F_L - 1)$ values of **94** (5 mM) and **^{t}BuSG** (5 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) The structure of **94**.

3.15 Retesting false positives using a fluorescence polarisation assay

A number of assays were developed, one of which was a fluorescence polarisation assay that is described in Chapter 4. A selection of the compounds from each of the series that were suspected to be false positives were re-tested using this fluorescence polarisation assay (Figure 3.25). The key concept behind this assay is that the fluorescent probe, **95**, has a high polarisation when bound to the protein and a low polarisation when it is displaced by a competing ligand, as demonstrated by the positive control **^{t}BuSG** (Figure 3.25 A). Unfortunately, fluorescence polarisation suggested that the compounds previously thought to be binding with a similar affinity to **^{t}BuSG** by the competition fluorescence assay were all false positives, with none

of them showing a statistically significant difference from the negative control of buffer except for **94**. **94** could be binding given that it showed a statistically significant difference to the negative control of buffer, however, it was not deemed to be worth following up given successes in other areas of the project, and that it was likely to be a very weak binder.

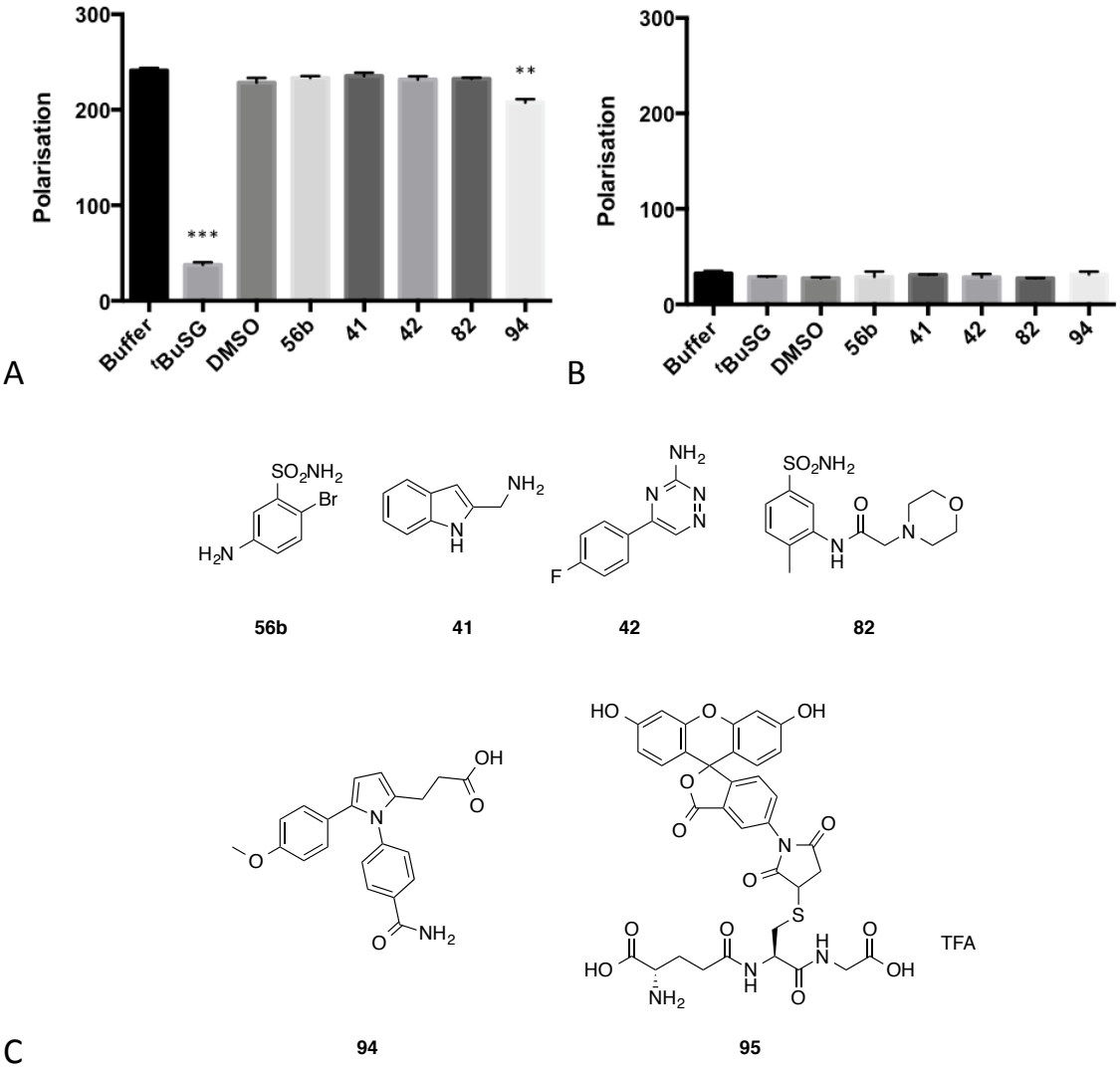


Figure 3.25 Retesting suspected false positives using a fluorescence polarisation assay described in Chapter 4. (A) The polarisation of the fluorescein probe, **95**, (25 nM) in the presence of *Sd*KefCTD (180 µM) and competing ligands (1 mM) or the negative control of buffer. (B) The polarisation of fluorescein probe, **95**, (25 nM) in the absence of *Sd*KefCTD and presence of competing ligands (1 mM). Error bars indicate one standard deviation of uncertainty (n = 3 for A, n = 2 for B). Significance of changes relative to the buffer were evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (C) Structures of compounds retested and the fluorescent probe **95**.

3.16 Conclusions

Although a number of approaches had been taken to developing non-peptidic ligands of *Sd*KefCTD, it was not possible to assess their ability to bind due to flaws in the assays implemented in the *in vitro* screening stage. The competition fluorescence assay had been validated using peptides, which do not interfere with the fluorescent probe **DNGSH**, however, as soon as non-peptidic ligands were tested, interference with **DNGSH** became a problem. There was therefore a need to develop alternative assays that could be applied in a quantitative manner, a problem that was addressed in Chapter 4.

In addition, protein-ligand docking was performed against a homology model, which was constructed using the conformation adopted by *Ec*KefCCTD when glutathione is binding as a template (PDB code: 3L9W),[27] and ligand-based screening was performed against the conformation adopted by glutathione when bound to *Ec*KefCCTD as observed in the X-ray crystal structure (PDB code: 3L9W).[27] Given that glutathione has an affinity of 900 μM (measured by fluorescence emission spectra),[21] it was unlikely that *in silico* virtual screening would yield a compound with an affinity lower than 900 μM. In order to identify potent starting points, future *in silico* studies needed to focus on using glutathione S-conjugates that have tighter binding for *Sd*KefCTD, such as **^{t}BuSG** (reported K_D of 400 ± 200 nM; measured by fluorescence emission spectra).[21] **^{t}BuSG** was dissected in Chapter 5, and the group efficiencies of each section of the peptide backbone were determined. This information could be used to help develop a pharmacophore model for future *in silico* ligand-based screening efforts.

3.17 Chapter 3 References:

[1] O. Trott, A. J. Olson, *J Comput Chem* **2010**, *31*, 455–461.
[2] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727–748.
[3] M. S. Armstrong, G. M. Morris, P. W. Finn, R. Sharma, L. Moretti, R. I. Cooper, W. G. Richards, *J. Comput. Aided Mol. Des.* **2010**, *24*, 789–801.
[4] P. Willett, J. M. B. and, G. M. Downs, *J. Chem. Inf. Model.* **1998**, *38*, 983–996.
[5] G. B. McGaughey, R. P. Sheridan, A. C. Good, *Journal of medicinal …* **2010**.
[6] S. Kortagere, M. D. Krasowski, S. Ekins, *Trends Pharmacol. Sci.* **2009**, *30*, 138–147.
[7] D. E. Patterson, R. D. Cramer, A. M. Ferguson, R. D. Clark, L. E. Weinberger, *J. Med. Chem.* **1996**, *39*, 3049–3059.
[8] P. C. D. Hawkins, A. A. Geoffrey Skillman, A. Nicholls, *J. Med. Chem.* **2006**, *50*, 74–82.
[9] J. A. Grant, M. A. Gallardo, B. T. Pickup, *J Comput Chem* **1996**, *17*, 1653–1666.
[10] T. S. Rush III, J. A. Grant, A. L. Mosyak, A. Nicholls, *J. Med. Chem.* **2005**, *48*, 1489–1495.
[11] R. P. Sheridan, G. B. McGaughey, W. D. Cornell, *J. Comput. Aided Mol. Des.* **2008**, *22*, 257–265.
[12] P. J. Ballester, P. W. Finn, W. G. Richards, *J. Mol. Graph. Model.* **2009**, *27*, 836–845.
[13] P. J. Ballester, W. G. Richards, *J Comput Chem* **2007**, *28*, 1711–1723.
[14] M. S. Armstrong, G. M. Morris, P. W. Finn, R. Sharma, W. G. Richards, *J. Mol. Graph. Model.* **2009**, *28*, 368–370.
[15] G. M. Morris, D. S. Goodsell, R. S. Halliday, *Journal of …* **1998**.
[16] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J Comput Chem* **2009**, *30*, 2785–2791.
[17] R. Huey, G. M. Morris, A. J. Olson, D. S. Goodsell, *J Comput Chem* **2007**, *28*, 1145–1152.
[18] O. Korb, T. Stützle, T. E. Exner, *J. Chem. Inf. Model.* **2009**, *49*, 84–96.
[19] J. W. Liebeschuetz, J. C. Cole, O. Korb, *J. Comput. Aided Mol. Des.* **2012**, *26*, 737–748.
[20] X. Sun, J. W. F. Wasley, J. Qiu, J. P. Blonder, A. M. Stout, L. S. Green, S. A. Strong, D. B. Colagiovanni, J. P. Richards, S. C. Mutka, et al., *ACS Med Chem Lett* **2011**, *2*, 402–406.
[21] J. Healy, S. Ekkerman, C. Pliotas, M. Richard, W. Bartlett, S. C. Grayer, G. M. Morris, S. Miller, I. R. Booth, S. J. Conway, et al., *Biochemistry* **2014**, *53*, 1982–1992.
[22] B. Meyer, T. Peters, *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 864–890.
[23] A. Viegas, J. Manso, F. L. Nobrega, E. J. Cabrita, *J. Chem. Educ.* **2011**, *88*, 990–994.
[24] L. Lyngberg, J. Healy, W. Bartlett, S. Miller, S. J. Conway, I. R. Booth, T. Rasmussen, *J. Bacteriol.* **2011**, *193*, 4925–4932.
[25] S. Miller, L. S. Ness, C. M. Wood, B. C. Fox, I. R. Booth, *J. Bacteriol.* **2000**, *182*, 6536–6540.
[26] C. Chothia, A. M. Lesk, *EMBO J.* **1986**, *5*, 823–826.
[27] T. P. Roosild, S. Castronovo, J. Healy, S. Miller, C. Pliotas, T. Rasmussen, W. Bartlett, S. J. Conway, I. R. Booth, *PNAS* **2010**, *107*, 19784–19789.
[28] The PyMOL Molecular Graphics System, Version 1.7.4 Schrodinger, LLC, **2010**.
[29] T. P. C. Rooney, P. Filippakopoulos, O. Fedorov, S. Picaud, W. A. Cortopassi, D. A. Hay, S. Martin, A. Tumber, C. M. Rogers, M. Philpott, et al., *Angewandte Chemie* **2014**, *126*, 6240–6244.
[30] G. A. Ross, G. M. Morris, P. C. Biggin, *PLoS ONE* **2012**, *7*, e32036.
[31] A. P. Graves, R. Brenk, B. K. Shoichet, *J. Med. Chem.* **2005**, *48*, 3714–3728.
[32] K. C. Joshi, V. N. Pathak, S. Sharma, *J. Fluorine Chem.* **1986**, *32*, 299–307.
[33] J. Limanto, R. A. Desmond, D. R. Gauthier Jr., P. N. Devine, R. A. Reamer, R. P. Volante, *Org. Lett.* **2003**, *5*, 2271–2274.
[34] N. Desroy, A. Denis, C. Oliveira, D. Atamanyuk, S. Briet, F. Faivre, G. LeFralliec, Y. Bonvin, M. Oxoby, S. Escaich, et al., *J. Med. Chem.* **2013**, *56*, 1418–1430.
[35] P. Politzer, J. S. Murray, *Chemphyschem* **2013**, *14*, 278–294.

[36] L. A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, et al., *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 314–318.
[37] L. Pan, Y. Jiang, Z. Liu, X.-H. Liu, Z. Liu, G. Wang, Z.-M. Li, D. Wang, *Eur J Med Chem* **2012**, *50*, 18–26.
[38] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discovery Today* **2004**, *9*, 430–431.

Chapter 4:

Developing quantitative assays to assess ligand affinity for *Sd*Kef

4. Developing quantitative assays to assess ligand affinity for *Sd*Kef

4.1 Introduction and aims

As described in Chapter 3, Healy *et al.*[1] previously demonstrated the use of a competition fluorescence assay as a qualitative assessment of peptide ligand binding to the C-terminal domain of *Shewanella denitrificans* Kef (*Sd*KefCTD). However, when applied to non-peptidic ligands this assay was found to have limited applicability, presumably due to the ligands interfering with the fluorescent probe. It was, therefore, necessary to develop alternative quantitative assays for screening compounds against *Sd*KefCTD. A number of options were explored in parallel, namely: Surface Plasmon Resonance; AlphaScreen®; ^{1}H Carr-Purcell-Meiboom-Gill (CPMG) NMR; ^{19}F NMR; Microscale Thermophoresis and Fluorescence Polarisation. The aim was to have several reliable assays that could be run in tandem to screen a library of compounds. The availability of multiple assays would enable more efficient identification of false positives through the crosschecking of results.

4.2 Surface Plasmon Resonance (SPR)

All SPR studies were performed under the guidance of Dr. David Staunton from the Molecular Biophysics Suite, Department of Biochemistry, University of Oxford.

4.2.1 Introduction to Surface Plasmon Resonance (SPR)

Surface Plasmon Resonance (SPR) is an optical phenomenon used to perform a range of different binding studies through monitoring the intensity of polarised light undergoing total internal reflection at the interface of two media with opposing dielectric constants (Figure 4.1).[2] When

polarised light is reflected under the conditions of total internal reflection, surface plasmons (electron charge density waves) are excited. When the energy and momentum (along the interface) for the incident photons and the surface plasmon coincide, the intensity of the reflected light decreases due to the photons' energy being transferred to the charge density waves of the plasmon.[3] The angle at which this decrease in intensity occurs is called the resonance angle, a parameter that varies in accord with the refractive index. Increasing the mass at the surface of the interface causes a proportional increase in the refractive index, and so by immobilising a protein or reporter ligand to the surface, it is possible to obtain quantitative binding measurements through detecting changes in the resonance angle and thus the refractive index that are caused by a ligand or protein binding to the surface respectively. The modification of the refractive index caused by such binding events is represented on sensorgrams by changes in response units (RU).

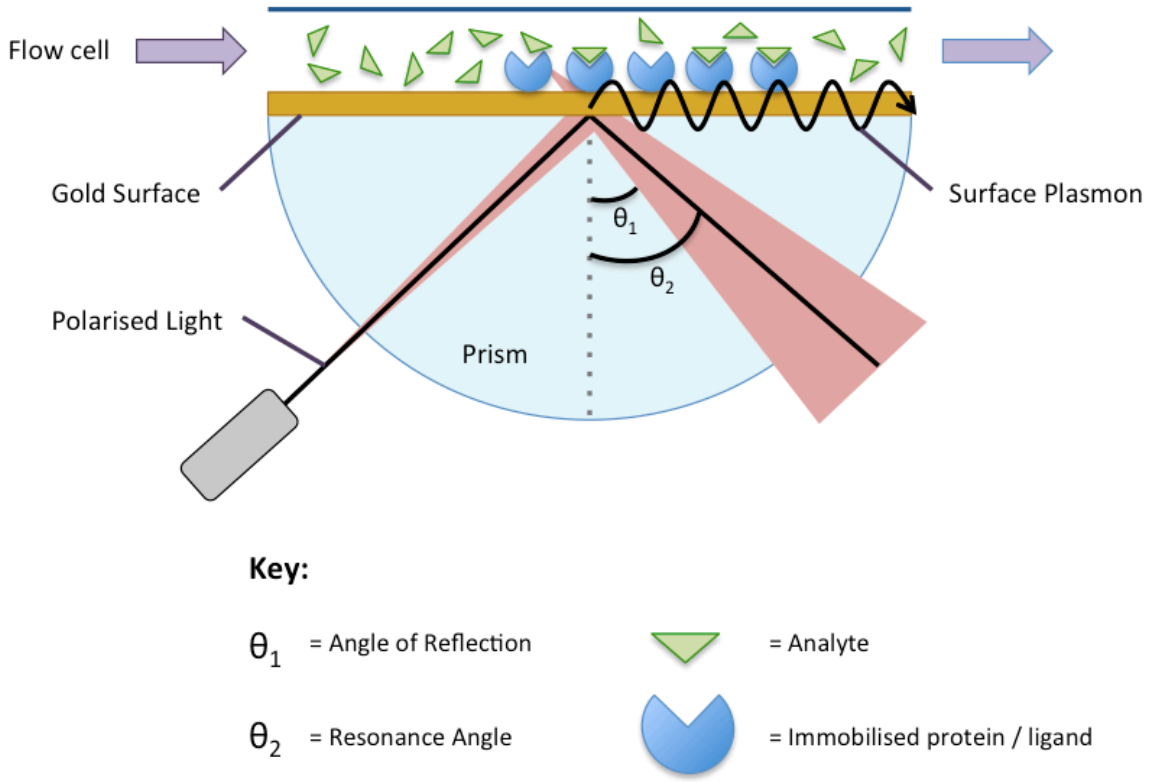


Figure 4.1 A visual representation of Surface Plasmon Resonance (SPR).

4.2.2 Nickel affinity SPR

Initial attempts at using SPR involved immobilisation of the His_6-tagged *Sd*KefCTD onto a Series S Biacore™ Sensor Chip NTA. This sensor chip is pre-immobilised with nitrilotriacetic acid (NTA) to enable the capture of His_6-tagged protein onto the surface *via* Ni^{2+}/NTA chelation. Once His_6-tagged *Sd*KefCTD was immobilised on the surface (as determined by an increase in response), the aim was to measure the surface dissociation constants (K_Ds), through detecting changes in the response, caused by ligand binding events to *Sd*KefCTD.[4] The glutathione S-conjugate **^{t}BuSG** was selected as the analyte for a trial run at obtaining K_D values, whereby the surface of immobilised *Sd*KefCTD was subjected to a serial dilution of **^{t}BuSG** by a factor of 0.5 from 200-0.05 μM. The initial step of adhering His_6-tagged *Sd*KefCTD showed some promise, with an overall increase in response (highlighted by the blue arrow in Figure 4.2 A), indicating increased mass at the surface, after the injection of *Sd*KefCTD (highlighted by the red arrow in Figure 4.2 A). The initial drop at the red arrow is caused by the injected sample being in a different buffer to the running buffer, altering the refractive index and triggering a drop in response. Once the injection has finished, the running buffer replaces the sample buffer causing the response to increase at the blue arrow to the original value plus the contribution from the immobilised protein.

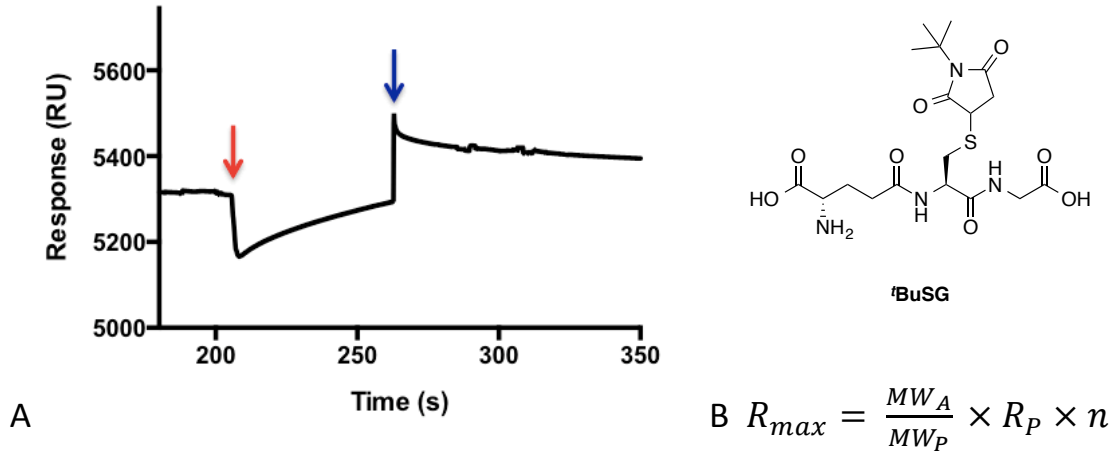


Figure 4.2 (A) A sensorgram demonstrating *Sd*KefCTD being loaded onto the chip. The red arrow indicates the time point at which the protein was injected and the blue arrow highlights the increase in response corresponding to protein adhering to the surface. (B) An equation for calculating the maximum theoretical response of the analyte for a given protein level (R_{max}), where R_P is the amount of protein immobilised, MW_A is the molecular weight of the analyte, MW_P is the molecular weight of the protein and n is the stoichiometry of the reaction. The structure of the glutathione S-conjugate **^{t}BuSG** is highlighted in the top right of the figure.

However, *Sd*KefCTD was observed to slowly wash off the surface after loading (blue arrow; Figure 4.2 A), which could complicate the interpretation of on and off rates due to the base line not remaining flat. Furthermore, the change in response units (RU) corresponding to *Sd*KefCTD adhering to the surface was only ≈ 120 RU, meaning that the maximum theoretical response (R_{max}; Figure 4.2 B) using **^{t}BuSG** as the analyte was ≈ 1.0 RU, assuming the stoichiometry of the reaction to be 1. In order to provide a good level of sensitivity for determining K_D values, the preferred maximum response caused by analyte binding is generally ≥ 100 RU.[5] To obtain an R_{max} = 100 RU, using **^{t}BuSG** as the analyte, the protein concentration on the surface needed to be increased to R_P ≈ 11500, which is ≈ 100 fold the original concentration loaded. Given that the protein was washing off the surface at the original concentration, it was advisable not to increase the protein concentration, as this would attenuate the gradient of the baseline. As expected, the titration experiment that was performed using **^{t}BuSG** as the analyte did not produce interpretable results, as any changes observed were within experimental error/noise.

4.2.3 Development and synthesis of biotinylated glutathione S-conjugates

One approach to increasing the magnitude of R_{max} is to tether the ligand to the chip's surface, and make the protein the analyte, so that the *MW* of the ligand and protein are switched in the equation for R_{max} in Figure 4.2 B. Attachment of a ligand to a chip can be achieved by coating the surface of the chip with streptavidin, and tagging the ligand with a biotin linker. It was therefore decided to tether biotin to a glutathione S-conjugate to investigate whether these types of studies were viable. Analysis of the X-ray crystal structure of the C-terminal domain of *Escherichia coli* KefC (*Ec*KefC) co-crystallised with **ESG** (PDB code: 3L9X)[6] suggested that both the Glu of the peptide backbone and *N*-ethylsuccinimido ring are pointing out into solvent, providing potential sites for the attachment of biotinylated linkers that will not disrupt *Sd*KefCTD binding (Figure 4.3 A & B). The Gly-COOH was ruled out as a point of attachment due to the significance of its binding interactions with R416, R516 and N551 in *Ec*KefC, as demonstrated by mutation studies.[7,8] Three possible sites for attachment were therefore identified: the Glu-NH_2; the Glu-COOH; the Cys thiol.

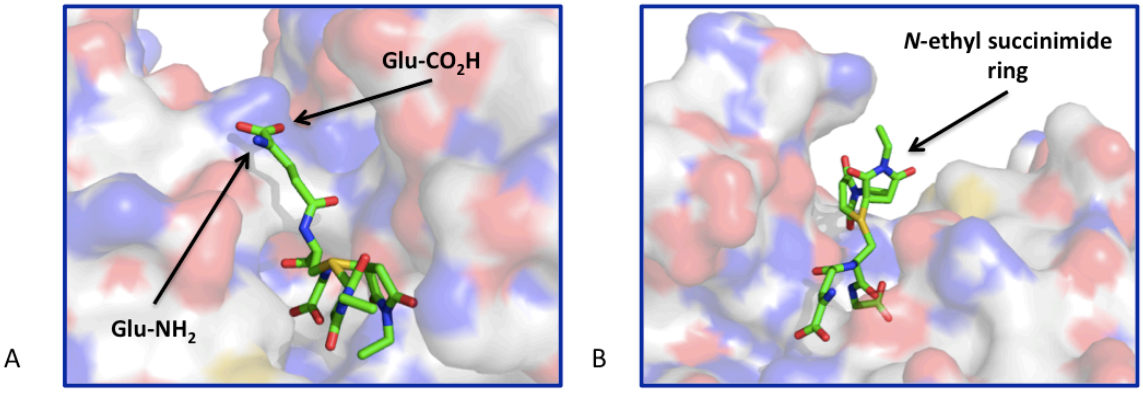
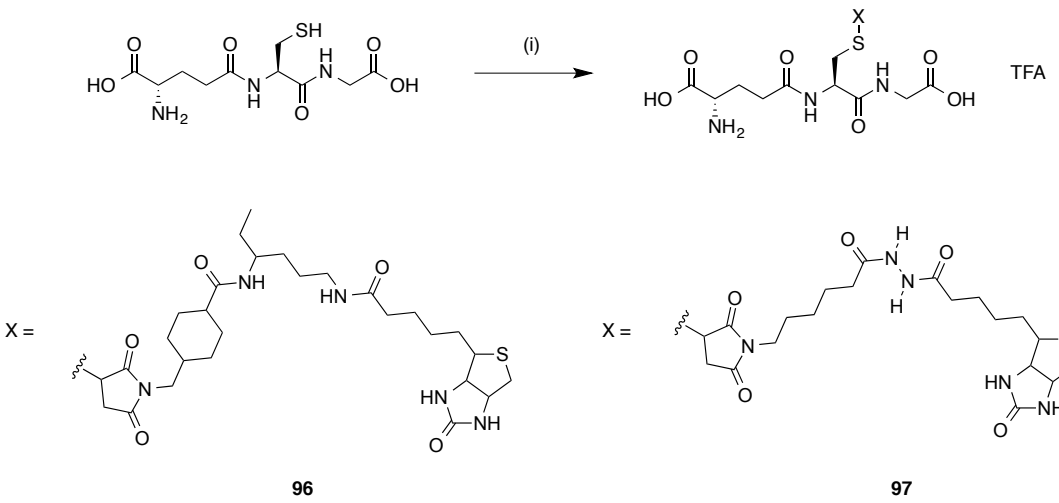


Figure 4.3 PyMOL[9] images of X-ray crystal structure of glutathione S-conjugate **ESG** (shown as sticks: C = green) co-crystallised with *Ec*KefC (PDB code: 3L9X)[6]. (A) demonstrates the solvent exposed Glu; (B) demonstrates the solvent exposed *N*-ethylsuccinimido ring.

The *N*-ethylsuccinimido ring was selected as the point of attachment, due to the synthetic tractability of coupling a biotinylated maleimide to the Cys, in a single conjugate addition step. Two commercially available biotinylated maleimides were thus employed with different linkers, in order to synthesise the biotinylated glutathione S-conjugates **96** and **97** (Scheme 4.1). The linker used to make **96** was selected as it has a bulky cyclohexyl group in close proximity to the maleimide; a feature that is important for Kef affinity in published structural studies on the succinimido ring of glutathione S-conjugates.[1] Compound **97** was synthesised using a linker that is simpler in structure, because this might help with tethering the conjugate to the streptavidin surface.



Scheme 4.1 Synthesis of biotinylated glutathione S-conjugates. *Reagents and conditions*: (i) For **96**: (1-biotinamido)-4-[4'-(maleimidomethyl)cyclohexanecarboxamido]hexane, RT, 49 h, 44%.; For **97**: biotinmaleimide, RT, 55 h, 31%.

The biotinylated glutathione S-conjugates were synthesised successfully by modifying the conditions used for **^{t}BuSG**, which involved removing sodium hydroxide and increasing the reaction time. The desired purified products **96** and **97** were isolated using semi-preparative HPLC in moderate yields of 44% and 31%, respectively (Scheme 4.1).

4.2.4 Binding of the biotinylated glutathione S-conjugates as determined using the competition fluorescence assay

As an initial qualitative assessment of binding, the biotinylated glutathione S-conjugates **96** and **97** were tested using the competition fluorescence assay (Chapter 3). The results suggested that both **96** and **97** were binding to *Sd*KefCTD more tightly than glutathione (K_D = 900 ± 200 μM; fluorescence emission spectra)[1] but less tightly than **^{t}BuSG** (K_D = 0.40 ± 0.2 μM; fluorescence emission spectra)[1] (Figure 4.4 A). Both of the biotinylated glutathione S-conjugates were checked for interference with the fluorescent **DNGSH** probe by running the experiment in the absence of *Sd*KefCTD. Neither of the biotinylated glutathione S-conjugates were found to be quenching **DNGSH** (Figure 4.4 B). It was therefore decided to proceed with kinetic studies using SPR to see if quantitative affinities could be obtained.

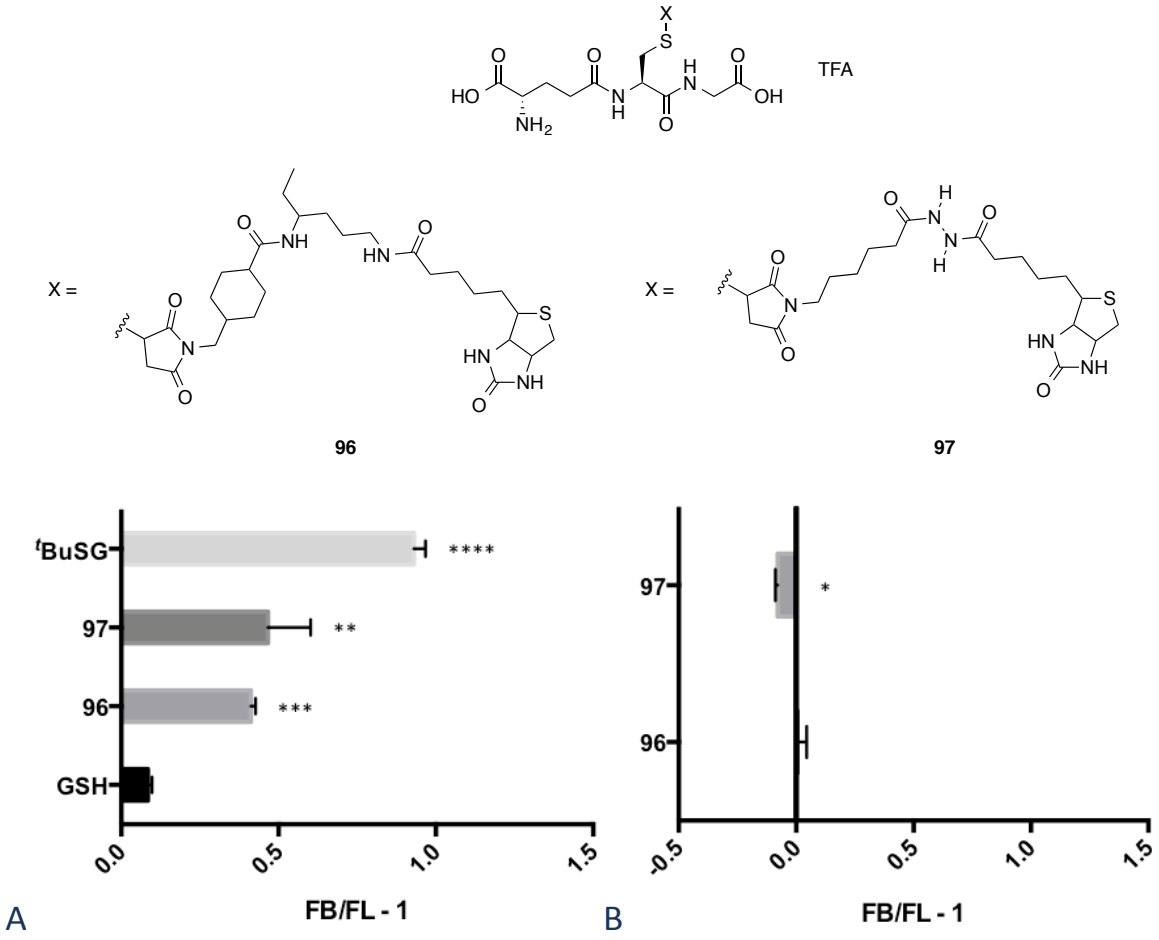


Figure 4.4 Analysis of the binding of the biotinylated glutathione S-conjugates using the competition fluorescence assay. (A) ($F_B/F_L - 1$) values of each biotinylated glutathione S-conjugate, **96** and **97**, (1 mM) relative to the positive control **^{t}BuSG** (1 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM); (B) ($F_B/F_L - 1$) values of each biotinylated glutathione S-conjugate, **96** and **97**, (1 mM) in the presence of **DNGSH** (5 µM) and absence of *Sd*KefCTD. n = 3, error bars indicate one standard deviation of uncertainty. Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$).

4.2.5 SPR with the biotinylated glutathione S-conjugates

The first step in this approach was to coat a CM5 sensor chip with streptavidin. This proceeded as expected (Figure 4.5), showing an increase in response of ≈ 7000 RU after the addition of streptavidin, and subsequent blocking of the surfaces with 1 M ethanolamine hydrochloride. With data in hand providing confidence that the streptavidin had been loaded onto the surface of

the chips, tethering of the biotinylated probes **96** and **97** to separate cells was investigated. Each probe was injected multiple times until a desirable increase in response of ≈ 220 RU and ≈ 180 RU was attained for **96** and **97** respectively. The observed increase in response has been demonstrated for the biotinylated probe **97** in the sensorgram in Figure 4.5 B.

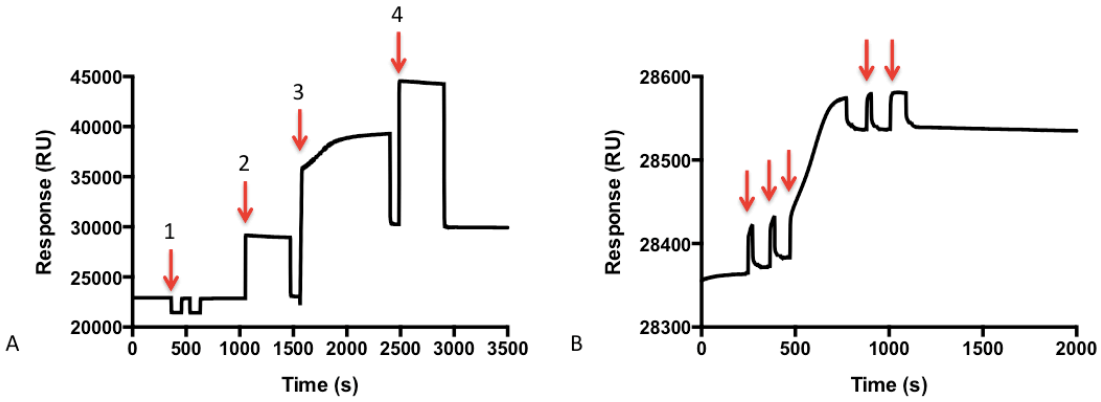


Figure 4.5 (A) A sensorgram demonstrating streptavidin being coated on the CM5 chip. Arrows indicate: 1) Injection of NaOH (50 mM; 2 × 8 µL; 5 µL/min); 2) Injection of NHS/EDC (420 s); 3) Injection of streptavidin (100 µg/mL) in sodium acetate (10 mM; pH 4.5; 840 s; 5 µL/min; 4) Injection of ethanolamine hydrochloride (1 M; pH 8.5; 420 s; 5 µL/min); (B) Shows a sensorgram demonstrating the increase in response upon addition of the biotinylated probe **97**. Arrows indicate time points at which **97** was injected.

Having successfully loaded the biotinylated glutathione S-conjugates **96** and **97** onto the streptavidin coated CM5 sensor chips; attempts were made to obtain dissociation constants of the probes using *Sd*KefCTD as the analyte. An initial serial dilution of *Sd*KefCTD (factor 0.5; 100-0.10 µM) was performed. Although binding of *Sd*KefCTD, which appeared to be concentration dependent, was observed on the flow cells for both probes **96** and **97** (Figure 4.6 B & E), binding of *Sd*KefCTD was also observed on the reference cells (Figure 4.6 A & D). These binding events occurred to a greater extent on the reference cell, such that a negative response was measured for both **96** and **97** when the reference cells were subtracted from the flow cells (Figure 4.6 C & F). The implication of this result is that *Sd*KefCTD was binding non-specifically, possibly forming interactions with the streptavidin. Furthermore, the R_{max} for **96** and **97** when *Sd*KefCTD is the analyte were calculated to be ≈ 11800 RU and

≈ 12500 RU respectively, assuming a stoichiometry of 1. Given that we were only seeing negative responses, it was clear that the protein was not binding selectively to the biotinylated conjugates to the extent predicted by the R_{max}. It was therefore apparent, with these data in hand, that this approach to obtaining quantitative kinetic data was not suitable.

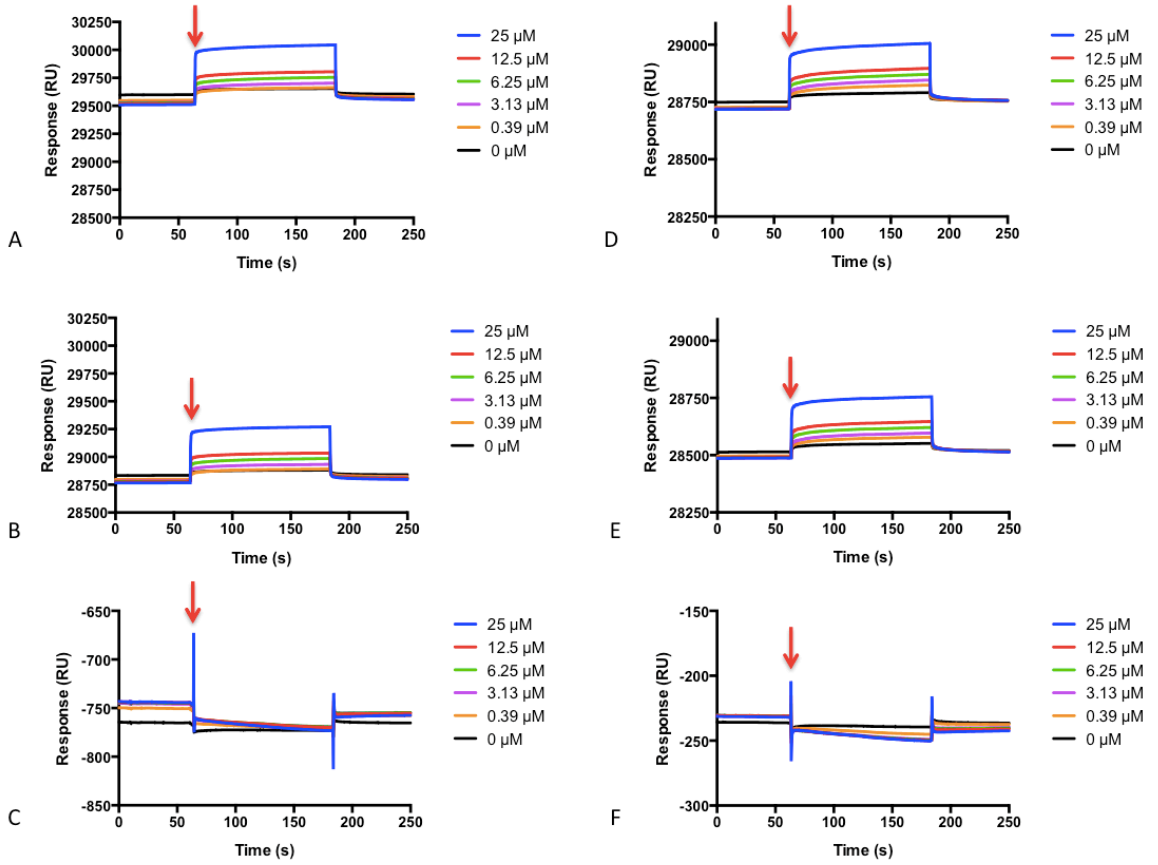


Figure 4.6 Sensorgrams demonstrating protein titration onto the chips loaded with biotinylated glutathione S-conjugates **96** and **97**, arrows indicate the time point at which protein was injected; (A) The reference cell for **96** (cell 1); (B) The flow cell for **96** (cell 2); (C) The reference - flow for **96** (2-1); (D) The reference cell for **97** (cell 3); (E) The flow cell for **97** (cell 4); (F) The reference cell subtracted from the flow cell for **97** (4-3).

4.3 AlphaScreen® with the biotinylated glutathione S-conjugates

4.3.1 Introduction to AlphaScreen®

Having established that SPR would not be a suitable assay for the generation of reliable quantitative data, the viability of applying the biotinylated glutathione S-conjugates to an alternative quantitative assay called AlphaScreen® was assessed (Figure 4.7).[10] AlphaScreen® employs the conjugation of donor and acceptor beads to a protein and a reporter ligand. The protein of interest (*Sd*KefCTD) is attached to Ni chelate acceptor beads *via* a His_6-tag and the reporter ligands (biotinylated glutathione S-conjugates **96** and **97**) are attached to streptavidin donor beads. The assay is run by the excitation of the donor beads at 680 nm, a process that results in the production of singlet oxygen. The singlet oxygen produced is only able to transfer its energy to acceptor beads that have proximity < 200 nm from the donor bead, due to its lifetime of 4 μs in aqueous solutions. Transfer of energy from singlet oxygen to the acceptor beads causes them to emit light in the range 520-620 nm. The light emitted in this range can be detected to quantify the proportion of reporter ligand bound to the protein. Quantitative binding data can therefore be obtained through the titration of ligands that compete with the reporter ligand for the protein binding site, resulting in a reduction in fluorescence.[10]

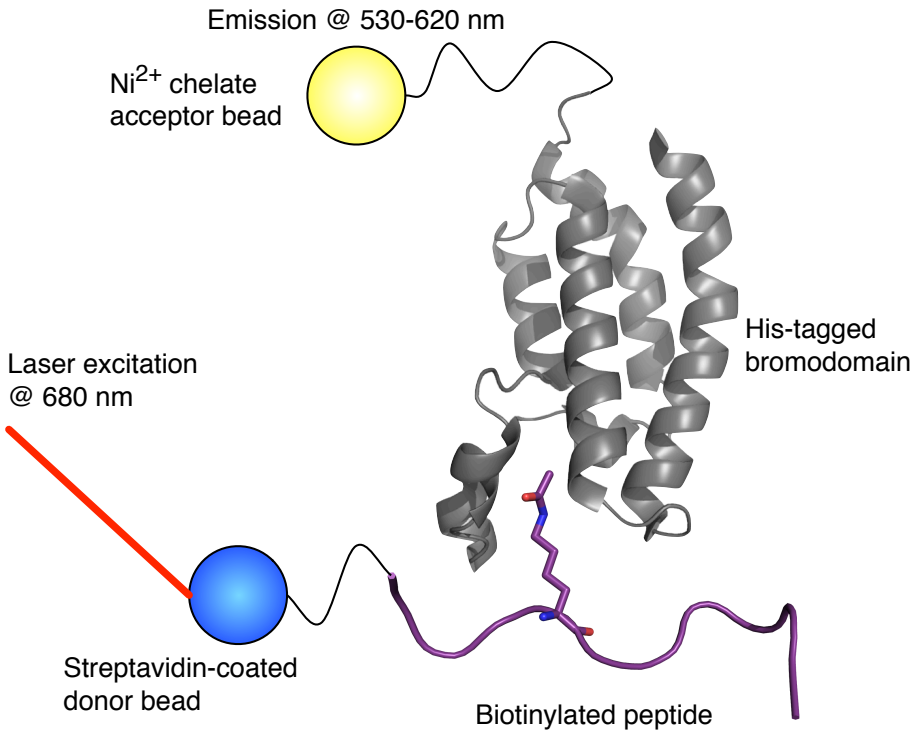


Figure 4.7 A visual representation of the AlphaScreen®, using a His-tagged bromodomain as an example. Excitation of the Donor bead (blue) by a laser at 680 nm results in the generation of singlet oxygen. Singlet oxygen can travel 200 nm is aqueous solution before decaying and so Acceptor beads (yellow) within this proximity can be excited by the singlet oxygen. Excited acceptor beads emit in the range: 530-620 nm. Image provided by Prof. Stuart Conway.

4.3.2 Trial conditions for AlphaScreen®

Trial conditions were selected with the aid of an AlphaScreen® start-up guide from Perkin Elmer®, which recommended using the acceptor and donor beads at a final concentration of ~20 µg/mL, and having both the protein and probe at final concentrations of 0-300 nM. The conditions used in the trial screen were as follows: both the protein and probe were at final concentrations of 40 nM and 200 nM; both the acceptor and donor bead were at final concentrations of 17.5 µg/mL; final well volumes of 20 µL were used. As positive controls, a biotinylated His_6 was employed to determine the maximum emission intensity expected, and the competing ligands, glutathione (**GSH**; K_D = 900 ± 200 µM; fluorescence emission spectra)[1] and **^{t}BuSG** (K_D = 0.40 ± 0.2 µM; fluorescence emission spectra)[1] were run at final concentrations of 62.5 µM

with the protein and probe to see if a decrease in the emission intensity could be observed that corresponded with the extent that the biotinylated probes were displaced from *Sd*KefCTD. The heat map in Figure 4.8 highlights the outcome of this trial run.

		Buffer + *Sd*KefCTD	Probe + *Sd*KefCTD	^{t}BuSG (62.5 µM) + Probe + *Sd*KefCTD	GSH (62.5 µM) + Probe + *Sd*KefCTD
200 nM final concentration of protein and probe	Probe 96	2094	1737	1708	1791
	Probe 97	2137	1921	1744	1883
40 nM final concentration of protein and probe	Probe 96	1714	1610	1589	1555
	Probe 97	1638	1397	1524	1274

Negative Control

Biotinylated His_6
56020

Positive Control

Figure 4.8 Heat map of the emission intensities obtained in the AlphaScreen® under various trial conditions. Red represents high emission intensity and green represents low emission intensity. An average of triplicate values is shown.

The results of the trial run suggested that neither of the biotinylated probes were binding to *Sd*KefCTD in the AlphaScreen® (Figure 4.8). The emission intensity of the positive control, biotinylated His_6, was over 30 fold in magnitude in comparison to all the other runs. Furthermore, the emission intensity of the negative control (Buffer + *Sd*KefCTD), where no probe was present, was comparatively higher than the run when probe was present (Probe + *Sd*KefCTD), suggesting that the probes were not eliciting the desired increase in emission intensity through binding to the protein. There was also no statistically significant difference (as determined by the Student's *t*-test) between the runs with just probe and protein, to when **^{t}BuSG** and glutathione were present (with the exception of 200 nM of **97** + *Sd*KefCTD + **^{t}BuSG**). This result suggests that the slight differences in the emission intensity were just noise. Although there was the possibility of trying to optimise the conditions of the AlphaScreen®, for

example by altering the order of addition of each component or the incubation time, it was decided not to pursue this avenue due to successes in other assays being developed in parallel.

4.4 "Pull-down" study with the biotinylated glutathione S-conjugates

It was hypothesised that the biotinylated glutathione S-conjugates, despite being observed to bind in the competition fluorescence assay, were not able to bind to *Sd*KefCTD once the biotin and streptavidin had bound, due to unfavourable interactions between the proteins. To investigate this theory, and assess the binding of the biotinylated glutathione S-conjugates to *Sd*KefCTD once they had bound to streptavidin, a "pull-down" study was performed using Dynabeads® MyOne™ Streptavidin C1 magnetic beads. Figure 4.9 shows the results of this study, which found that both of the biotinylated glutathione S-conjugates, **96** (Figure 4.9 A) and **97** (Figure 4.9 B), were unsuccessful at pulling-down *Sd*KefCTD (run on its own in lane 2 of Figure 4.9 A & B). *Sd*KefCTD was observed in the supernatant after incubation with **96**/**97** and the Dynabeads® for 2 h at 0 °C (lane 3 of Figure 4.9 A & B), but not in the denatured sample of the Dynabeads® (lane 4 of Figure 4.9 A & B). This result suggested that *Sd*KefCTD was not binding to the biotinylated probes, an outcome that was exactly the same as the negative control of *Sd*KefCTD + Dynabeads® without **96**/**97** (lanes 5 and 6 of Figure 4.9 A and B) and the positive control of *Sd*KefCTD + **96**/**97** + Dynabeads® + 10 eq. of **^{t}BuSG** (lanes 7 and 8 of Figure 4.9 A & B). A positive control of *Sd*KefCTD + the biotinylated maleimides used to synthesise **96**/**97** + Dynabeads® was run (lanes 9 and 10 of Figure 4.9 A & B) to check the beads were intact, as protein should be pulled down through conjugation of one of *Sd*KefCTD's cysteines to the biotinylated maleimides. The maleimides were observed to pull down the protein, as can be observed from the faint bands corresponding to mass of *Sd*KefCTD in lane 10 of Figure 4.9 A & B, suggesting that the beads were functioning. The overall implication of these results was that the

interaction of the streptavidin-biotinylated glutathione S-conjugate complex with *Sd*KefCTD was too weak to be exploited by SPR and AlphaScreen®.

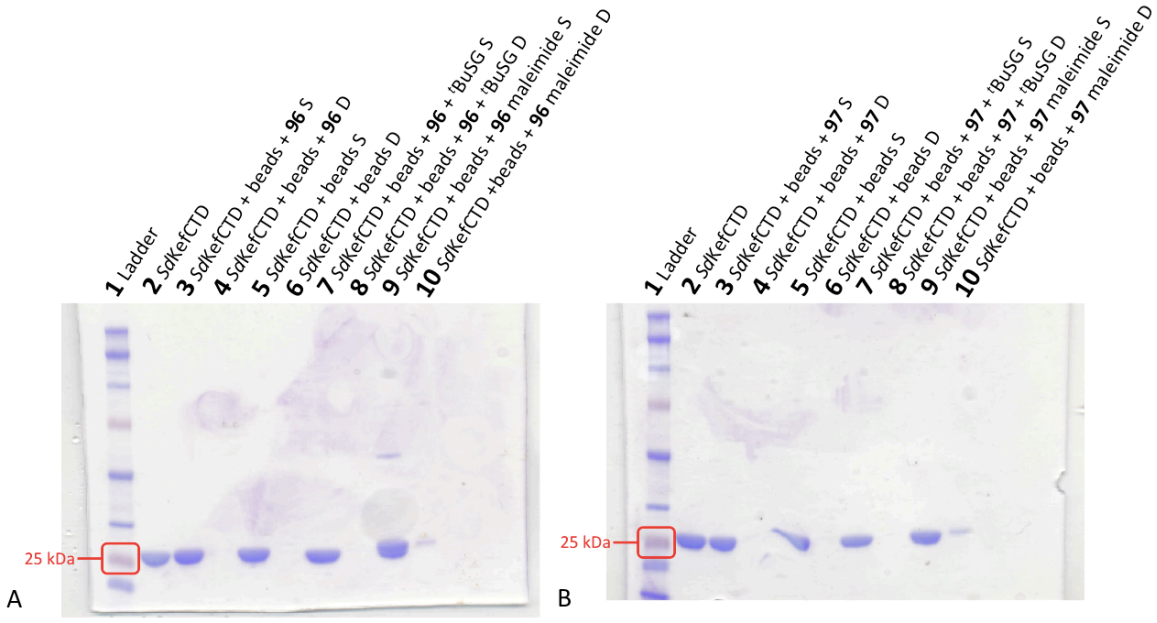


Figure 4.9 "Pull-down" analysis of *Sd*KefCTD using biotinylated probes **96** (A) and **97** (B) using Dynabeads® MyOne™ Streptavidin C1. S refers to the supernatant before denaturation; D refers to after denaturation. Lanes correspond to the following: 1) Ladder – Precision Plus Protein™ Dual Colour Standard #161-0374; 2) *Sd*KefCTD alone; 3) [*Sd*KefCTD + **96**/**97** + Dynabeads®] supernatant before denaturation; 4) [*Sd*KefCTD + **96**/**97** + Dynabeads®] after denaturation; 5) [*Sd*KefCTD + Dynabeads®] supernatant before denaturation; 6) [*Sd*KefCTD + Dynabeads®] after denaturation; 7) [*Sd*KefCTD + **96**/**97** + Dynabeads® + 10 eq. **^{t}BuSG**] supernatant before denaturation; 8) [*Sd*KefCTD + **96**/**97** + Dynabeads® + 10 eq. **^{t}BuSG**] after denaturation; 9) [*Sd*KefCTD + biotinylated maleimides use to make **96**/**97** + Dynabeads®] supernatant before denaturation; 10) [*Sd*KefCTD + biotinylated maleimide **96**/**97** + Dynabeads®] after denaturation.

4.5 SPR and AlphaScreen® conclusions

Although each of the biotinylated probes **96** and **97** were observed to bind qualitatively by the competition fluorescence assay, it was not possible to quantify their affinity to *Sd*KefCTD by either SPR or AlphaScreen®. The lack of success in both of these assays can be rationalised as

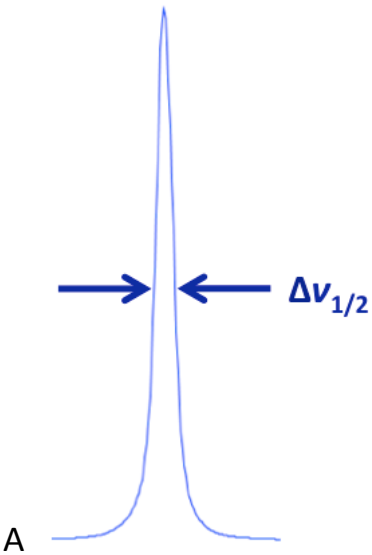
follows: the interaction between the biotinylated probes and *Sd*KefCTD may be too weak to be exploited by SPR and AlphaScreen®; once the biotinylated probes bind to the steptavidin beads they are no longer able to bind to *Sd*KefCTD; the results of the competition fluorescence assay could be misleading, and the biotinylated probes are not actually binding. Future work on these assays would therefore look to improve the affinity of the biotinylated probes for *Sd*KefCTD by either tethering the biotin to the Glu of the glutathione S-conjugate **^{t}BuSG**, as attachment here might not alter the affinity as much. Alternatively, the length of the linker could be increased in case the problem is that the biotinylated conjugates no longer bind once they have conjugated to the streptavidin beads.

4.6 Binding studies using ^{1}H CPMG NMR

All ^{1}H CPMG NMR data were obtained in collaboration with Amjad Khan in the Claridge Group, University of Oxford.

4.6.1 Introduction to ^{1}H CPMG NMR

An approach to detecting ligand-protein binding that is quantitative and free from false positive artifacts caused by fluorescence is a one-dimensional relaxation-edited ^{1}H NMR method outlined by Hajduk *et al.*[11] This technique allows direct observation of ligand binding to a protein, through monitoring the attenuation of ligand signals upon addition of protein. Formation of a ligand-protein complex causes the bound ligand to move and tumble at a slower rate, which consequently results in the ligand signals broadening due to the ligand having a higher ^{1}H spin-spin relaxation rate (shorter T_2^*) in the bound state.[12] The relationship between the time constant for transverse relaxation (T_2^*) and resonance broadening, which can be quantified using the half-height linewidth ($\Delta\nu_{1/2}$), is outlined in Figure 4.10 A & B.[13]



A B $\Delta\nu_{1/2} = \frac{1}{\pi T_2^*}$ C $\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{2\,(\Delta B_o)}}$

Figure 4.10 Demonstrates the effect that the transverse relaxation time T_2 has on the half-height linewidth ($\Delta\nu_{1/2}$) of a resonance. (A) A visual representation of a half-height linewidth ($\Delta\nu_{1/2}$) of a resonance; (B) The equation for the relationship between $\Delta\nu_{1/2}$ and T_2^*, where T_2^* is the combination of genuine transverse relaxation processes and relaxation caused by magnetic field inhomogeneity; (C) shows an equation highlighting the contributions of genuine relaxation processes, T_2, and relaxation processes from field inhomogeneity, $T_{2\,(\Delta B_o)}$, to T_2^*.[13]

The relaxation editing that is implemented in this technique is the Carr-Purcell-Meiboom-Gill (CPMG) sequence.[11,14,15] CPMG is a spin-echo sequence that looks to remove contributions from magnetic field inhomogeneity ($T_{2\,(\Delta B_o)}$) to T_2^* that arise from instrumental imperfections, so that the attenuation of signals upon addition of protein is enhanced (Figure 4.10 C).[13]

4.6.2 Validation of ^{1}H CPMG NMR

To validate the applicability of ^{1}H CPMG NMR to *Sd*KefCTD, a number of control experiments were performed using **^{t}BuSG** as the reporter ligand, as highlighted in Figure 4.11 **^{t}BuSG** was an ideal ligand for use in these studies as the ^{t}Bu signal does not overlap with any of the protein signals and integrates for nine protons, making it easier to track the ligand at low concentrations.

The first experiment performed was to run 10 μM of **^{t}BuSG** in the presence and absence of 20 μM of wild-type *Sd*KefCTD, and compare the intensity of the ^{t}Bu signal in both (Figure 4.11 B1-2 & C1-2) to quantify the extent of the signal broadening / attenuation in the presence of the protein. A 90% decrease in the intensity of the ^{t}Bu signal was observed in the presence of wild-type *Sd*KefCTD, relative to when the protein was absent. To determine whether this decrease was a result of specific binding to the glutathione binding-site of *Sd*KefCTD two controls were run. The first was to see if the ^{t}Bu signal intensity could be restored by displacing the **^{t}BuSG** bound to wild-type *Sd*KefCTD with 3 mM of the glutathione S-conjugate **ESG** (Figure 4.11 B3 & C3). The result of this control was promising as after adding **ESG** the ^{t}Bu signal intensity was returned to 70% of what it was in the absence of wild-type *Sd*KefCTD. The second control involved using the inactive mutant, Q419K,[1] which does not to bind glutathione S-conjugates,[1,6] to investigate whether the loss of signal intensity was a result of non-specific binding to the protein, or due to the viscosity of the sample being increased by the protein, which could lead to slower tumbling (Figure 4.11 B4 & C4). Comparing the signal intensity of the ^{t}Bu group of 10 μM of **^{t}BuSG** in the presence and absence of 20 μM of the *Sd*KefCTD mutant Q419K found that there was no statistically significant difference and, therefore, that the decrease in signal intensity observed when **^{t}BuSG** was in the presence of wild-type *Sd*KefCTD was a result of specific binding.

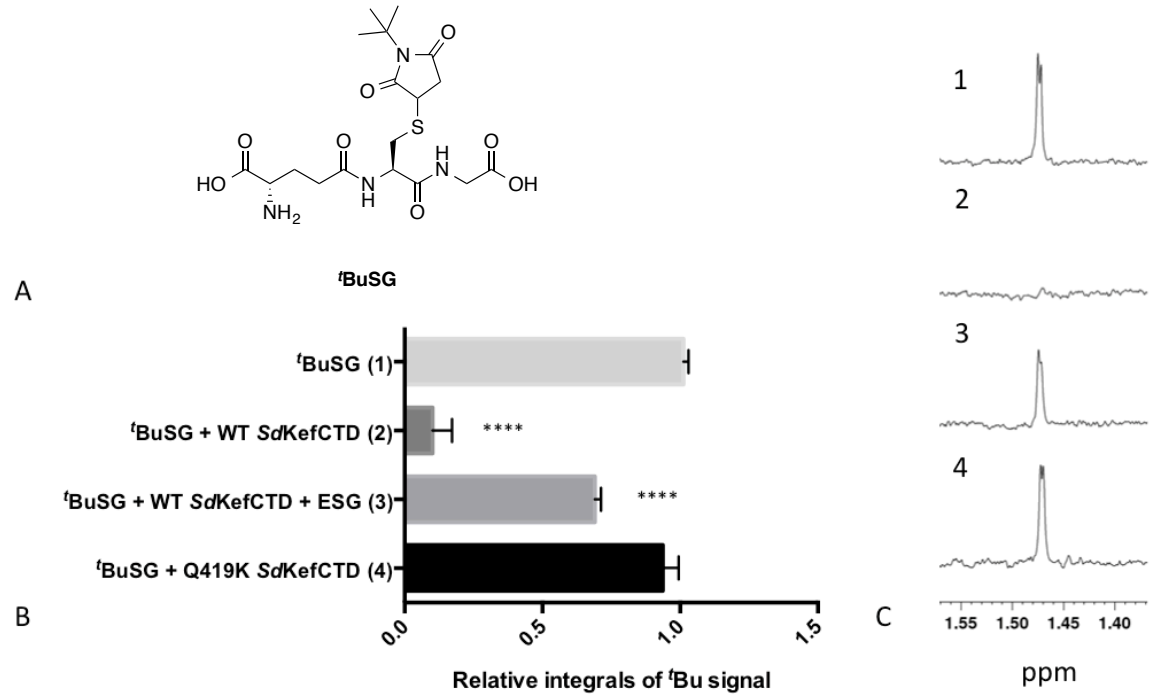


Figure 4.11 Control experiments performed to validate ^{1}H CPMG NMR for use with *Sd*KefCTD. (A) The structure of the glutathione S-conjugate **^{t}BuSG**, which was used as the reporter ligand in these experiments; (B) The integrals of the ^{t}Bu signal of **^{t}BuSG** in each of the following experiments relative to 1: (1) **^{t}BuSG** (10 µM); (2) **^{t}BuSG** (10 µM) + WT *Sd*KefCTD (20 µM); (3) **^{t}BuSG** (10 µM) + WT *Sd*KefCTD (20 µM) + **ESG** (3 mM); (4) **^{t}BuSG** (10 µM) + Q419K *Sd*KefCTD (20 µM); (C) The ^{t}Bu signal for **^{t}BuSG** for each of the experiments 1-4. n = 3, error bars indicate one standard deviation of uncertainty. Significance of changes in experiments 2-4 relative to experiment 1 evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). It should be noted that the ^{t}Bu signal in (C) appears as two singlets due to **^{t}BuSG** being tested as a mixture of diastereomers.

Having demonstrated that it was possible to apply ^{1}H CPMG NMR to the *Sd*KefCTD system, a dissociation constant for **^{t}BuSG** was obtained using ^{1}H CPMG NMR, which was compared to values reported in the literature (Figure 4.12). A K_D of 540 ± 33 nM for **^{t}BuSG** was obtained using ^{1}H CPMG NMR, assuming single site binding, which is comparable to the reported affinity measured by fluorescence emission spectra: K_D = 400 ± 200 nM.[1]

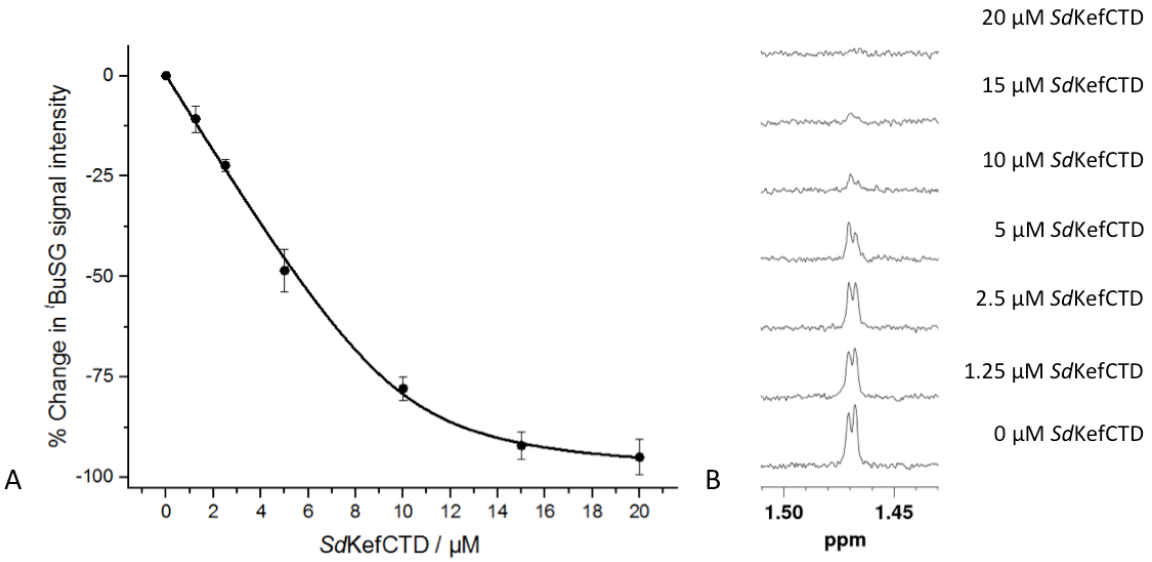


Figure 4.12 Analysis of the binding of **^{t}BuSG** by ^{1}H CPMG NMR. (A) Shows the saturation-binding curve of **^{t}BuSG** by ^{1}H CPMG NMR. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1; (B) Highlights the attenuation of the **^{t}BuSG** signal intensity upon titration of *Sd*KefCTD. It is worth noting that the ^{t}Bu signal in (B) appears as two singlets due to **^{t}BuSG** being tested as a mixture of diastereomers.

4.7 Binding studies using a ^{19}F NMR assay

All ^{19}F NMR data were obtained in collaboration with Amjad Khan in the Claridge Group, University of Oxford.

4.7.1 Introduction to ^{19}F NMR ligand-protein binding studies

^{1}H CPMG NMR provided a valuable approach to obtaining quantitative data for individual ligands, however, it was not going to be possible to apply it to a high-throughput screening set up, whereby multiple ligands are simultaneously screened in a single NMR tube. Problems could arise from the overlap of ligand signals with each other and signals from the protein, leading to complex interpretations of the ^{1}H NMR spectra obtained. Furthermore, the ligands screened might not have a suitably intense signal to track, as was the case with **^{t}BuSG**. Increasingly popular

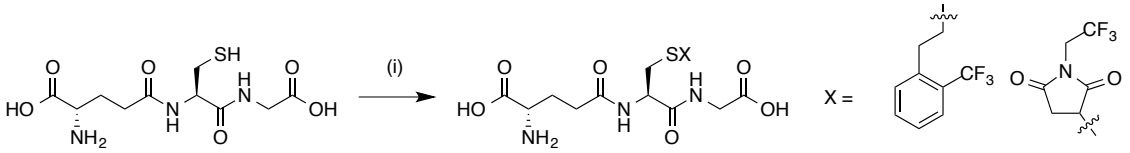
alternatives to ^{1}H NMR screening methods are ^{19}F NMR approaches, due to the advantages resulting from using ^{19}F over ^{1}H, for example: the natural abundance of ^{19}F is 100%; ^{19}F has a sensitivity of 83% relative to ^{1}H; ^{19}F is not naturally abundant in proteins, removing any background interference from the protein; ^{19}F chemical shifts are sensitive to the surrounding environment, making it possible to distinguish the bound and unbound states; the chemical shift dispersion of ^{19}F is large (~200 ppm), making screening libraries of fluorinated compounds possible without significant chemical shift overlap; ^{19}F can replace ^{1}H in ligands with minimal impact on biological activity.[16]

We therefore wanted to develop an ^{19}F NMR assay that applies the same principles used in ^{1}H CPMG NMR, whereby the ^{19}F signal of a fluorinated reporter ligand would undergo line broadening upon the addition of protein, an approach demonstrated by Tengel *et al.*[16] and Dalvit *et al.*[17] A high throughput screen could then be performed to identify ligands that are able to displace the fluorinated reporter ligand from *Sd*KefCTD and restore the ^{19}F signal, in a similar fashion to the FAXS method implemented by Dalvit *et al.*[17]

4.7.2 Synthesis of fluorinated glutathione S-conjugates

Two fluorinated probes were synthesised in parallel for use in the ^{19}F NMR assay. The first was *S*-(2-trifluoromethylbenzyl) glutathione **98**, a glutathione S-conjugate that had previously been made by Dr. Jess Healy, using a photochemical thiol-ene reaction to couple glutathione with 2-trifluoromethyl styrene. The reason *S*-(2-trifluoromethyl benzyl) glutathione, **98**, was synthesised is because Dr. Jess Healy had shown that it was able to bind to *Sd*KefCTD using both the competition fluorescence assay[1] and differential scanning fluorimetry[18] (unpublished data). The second fluorinated probe was *S*-*N*-2,2,2-trifluoroethylsuccinimido glutathione, **99**, which was synthesised by the conjugate addition of glutathione to 1-(2,2,2-trifluoroethyl)-2,5-dihydro-1*H*-

pyrrole-2,5-dione (Scheme 4.2). Compound **99** was synthesised because quantitative data had not been obtained for **98**, and there was more certainty that **99** would be able to bind with a reasonable affinity, due to it being a fluorinated analogue of the known glutathione S-conjugate **ESG** (K_D = 12 ± 3 μM; fluorescence emission spectra).[1]



Scheme 4.2 Synthesis of *S*-2-(2-(trifluoromethyl)phenyl)ethyl L-glutathione 2,2,2-trifluoroacetate, **98**, and *S*-*N*-2,2,2-trifluoroethylsuccinimido glutathione 2,2,2-trifluoroacetate, **99**; *Reagents and conditions*: (i) For **98**: 2-Trifluoromethyl styrene, 2,2-dimethoxyphenyl acetophenone, THF/H_2O, RT, hν , 5 h, 20%; For **99**: 1-(2,2,2-Trifluoroethyl)-2,5-dihydro-1*H*-pyrrole-2,5-dione, NaOH, RT, 9.5 h, 84%.

4.7.3 Validation of binding of the fluorinated glutathione S-conjugates by competition fluorescence and ^{1}H CPMG NMR

Before progressing the fluorinated ligands **98** and **99** to the ^{19}F NMR assay development stage, they were tested qualitatively using the competition fluorescence assay to confirm, or reconfirm in the case of **98**, that they bound specifically in the glutathione binding-site. The competition fluorescence assay showed that they both bound. As expected, **99** bound with an affinity comparable to the glutathione S-conjugate **ESG** (Figure 4.13).

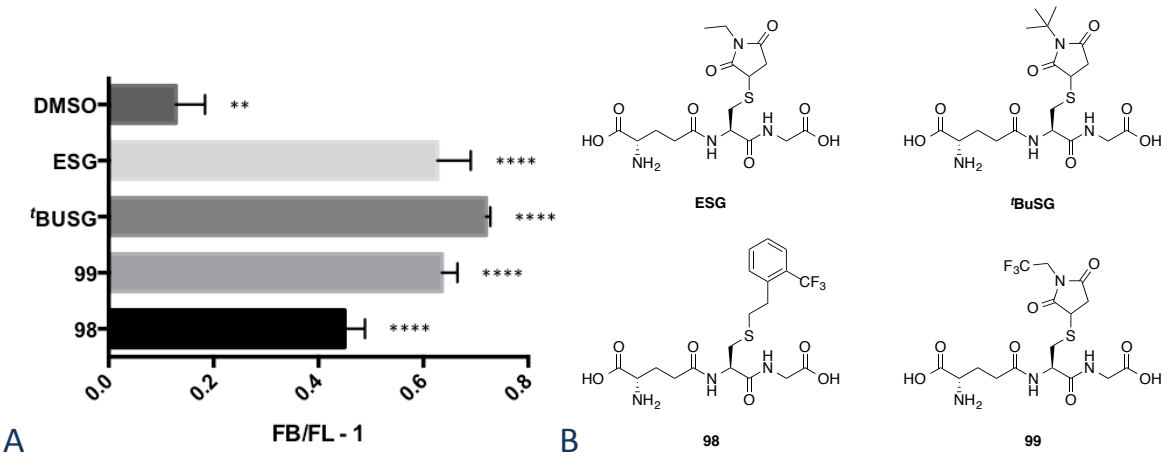


Figure 4.13 Analysis of the binding of the fluorinated glutathione S-conjugates by competition fluorescence assay. (A) Shows (F_B/F_L – 1) values of each fluorinated glutathione S-conjugate (1 mM) relative to the positive control **^{t}BuSG** (1 mM) and **ESG** (1 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) Structures of the compounds tested.

Having confirmed that both the probes bound using the competition fluorescence assay, a quantitative assessment of their binding was performed using ^{1}H CPMG NMR, which determined K_D values of 1.8 ± 0.3 µM and 4.1 ± 0.4 µM for **98** and **99**, respectively assuming single site binding (Figure 4.14). These affinities were deemed appropriate for performing the desired ^{19}F NMR studies and so both **98** and **99** were progressed.

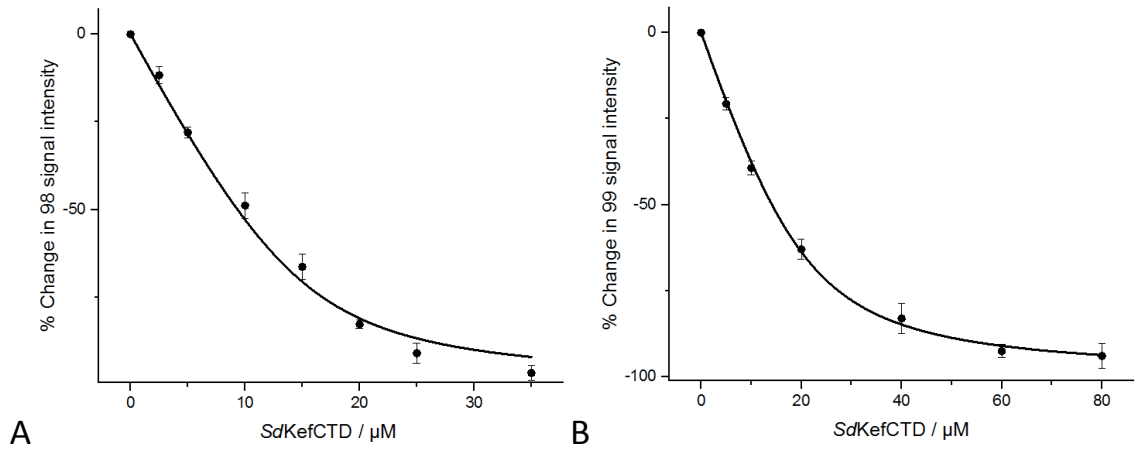


Figure 4.14 Saturation binding curves of **98** and **99** obtained by ^{1}H CPMG NMR. (A) The binding curve for **98**, from which a K_D of 1.8 ± 0.3 µM was extracted; (B) The binding curve for **99**, from which a K_D of 4.1 ± 0.4 µM was extracted. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1.

4.7.4 Development of the ^{19}F NMR assay using the fluorinated glutathione S-conjugates

As it was intended to use the ^{19}F NMR assay to perform high throughput screening, which would involve queuing samples overnight, Amjad Khan checked the stability of each of the probes using a time course experiment. The stability was checked because false positives could arise from the signal intensity of the fluorinated probes, **98** and **99**, decreasing due to the compounds breaking down. **98** was found to be stable over the time periods investigated (Figure 4.15 A), however, **99** was not, with the appearance of a new signal, and almost complete break down after 15 h (Figure 4.15 B). Furthermore, the peak for **99** at time = 0 h was much broader than **98** due to **99** being a mixture of diastereomers, resulting in a poor signal to noise ratio. Compound **99** was therefore ruled out for further use in the ^{19}F NMR studies and only **98** was progressed further.

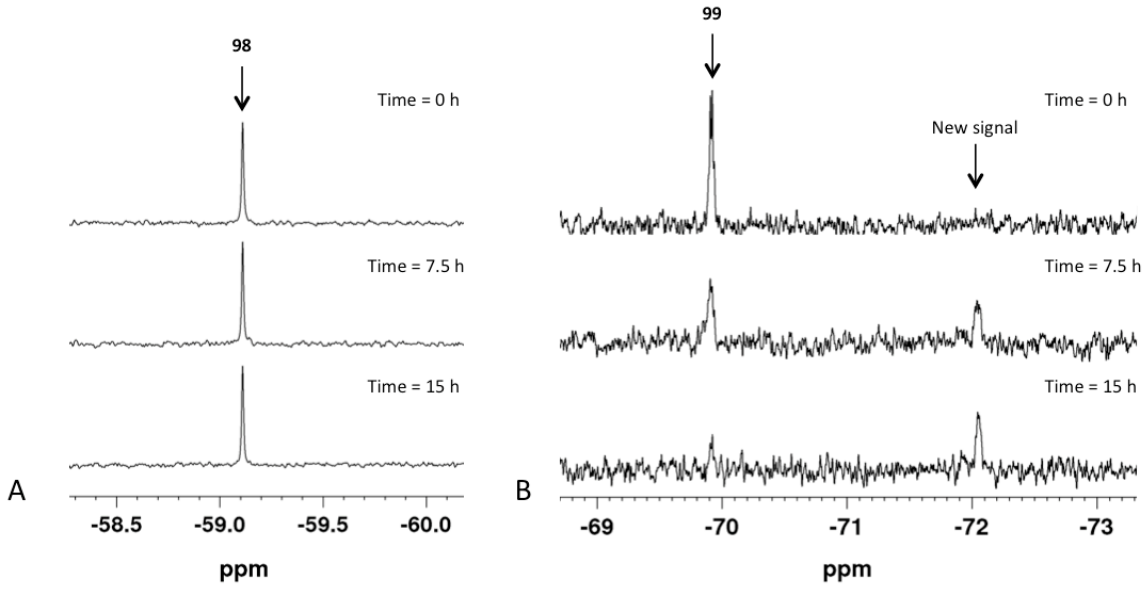


Figure 4.15 ^{19}F NMR time course experiment investigating the stability of the fluorinated ligands **98** and **99**. (A) The signal corresponding to the CF_3 of **98** (7 μM); (B) The signal corresponding to the CF_3 of **99** (7 μM).

The ^{19}F NMR assay was developed with the same principles of ^{1}H CPMG NMR in mind, namely that the NMR resonances of a bound ligand are attenuated and / or broadened due to faster nuclear spin transverse relaxation in the protein-ligand complex. An initial titration study of *Sd*KefCTD with **98**, performed by Amjad Khan, found that not only, did the signal corresponding to unbound **98** at −59.07 ppm attenuate, but a new broad peak appeared at −59.74 ppm when recording standard ^{19}F NMR spectra (Figure 4.16). Furthermore, increasing the concentration of *Sd*KefCTD caused the intensity of the new broad peak at −59.74 ppm to increase.

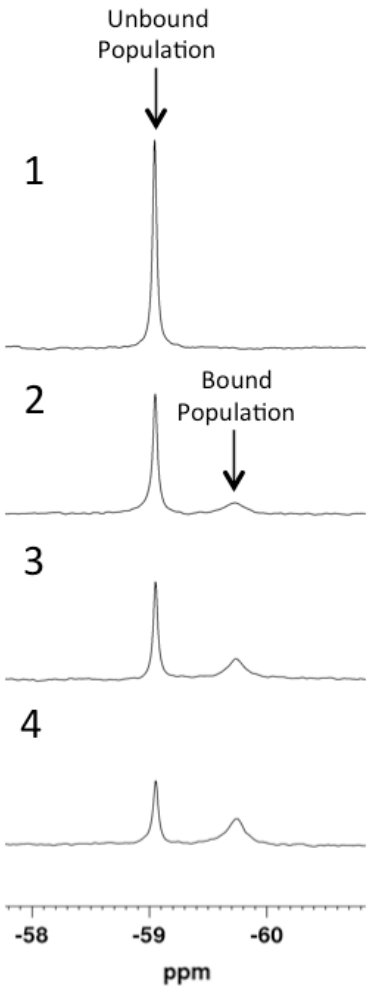


Figure 4.16 A ^{19}F NMR experiment monitoring the CF_3 signal of **98**, highlighting the unbound population of **98** (sharp peak; −59.07 ppm) and the appearance of a new ^{19}F signal, corresponding to the bound population of **98** complexed with *Sd*KefCTD (broad peak; −59.74 ppm). (1) **98** (50 μM); (2) **98** (50 μM) + *Sd*KefCTD (17 μM); (3) **98** (50 μM) + *Sd*KefCTD (25 μM); (4) **98** (50 μM) + *Sd*KefCTD (30 μM);

Prof. Tim Claridge and Amjad Khan attributed the new peak to a ^{19}F slow exchange system, similar to that reported by *Matei et al.*,[19] whereby the unbound population of **98** is in slow exchange with a bound **98**-*Sd*KefCTD complex, giving rise to the two corresponding peaks of −59.07 ppm and −59.74 ppm respectively. These two distinct peaks arise due to the sensitivity of ^{19}F chemical shifts to the surrounding environment. Prof. Tim Claridge and Amjad Khan proposed to exploit this slow exchange system to extract a K_D for **98** in a single shot experiment, as all of the components required to calculate a K_D could be determined due to the intensity of the broad

peak at −59.74 ppm corresponding to the concentration of the **98**-*Sd*KefCTD complex, [PL], (Equation 4.1).

$$K_D = \frac{[P][L]}{[PL]}$$

Equation 4.1 The equation for the dissociation constant K_D, where [P] is the concentration of the protein; [L] is the concentration of the ligand; [PL] is the concentration of the protein-ligand complex.

Amjad Khan obtained a K_D of 1.4 ± 0.5 μM for **98** using this single shot ^{19}F NMR approach, which is comparable to that obtained using ^{1}H CPMG NMR: K_D = 1.8 ± 0.3 μM. With a reasonable K_D value for **98** being obtained using this approach, attention was turned to determining binding constants of competing ligands, K_I values.

It was proposed that the binding constants for competing ligands, K_I, could be calculated in single shot experiments through determining the apparent binding constant, K_D^{app}, of **98** upon the introduction of a competing ligand using the equation in Equation 4.2 A, reported by Dalvit.[20] The values for K_D, K_D^{app}, and the concentration of the competing ligand introduced could thus be used in Equation 4.2 B, also reported by Dalvit[20], to calculate the K_I values.

$$\text{A: } K_D^{app} = \frac{[P_T][L_T] - [P_T][PL] + [PL]^2 - [L_T][PL]}{[PL]} \qquad \text{B: } K_I = \frac{[I]K_D}{K_D^{app} - K_D}$$

Equation 4.2 (A) The equation for the apparent binding constant, K_D^{app}, where $[P_T]$ is the total concentration of the protein; $[L_T]$ is the total concentration of the reporter ligand; [PL] is the concentration of the protein-ligand complex. (B) The equation for the binding constant of a competing ligand K_I, where [I] is the concentration of the competing ligand; K_D^{app} is the apparent binding constant of the reporter ligand; K_D is the dissociation constant of the reporter ligand.

As a proof of concept, Amjad Khan observed whether the bound and unbound populations of **98** could be altered through the introduction of the competing ligand **^{t}BuSG**. Pleasingly, it was found that the introduction of **^{t}BuSG** caused an increase in the signal intensity of the unbound population of **98** and a decrease in the signal intensity corresponding to the bound **98**-*Sd*KefCTD population (Figure 4.17).

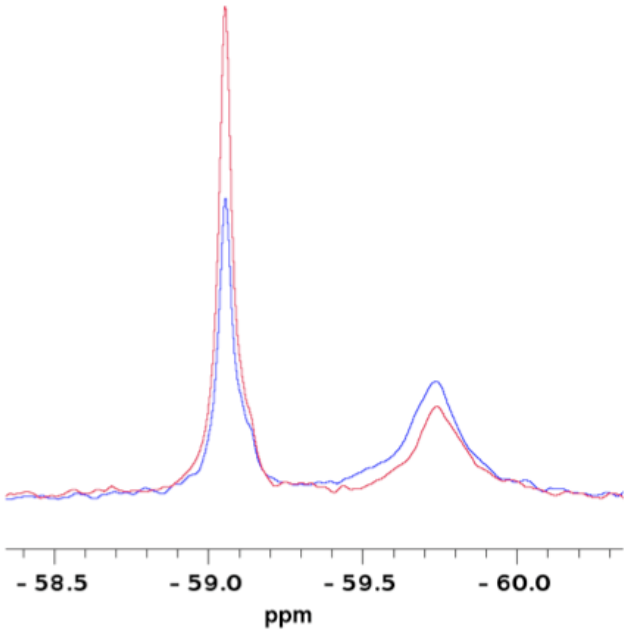
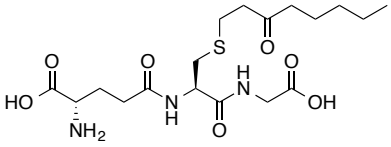


Figure 4.17 Highlights the ability to alter the populations of the bound and unbound states through the displacement of **98** from *Sd*KefCTD with the competing ligand **^{t}BuSG** (40 µM) when recording a standard ^{19}F NMR spectrum. The blue ^{19}F NMR spectrum corresponds to **98** (50 µM) in the presence of *Sd*KefCTD (35 µM) and the red ^{19}F NMR spectrum corresponds to **98** (50 µM) in the presence of *Sd*KefCTD (35 µM) and **^{t}BuSG** (40 µM).

The ^{19}F NMR assay was then validated using a set of training ligands that were reported in the paper published by Healy *et al.*[1] because a number of affinities were available for each of these compounds that had been obtained using up to three different quantitative assays: fluorescence emission spectra; fluorescence anisotropy; ITC. In addition to the literature affinities, the K_I value obtained for **^{t}BuSG** using the ^{19}F NMR assay were also compared to the K_D values obtained using

^{1}H CPMG NMR. All of the data processing and analysis of the ^{19}F NMR assay was performed by Amjad Kahn, and the results obtained by him are highlighted in Table 4.1.

Table 4.1 Comparison of the binding constants, K_I, for a range of competing ligands that bind to *Sd*KefCTD determined using the ^{19}F NMR assay to K_D values reported in the literature[1] and obtained using ^{1}H CMPG NMR. The literature values were obtained using one or more of the following techniques: fluorescence emission spectra; fluorescence anisotropy; ITC. The structure of S-octan-3-on-1-yl glutathione, **100**, is shown below.



100

Ligand	K_D / μM K_I / μM*	Method
98	1.4 ± 0.5	^{19}F NMR assay
	1.8 ± 0.3	^{1}H CPMG NMR
^{t}BuSG (24)	2.0 ± 0.2*	^{19}F NMR assay
	0.54 ± 0.3	^{1}H CPMG NMR
	0.4 ± 0.2	emission peak
	6.7 ± 0.27	ITC
ESG (18)	21 ± 1.0*	^{19}F NMR assay
	12 ± 3	emission peak
	23 ± 4	ITC
DNGSH (35)	19 ± 1.5*	^{19}F NMR assay
	6 ± 2	emission peak
	8 ± 2	anisotropy
	19 ± 6	ITC
S-Octan-3-on-1-yl glutathione (100)	23 ± 1.5*	^{19}F NMR assay
	4.4 ± 0.5	emission peak
GSH (11)	881 ± 159*	^{19}F NMR assay
	900 ± 200	emission peak
S-lactoyl glutathione (SLG; 16)	1500 ± 192*	^{19}F NMR assay
	900 ± 200	emission peak

As can be observed from Table 4.1, the binding constants of the competing ligands, K_I, obtained using the ^{19}F NMR assay developed by Amjad Khan are in good agreement with the reported literature values.

4.8 ^{1}H CPMG NMR and ^{19}F NMR assay conclusions

In conclusion, ^{1}H CPMG NMR was found to be a reliable, fluorescence-free approach to obtaining K_D values of ligands, however, it was heavily reliant on the ligand having a suitable high intensity signal that could be tracked in the ^{1}H NMR spectrum that did not overlap with signal from the protein. Furthermore, ^{1}H CPMG NMR did not show promise for performing high throughput screening, as it would involve interpretation of complex ^{1}H NMR spectra. A ^{19}F NMR assay was thus developed with a fluorinated glutathione S-conjugate acting as a reporter ligand to circumvent the aforementioned problems with ^{1}H CPMG NMR. This ^{19}F NMR assay was validated using a set of training ligands and shows promise for identifying novel ligands of Kef in a high throughput-screening program.

4.9 Microscale thermophoresis and fluorescence polarisation

4.9.1 Introduction to microscale thermophoresis and fluorescence polarisation

4.9.1.1 Microscale thermophoresis introduction

Microscale thermophoresis is a technique that can be used to detect the directed movement of a fluorescently tagged ligand along a temperature gradient. This directed movement is termed thermophoresis, a phenomenon first observed by Ludwig[21] that is dependent upon the size, charge and solvation shell of the ligand. Binding of the fluorescently tagged ligand to a protein

changes one or more of the aforementioned properties of the ligand, such that measuring the change in thermophoresis upon titration of the protein allows a quantitative insight into the ligand's affinity for the protein. After increasing the temperature of a sample locally by ≈2-6 K with an IR-laser, a steady state is attained due to the effects of thermophoresis and mass diffusion on the fluorescent ligand balancing. The stationary spatial concentration distribution of the fluorescent ligand in this steady state can be quantified using the Soret coefficient, S_T, (Equation 4.3).

$$\frac{c_{hot}}{c_{cold}} = e^{-S_T \Delta T}$$

Equation 4.3 The relationship between the stationary spatial concentration and the Soret coefficient, where c_{hot}/c_{cold} is the stationary spatial concentration ratio; S_T is the Soret coefficient; ΔT is the temperature rise.[22]

Measuring the fluorescence distribution of the fluorescent ligand before (F_{cold}) and after (F_{hot}) IR-laser heating allows the normalised fluorescence F_{norm} to be calculated, which is related to the Soret coefficient and thus the stationary spatial concentration ratio by the equation in Equation 4.4.

$$F_{norm} = \frac{F_{hot}}{F_{cold}} = 1 - [S_T - \frac{dF}{dT}]\Delta T$$

Equation 4.4 The relationship between the normalised fluorescence, F_{norm}, and the Soret coefficient, S_T, where F_{cold} is the fluorescence distribution of the fluorescent ligand before IR-laser heating; F_{hot} is the fluorescence distribution of the fluorescent ligand after IR-laser heating; S_T is the Soret coefficient; ΔT is the temperature rise.

A dissociation constant for the fluorescent ligand can therefore be determined by deconvoluting the contributions to F_{norm} from the bound (F_{norm}(bound)) and unbound (F_{norm}(unbound)) states of the fluorescent ligand through titration of the protein until all of the binding sites are saturated using the equation in Equation 4.5.

$$F_{norm} = (1 - x)F_{norm}(unbound) + xF_{norm}(bound)$$

Equation 4.5 Where, F_{norm}(bound) is the contribution to fluorescence from the complex formed by the fluorescent ligand with the protein; F_{norm}(unbound) is the contribution from the unbound fluorescent ligand; x is the fraction of fluorescent ligands that are bound in the complex.[23]

4.9.1.2 Fluorescence polarisation introduction

The theory underlying fluorescence polarisation was first outlined by Perrin in 1926.[24] Fluorescence polarisation is founded on the principle that when a fluorescent ligand is bound to a large macromolecule, it has a larger effective volume (V) and so tumbles more slowly than when it is free in solution (when it has a small V), as highlighted by Stoke's equation (Figure 4.6).

$$\rho = \frac{3\eta V}{RT}$$

Equation 4.6 The Stoke's equation, where ρ is the rotational relaxation time (the time taken to rotate through 68.5°), V is the molecular volume, η is the viscosity, R is the ideal gas constant and T is the temperature.[25]

This phenomenon can, in theory, be used to quantitatively assess the binding of small fluorescent ligands to large proteins as well as the binding of ligands that compete with the fluorescent probe. In short, the fluorescence polarisation assay is performed by exciting a fluorescent ligand with plane-polarised light in a single plane and then measuring the intensity of the light emitted in the planes that are parallel ($I_{||}$) and perpendicular ($I_{\perp}$) to the excitation plane. $I_{||}$ and $I_{\perp}$ are related to the polarisation (P) by the equation outlined in Equation 4.7.

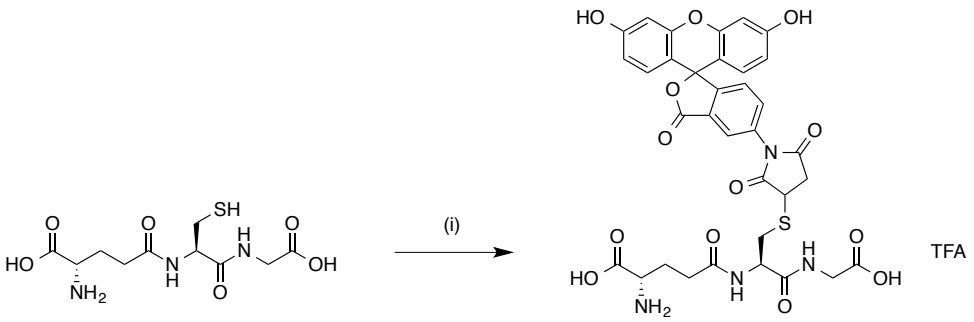
$$P = (I_{//} - I_{\perp})/(I_{//} + I_{\perp})$$

Equation 4.7 The equation relating polarisation (P) to the fluorescence intensity perpendicular to the excitation plane ($I_{\perp}$) and the fluorescence intensity parallel to the excitation plane ($I_{||}$).[25]

P is thus large when the fluorescent probe is bound to the protein ($I_{||} > I_{\perp}$) due to the larger protein-ligand complex retaining the same orientation during the period of absorbing and emitting a photon, whereas, P tends to zero when the fluorescent probe has been completely displaced from the protein and is tumbling rapidly in solution ($I_{||} = I_{\perp}$).

4.9.2 Development and synthesis of a fluorescein-tagged glutathione S-conjugate

After failed attempts to apply the fluorescent probe **DNGSH, 35,** to microscale thermophoresis and fluorescence polarisation (data not shown), it was decided to tag glutathione with fluorescein, a fluorophore typically used in these studies. Commercially available *N*-(5-fluoresceinyl)malemide was used to tag glutathione with fluorescein *via* a conjugate addition to afford **95** in a moderate yield of 56% (Scheme 4.3).



11 **95**

Scheme 4.3 Synthesis of *S*-*N*-(5-fluoresceinyl)succinimido glutathione 2,2,2-trifluoroacetate; *Reagents and conditions*: (i) *N*-(5-Fluoresceinyl)maleimide, DMF RT, 144 h, 56%.

4.9.3 Validation of the fluorescein glutathione S-conjugate binding using ^{1}H CPMG NMR and microscale thermophoresis

Once synthesised, the ability of the fluorescein glutathione S-conjugate **95** to bind to *Sd*KefCTD was assessed. The competition fluorescence assay was ruled out for performing such analysis due to the problems encountered in Chapter 3 regarding the interference with the fluorescent probe **DNGSH** when testing non-peptidic ligands. Two alternative biophysical assays were instead employed to obtain quantitative K_Ds for **95**, namely ^{1}H CPMG NMR and microscale thermophoresis. ^{1}H CPMG NMR was the first approach implemented and found that **95** bound to *Sd*KefCTD with a K_D of 1.4 ± 0.01 µM, assuming single site binding (Figure 4.18).

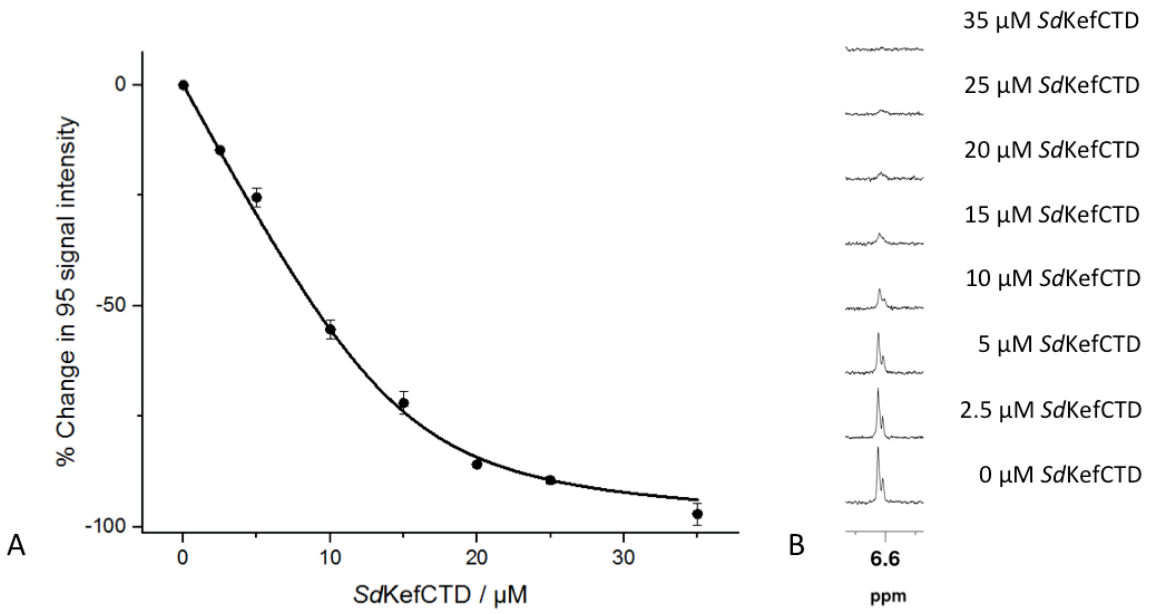


Figure 4.18 Analysis of the binding of **95** by ^{1}H CPMG NMR. (A) Shows the saturation-binding curve of **95** by ^{1}H CPMG NMR. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1; (B) Highlights the attenuation of the signal intensity of a well-defined aromatic peak of the fluorescein component of **95** upon the titration of *Sd*KefCTD.

Microscale thermophoresis was then used to corroborate the K_D obtained by ^{1}H CPMG NMR. The concentration of **95** for use in the microscale thermophoresis experiments was determined with the aid of the User Starting Guide for the Monolith NT.115 provided by NanoTemper

Technologies. The start up guide recommended performing a dye calibration curve with a "Capillary Scan" with 50% LED power and to select a concentration of fluorescent probe that results in a fluorescent count that lies between 200-1500 counts. A concentration of 25 nM of **95** gave a fluorescent count between 700-800 counts and was thus deemed appropriate to proceed to K_D measurements. In order to determine the K_D, 25 nM of **95** was subjected to a serial dilution, from 300.5-0.02 µM, of *Sd*KefCTD, with an MST Power of 80% and LED Power of 20%, to obtain a saturation-binding curve (Figure 4.19) A K_D of 9.1 ± 0.3 µM was extracted from this curve, which agreed reasonably well with the K_D of 1.4 ± 0.01 µM obtained using ^{1}H CPMG NMR. Although microscale thermophoresis showed potential for use as a quantitative assay in itself, it was decided to progress **95** to fluorescence polarisation studies and focus on the development of the fluorescence polarisation assay instead.

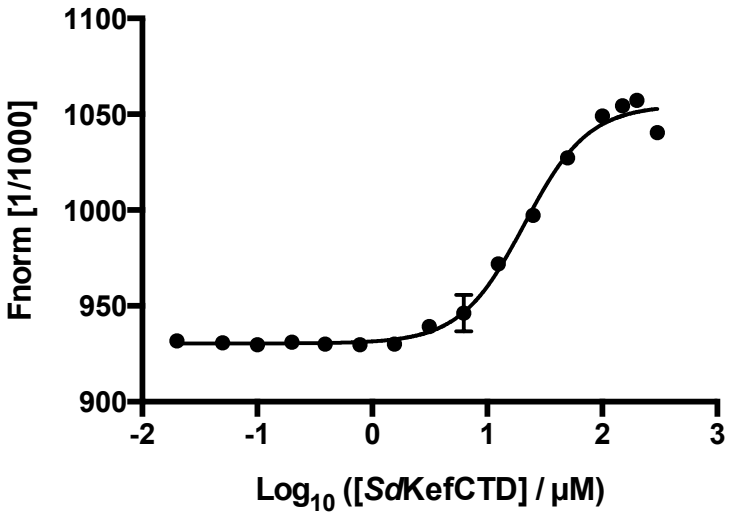


Figure 4.19 Saturation binding curve of **95** (25 nM) in the presence of *Sd*KefCTD (300.5-0.02 µM) measured using microscale thermophoresis. Error bars indicate one standard deviation of uncertainty (n = 3).

4.9.4 Developing the fluorescence polarisation assay

Confident that the fluorescein-tagged glutathione S-conjugate, **95**, was binding with good affinity to *Sd*KefCTD, as validated by ^{1}H CMPG NMR and microscale thermophoresis, it was decided to proceed with the fluorescence polarisation studies, which were all aided in their design by the *Nature Protocols* paper published by Rossi *et al.*[25] The studies commenced with an attempt at obtaining a saturation-binding curve, whereby 25 nM of **95** (the same concentration employed in the microscale thermophoresis studies) was subjected to a serial dilution, from 260-0.25 µM of *Sd*KefCTD (Figure 4.20). It was not possible to reach a plateau at the higher concentrations of protein, even when halving the concentration of **95** to 12 nM and increasing the starting concentration of the dilution to 310 µM. The curve obtained from making these changes was analogous to that highlighted in Figure 4.20. Consultation with Dr. David Staunton revealed that this observation is not uncommon in saturation-binding studies performed using fluorescence polarisation, due to the effect that increasing the protein concentration has on the viscosity and thus on the tumbling rate of the fluorescent probe, as highlighted by the contribution of viscosity (η) to the Stoke's equation (Equation 4.6).[25]

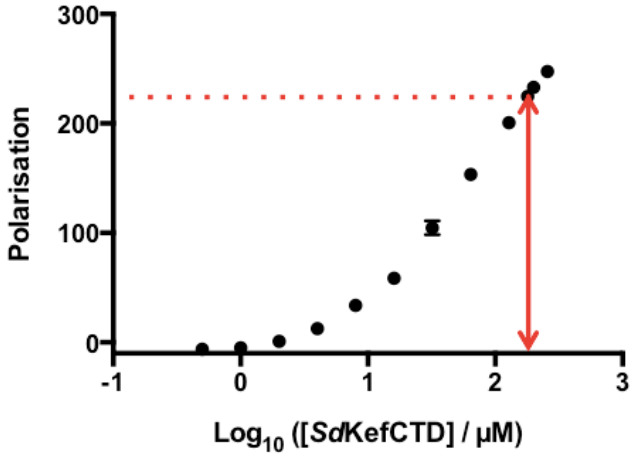


Figure 4.20 Attempted saturation-binding curve of 25 nM of **95** by fluorescence polarisation (the polarisation measured for 25 nM of **95** in the absence of protein has been subtracted from each titration point). The red arrow and dashed red line highlight the arbitrary window selected for displacement studies. Error bars indicate one standard deviation of uncertainty (n = 3).

Nevertheless, the graph obtained from the saturation-binding studies in Figure 4.20. allowed a window to be determined for the displacement of **95** from *Sd*KefCTD, for use in competition-binding studies.[25] It was arbitrarily decided that using 25 nM of **95** and a concentration of 180 μM of protein provided a suitable window of ≈200 polarisation units for such studies, as highlighted by the red arrow and dashed line in Figure 4.20. A displacement study was initially performed with single shots of competing ligands that have a range of known dissociation constants, as determined by ^{1}H CPMG NMR in Chapter 5 (Figure 4.21).

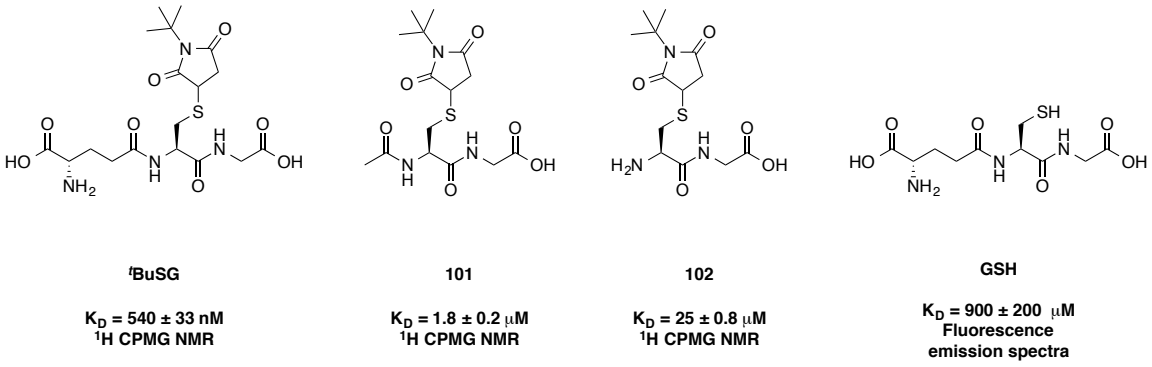


Figure 4.21 Structures of competing ligands with known affinities and their corresponding dissociation constants, as determined by ^{1}H CPMG NMR and fluorescence emission spectra.[1]

It was found that the fluorescence polarisation assay provided a qualitative reading that corresponded to the affinity of the competing ligands, with the more potent **^{t}BuSG** and **101** showing almost complete displacement of **95** from the protein and the weaker glutathione showing no statistically significant difference when compared to the polarisation reading for the negative control of buffer alone. It was also positive to see that there was still some level of sensitivity for ligands of medium potency, such as **102**. Furthermore, when performing the negative control of running the experiment in the absence of protein, the ligands were not shown to be affecting the polarisation of **95** (Figure 4.22 B), and that the effects on polarisation observed in Figure 4.22 A were protein dependent.

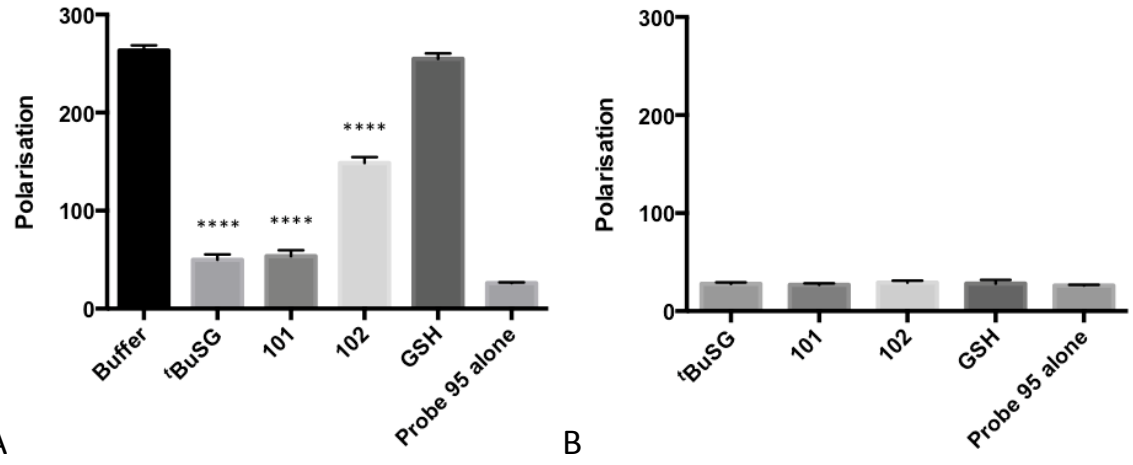


Figure 4.22 (A) The polarisation of the fluorescein probe, **95**, (25 nM) in the presence of *Sd*KefCTD (180 μM) and competing ligands (1 mM) or the negative control of buffer. The negative control buffer refers to the polarisation **95** (25 nM) in the presence of *Sd*KefCTD (180 μM) and 10 μL of assay buffer. (A) The polarisation of the fluorescein probe, **95**, (25 nM) in the absence of *Sd*KefCTD and presence of competing ligands (1 mM). Error bars indicate one standard deviation of uncertainty (n = 3 for A, n = 2 for B). Significance of changes relative to the buffer were evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$).

To assess the robustness of the fluorescence polarisation assay the Z'-factor was calculated using the control data generated in the displacement study in Figure 4.22. The Z'-factor is a statistical measure of the size of the response of a given assay and is defined in Equation 4.8.

$$Z' = 1 - \frac{(3\sigma_{c+} - 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|}$$

Equation 4.8 The equation for calculating the Z'-factor, where σ_{c+} is the standard deviation of the positive control; σ_{c-} is the standard deviation of the negative control; μ_{c+} is the mean of the positive control; μ_{c-} is the mean of the negative control.[26]

By using the standard deviation and mean for **^{t}BuSG** as the positive control and for the buffer as the negative control, a Z'-factor value of 0.85 was calculated. *Zhang et al.*[26] describe an assay with $1 > Z' \geq 0.5$ as "an excellent assay" and so confidence was taken with progressing the assay.

The fluorescence polarisation assay was next implemented in Chapter 3, Section 3.15, to analyse whether the hits identified in this chapter were false positives. The majority of these hits were confirmed to be false positives (Figure 4.23 A) and the ligands were not found to interfere with the fluorescence polarisation assay in the same way that they affected the competition fluorescence assay (Figure 4.23 B). **94** showed a statistically significant decrease in polarisation, suggesting that it is a weak binder of *Sd*KefCTD. However, the strength of the binding by fluorescence polarisation was not as extensive as was found by the competition fluorescence assay, showing that there was still a degree of interference from **94** in the competition fluorescence assay.

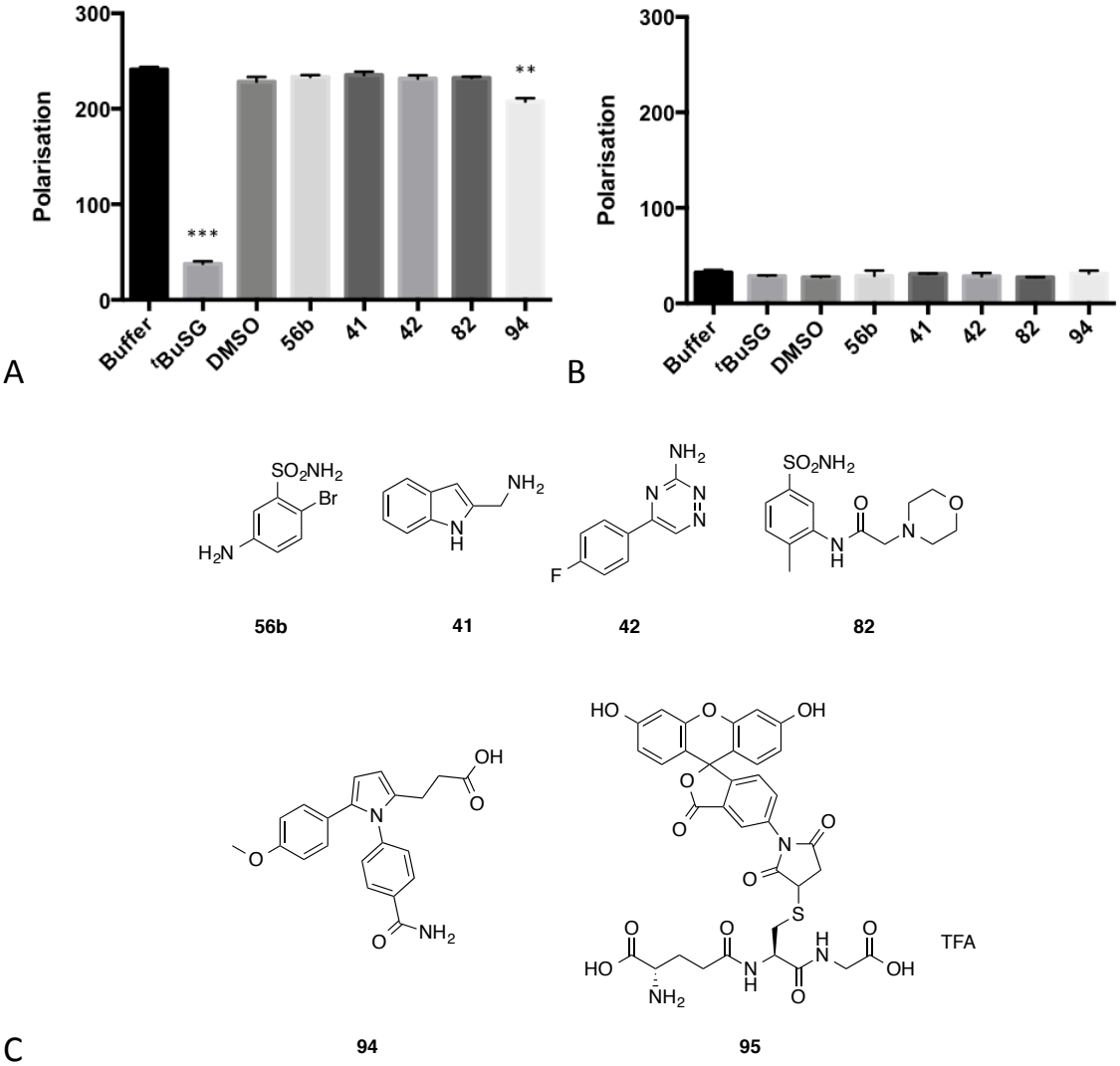


Figure 4.23 Retesting suspected false positives using a fluorescence polarisation assay described in Chapter 4. (A) The polarisation of the fluorescein probe, **95**, (25 nM) in the presence of *Sd*KefCTD (180 µM) and competing ligands (1 mM) or the negative control of buffer. (B) The polarisation of fluorescein probe, **95**, (25 nM) in the absence of *Sd*KefCTD and presence of competing ligands (1 mM). Error bars indicate one standard deviation of uncertainty (n = 3 for A, n = 2 for B). Significance of changes relative to the buffer were evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (C) Structures of compounds retested and the fluorescent probe **95**.

Having shown that the fluorescence polarisation assay could be applied in a qualitative manner, it was decided to see if it could be used quantitatively in a competition binding analysis. The concentrations of *Sd*KefCTD and the fluorescent ligand **95** were kept at 180 µM and 25 nM respectively and **^{t}BuSG** was used as the competing ligand in a serial dilution, from 12800-0.39 µM

(Figure 4.24). The resulting sigmoidal curve from plotting the polarisation against the logarithmic of the concentration of **^{t}BuSG** was fitted using GraphPad Prism 6 and the half maximal inhibitory concentration (IC_{50}) of **^{t}BuSG** was determined to be 60 ± 4 μM, using the non-linear regression, one site – Fit logIC50 curve fitting equation, which assumes that only one site of *Sd*KefCTD is being occupied. This result is promising as the IC_{50} for an unlabelled ligand that is competitively displacing a labelled ligand is expected to be higher than the unlabelled ligand's K_D.[27,28]

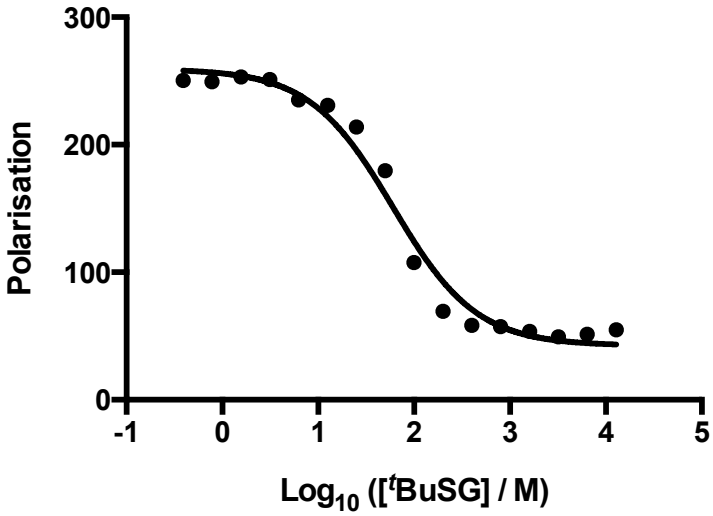


Figure 4.24 Fluorescence polarisation competition binding analysis, using **95** as the fluorescent ligand (25 nM) and with **^{t}BuSG** as the competing ligand in the presence of *Sd*KefCTD (180 μM). Error bars are plotted to indicate one standard deviation of uncertainty but are too small to be observed (n = 2). Curve fitted using Prism 6.

4.9.5 Fluorescence polarisation conclusions

A fluorescein-tagged glutathione S-conjugate, **95**, was successfully synthesised and applied to the quantitative assays microscale thermophoresis and fluorescence polarisation. Although microscale thermophoresis was not explored extensively, future work could look to perform displacement and competition-binding studies using the technique. The fluorescence polarisation assay now stands in good stead to be used in a high throughput screen format, using the

displacement approach to identify novel ligands of *Sd*KefCTD. The affinities of any novel ligands identified could be compared quantitatively by performing competition-binding studies to obtain IC_{50}s.

4.10 Overall conclusions

The application of a number of different assays to obtain quantitative affinities of ligands binding to *Sd*KefCTD was explored, four of which showed promise for future application to the screening of novel ligands. Two of these were fluorescence based, microscale thermophoresis and fluorescence polarisation, and two were NMR based, ^{1}H CPMG NMR and ^{19}F NMR. Employing each of these assays in tandem should allow the identification of novel ligands of *Sd*KefCTD and hopefully distinguish genuine hits from false positives more reliably. The project is now in a good position to perform high throughput screening with either the displacement version of the fluorescence polarisation assay or the ^{19}F NMR assay. Any promising hits can be verified quantitatively using ^{1}H CPMG NMR to obtain a K_D, fluorescence polarisation competition-binding studies to obtain an IC_{50} or the ^{19}F NMR assay to obtain a K_I. It should also be possible, in principle, to extract quantitative data for competing ligands using microscale thermophoresis.

4.11 Chapter 4 References:

[1] J. Healy, S. Ekkerman, C. Pliotas, M. Richard, W. Bartlett, S. C. Grayer, G. M. Morris, S. Miller, I. R. Booth, S. J. Conway, et al., *Biochemistry* **2014**, *53*, 1982–1992.

[2] J. Homola, S. S. Yee, G. Gauglitz, *Sensors and Actuators B: Chemical* **1999**, *54*, 3–15.

[3] B. Liedberg, C. Nylander, I. Lundström, *Biosensors and Bioelectronics* **1995**, *10*, i–ix.

[4] L. Nieba, S. E. Nieba-Axmann, A. Persson, M. Hämäläinen, F. Edebratt, A. Hansson, J. Lidholm, K. Magnusson, Å. F. Karlsson, A. Plückthun, *Analytical Biochemistry* **1997**, *252*, 217–228.

[5] D. G. Myszka, *Survey of the 1998 Optical Biosensor Literature.*, John Wiley & Sons, Ltd., **1999**.

[6] T. P. Roosild, S. Castronovo, J. Healy, S. Miller, C. Pliotas, T. Rasmussen, W. Bartlett, S. J. Conway, I. R. Booth, *PNAS* **2010**, *107*, 19784–19789.

[7] S. Miller, R. M. Douglas, P. Carter, I. R. Booth, *J. Biol. Chem.* **1997**, *272*, 24942–24947.

[8] T. P. Roosild, S. Castronovo, S. Miller, C. Li, T. Rasmussen, W. Bartlett, B. Gunasekera, S. Choe, I. R. Booth, *Structure* **2009**, *17*, 893–903.

[9] The PyMOL Molecular Graphics System, Version 1.7.4 Schrodinger, LLC, **2010**.

[10] F. Ganske, M. Orban, B. *BMG LABTECH* **2007**, Application Note 153.

[11] P. J. Hajduk, E. T. Olejniczak, S. W. Fesik, *J. Am. Chem. Soc.* **1997**, 119 12257–12261.

[12] I. K. H. Leung, M. Demetriades, A. P. Hardy, C. Lejeune, T. J. Smart, A. Szöllössi, A. Kawamura, C. J. Schofield, T. D. W. Claridge, *J. Med. Chem.* **2013**, *56*, 547–555.

[13] T. Claridge, *High-Resolution NMR Techniques in Organic Chemistry*, **2008**.

[14] H. Y. Carr, E. M. Purcell, *Physical Review* **1954**.

[15] S. Meiboom, D. Gill, *Review of scientific instruments* **2004**.

[16] T. Tengel, T. Fex, H. Emtenas, F. Almqvist, I. Sethson, J. Kihlberg, *Org. Biomol. Chem.* **2004**, *2*, 725–731.

[17] C. Dalvit, P. E. Fagerness, D. T. A. Hadden, R. W. Sarver, B. J. Stockman, *J. Am. Chem. Soc.* **2003**, *125*, 7696–7703.

[18] F. H. Niesen, H. Berglund, M. Vedadi, *Nat Protoc* **2007**, DOI 10.1038/nprot.2007.321.

[19] E. Matei, S. André, A. Glinschert, A. S. Infantino, S. Oscarson, H.-J. Gabius, A. M. Gronenborn, *Chemistry* **2013**, *19*, 5364–5374.

[20] C. Dalvit, *Concepts in Magnetic Resonance Part A* **2008**, *32A*, 341–372.

[21] C. Ludwig, K. Staatsdruckerei, *Diffusion Zwischen Ungleich Erwärmten Orten Gleich Zusammengesetzter Lösung*, **1856**.

[22] C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun, S. Duhr, *Nat Commun* **2010**, *1*, 100.

[23] M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske, S. Duhr, *Assay Drug Dev Technol* **2011**, *9*, 342–353.

[24] F. Perrin, *J. phys. radium* **1926**, *7*, 390–401.

[25] A. M. Rossi, C. W. Taylor, *Nat Protoc* **2011**, *6*, 365–387.

[26] J. Zhang, T. Chung, K. Oldenburg, *J Biomol Screen* **1999**, *4*, 67–73.

[27] Y. Cheng, W. H. Prusoff, *Biochem Pharmacol* **1973**, *22*, 3099–3108.

[28] R. B. Barlow, S. M. Bond, E. Bream, L. Macfarlane, D. S. McQueen, *Br. J. Pharmacol.* **1997**, *120*, 13–18.

Chapter 5:

Dissection of the glutathione S-conjugate ^{t}BuSG, and validation of Kef as a novel target for antibacterial agents

5. Dissection of the glutathione S-conjugate ^{t}BuSG, and validation of Kef as a novel target for antibacterial agents

5.1 Introduction and aims

In Chapter 3 ligand-based *in silico* virtual screening was performed on glutathione to identify non-peptidic inhibitors of Kef. It later became apparent that glutathione is a very weak binder with a K_D of 900 ± 200 μM (fluorescence emission spectra),[1] and so this approach was unlikely to yield potent starting points. It was therefore proposed to perform future ligand-based screening efforts on the more potent glutathione S-conjugates. Before proceeding, however, it was important to understand the key binding interactions made by glutathione S-conjugates to Kef.

The X-ray crystal structure of the *Escherichia coli* KefC C-terminal domain (*Ec*KefCCTD) in complex with KefF and the glutathione S-conjugate **ESG** (PDB code: 3L9X)[2] predicts that the peptidic core of **ESG** forms polar contacts with *Ec*KefCCTD as follows: 1) The Gly-carboxylate interacts with R416, R516 and N551 *via* a structural water; 2) The Cys-Gly amide N*H* donates a hydrogen bond to Q412; 3) The Cys-Gly amide carbonyl oxygen accepts a hydrogen bond from the D499-V500 amide N*H*; 4) The Glu-Cys amide carbonyl oxygen accepts a hydrogen bond from Q412; 5) The Glu-amine interacts with D499; 5) The Glu-carboxylate forms a salt bridge with R498 (Figure 5.1).

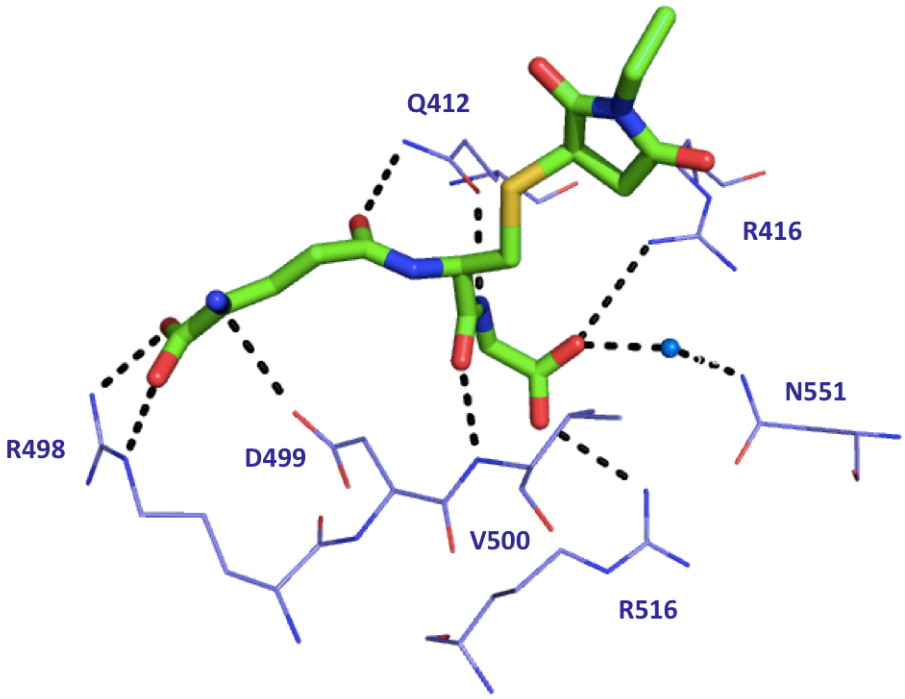


Figure 5.1 The glutathione binding site of *Ec*KefCCTD (PDB code: 3L9X; purple). Interactions between the glutathione S-conjugate **ESG** (carbon = green, one diastereomer shown) and key residues are highlighted by black dashed lines. Image generated in PyMOL.[3]

These predicted polar contacts are in agreement with the findings of Miller *et al.*[4] and Roosild *et al.*[5], who established through mutational studies that R416S, R516C (Miller *et al.*) and N551D (Roosild *et al.*) have reduced sensitivity to glutathione inhibition; suggesting that these mutations disrupt glutathione binding to *E. coli* KefC (*Ec*KefC). Subsequent functional studies performed by Roosild *et al.*[2] demonstrated the importance of the proposed interactions with Q412 and D499. The mutant Q412K showed a reduced sensitivity to inhibition by glutathione that was equivalent to introducing the triple mutation of R416A, R516A, and N551A. Mutational analysis of D499 indicated that it is unimportant for glutathione binding, with D499A and D499G not affecting glutathione inhibition. D499A and D499G were, however, found to be important for glutathione S-conjugate binding, with both mutations diminishing **ESG**'s ability to activate *Ec*KefC. The importance of five out of seven residues proposed to form polar contacts with glutathione and glutathione S-conjugates has thus been verified through mutational studies.

Given that the interaction of the D499-V500 amide N*H* would remain intact regardless of the mutation performed, only R498 was left open to mutational investigation.

With the importance of the above interactions well established from the protein's viewpoint, it was proposed to understand them from a chemical probe perspective through the dissection of the glutathione S-conjugate **^{t}BuSG** (Figure 5.2).

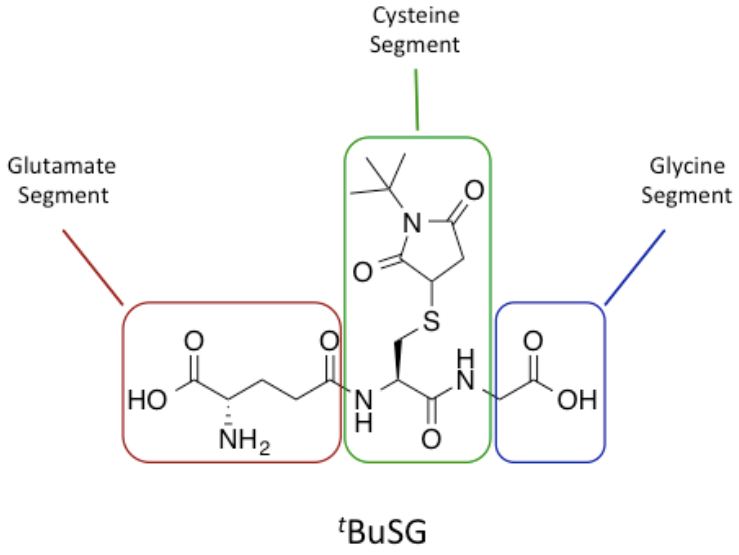


Figure 5.2 The three segments of the glutathione S-conjugate **^{t}BuSG** investigated in the molecular dissection studies.

The affinity of each fragment of **^{t}BuSG** for the C-terminal domain of Kef from *Shewanella denitrificans* (*Sd*KefCTD), was assessed qualitatively using the competition fluorescence assay[1] (described in Chapter 3) and quantitatively using ^{1}H CMPG NMR (described in Chapter 4). **^{t}BuSG** was selected as the basis for this study, as it is the most potent glutathione S-conjugate known to date and the *tert*-butyl group provided a convenient ^{1}H NMR signal to determine dissociation constants using ^{1}H CMPG NMR. Once affinities had been assigned to each fragment, it was possible to assess the contribution to binding made by each segment of the peptide through determining their group efficiency (GE), which is defined as the binding energy per heavy atom of the group. The information collated was used to develop cell membrane permeant analogues of **^{t}BuSG** for use in *in vivo* studies, namely K^{+} efflux and a Kirby-Bauer disc diffusion assay.[6] The

implementation of these tool compounds in the *in vivo* studies has provided the first validation of Kef as a novel target for antibacterial agents.

5.2 Comparison of the glutathione binding sites of *Ec*KefCCTD and a homology model of *Sd*KefCTD

As the biophysical screening of the synthesised probes were to be performed on *Sd*KefCTD, it was important to compare the binding site of *Sd*KefCTD with *Ec*KefCCTD (Figure 5.3). A homology model of *Sd*KefCTD was thus constructed using the X-ray crystal structure of *Ec*KefCCTD (PDB code: 3L9X) as a template. Molecular Operating Environment (MOE; Chemical Computing Group) software was used for the construction of the homology model. MOE-Align was used to align the query sequence of *Sd*KefCTD with the template sequence of *Ec*KefCCTD. A BLOSUM62 comparison matrix was used to give a sequence identity of 37.3%. MOE-Homology was then used to produce the homology model (Hm2*Sd*KefCTD) using an MMFF94x forcefield, with R-field solvation. **ESG** from 3L9X was preserved in the modeling of Hm2*Sd*KefCTD, with the aim of conserving the shape of the binding site as well as the binding mode of **ESG**.

Comparison of the glutathione binding site of the homology model Hm2*Sd*KefCTD with that of the X-ray crystal structure 3L9X showed that all of the key binding residues are conserved, with the exception of N551 (which is D558 in SdKef) and R498 (which is I505 in SdKef; Figure 5.3).

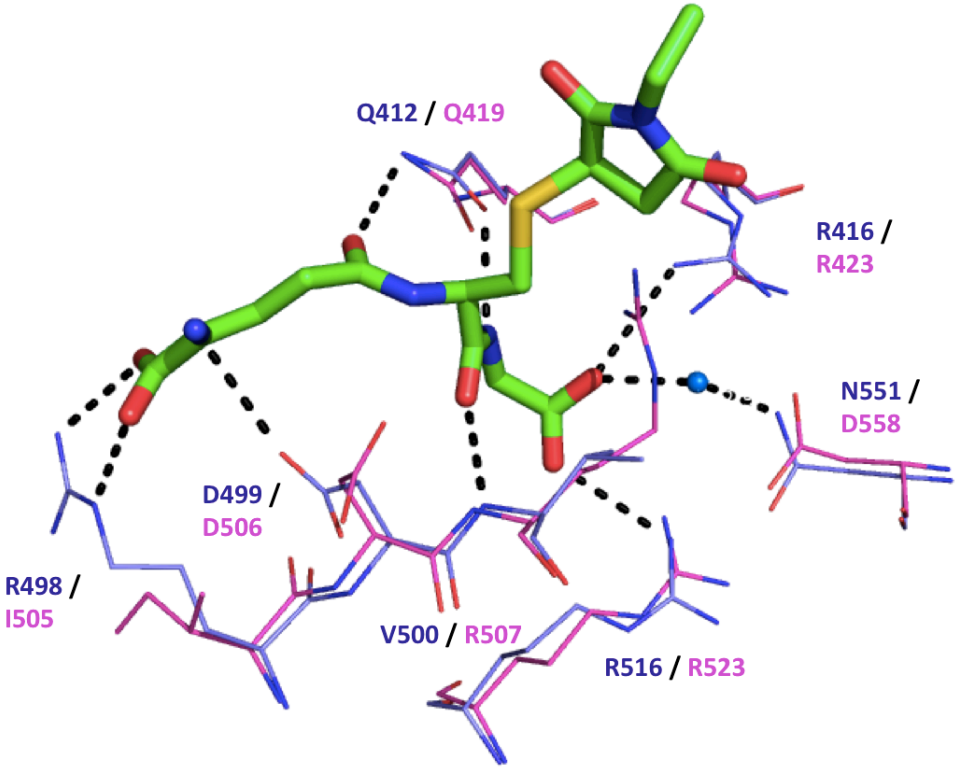


Figure 5.3 Overlay of the X-ray crystal structure of *Ec*KefCCTD (PDB code: 3L9X; carbon = purple) with the homology model of *Sd*KefCTD (Hm2*Sd*KefCTD; carbon = pink). Interactions between the glutathione S-conjugate **ESG** (one diastereomer shown) and key residues are highlighted by black dashed lines. Image generated in PyMOL.[3]

The lack of conservation of the R498 residue was of particular interest as this was one of the residues yet to be scrutinised by mutational analysis. In light of the difference, Wendy Bartlett in the Miller Group, University of Aberdeen, performed a mutational analysis on R498 to assess the importance of the proposed interaction between the Glu-carboxylate of glutathione S-conjugates and R498 in *Ec*KefC (unpublished data). As can be observed from Figure 5.4, the relative rate constant for **ESG** elicited K^{+} efflux of the mutant R498A was not statistically different to that of wild-type *Ec*KefC. The implication of this observation is that R498 is not required to bind **ESG**. This result is in keeping with R498 not being conserved in *Sd*KefCTD and brought the implied interaction between the Glu-carboxylate of glutathione S-conjugates and R498 into question. It was therefore proposed that the Glu-carboxylate of glutathione S-conjugates is redundant, and so the dissection of **^{t}BuSG** commenced with the glutamate region.

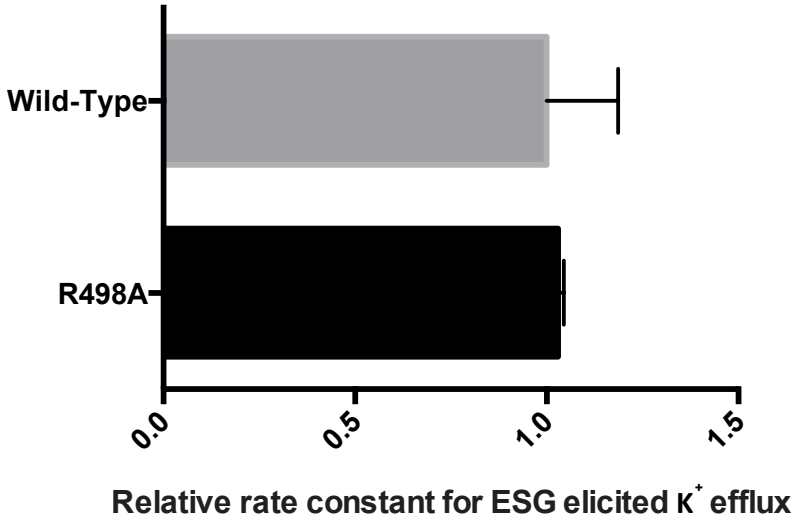


Figure 5.4 Functional analysis of R498 in *Ec*KefC. The graph shows the first-order relative rate constant of **ESG**-elicited K^+ efflux for R498A compared to wild-type *Ec*KefC. Bars show mean values normalised relative to wild-type *Ec*KefC and error bars indicate the proportional error (n = 2 for R498A and n = 4 for wild-type *Ec*KefC). Significance of changes evaluated by a Student's *t*-test.

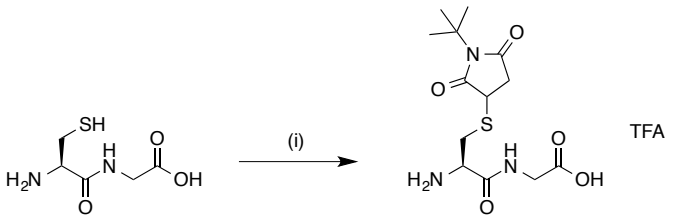
5.3 Analysis of the glutamate segment of ^{t}BuSG

To investigate the redundancy of the glutamate in glutathione S-conjugates, the glutamate region of **^{t}BuSG** was sequentially truncated to investigate the contribution from each of the functional groups. The affinity of each fragment for *Sd*KefCTD was assessed qualitatively using the competition fluorescence assay[1] (described in Chapter 3) and quantitatively using ^{1}H CMPG NMR (described in Chapter 4).

5.4 Synthesis of the probes for the glutamate segment of ^{t}BuSG

The first truncation involved complete removal of the glutamate to give compound **102** (Scheme 5.1). A conjugate addition of commercially available L-Cys-Gly, **103**, to

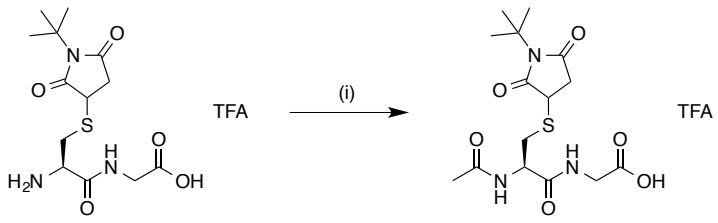
N-tert-butylmaleimide, under the same conditions used to synthesise **^{t}BuSG**, afforded **102** in a moderate yield of 58% (Scheme 5.1).



103 **102**

Scheme 5.1 Synthesis of **102**. *Reagents and conditions*: (i) *N-tert*-Butylmaleimide, NaOH, H_2O, RT, 1.5 h, 58%.

Analysis of the binding mode of **ESG** shown in Figure 5.3 indicated that removing the glutamate would prevent the following polar contacts with *Sd*KefCTD: 1) The Glu-Cys amide carbonyl oxygen accepting a hydrogen bond from Q419; 2) The Glu-amine donating a hydrogen bond to D506. The first interaction lost upon removal of the glutamate, the Glu-Cys amide carbonyl oxygen accepting a hydrogen bond from Q419, was therefore reintroduced through the acetylation of the dipeptide **102** with acetyl chloride to afford **101** (Scheme 5.2).

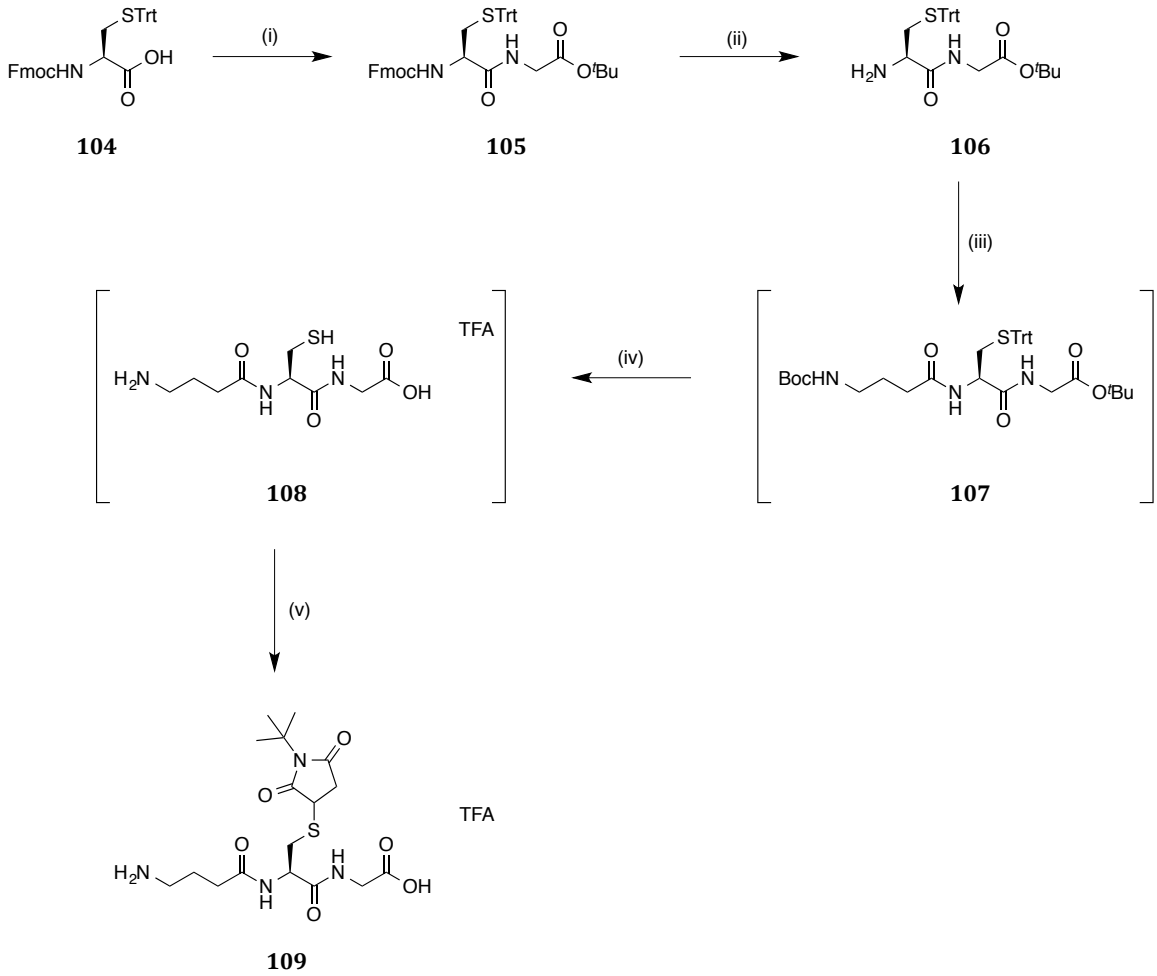


102 **101**

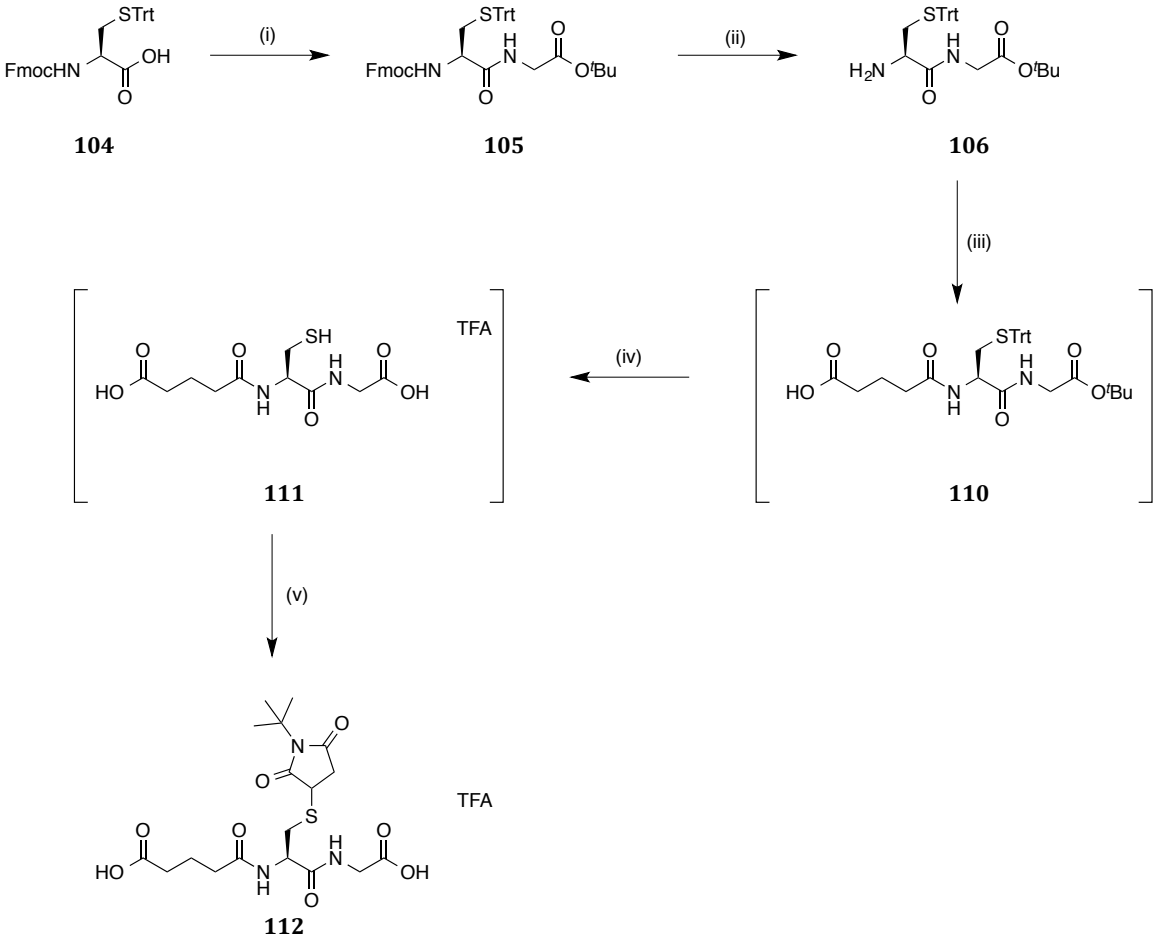
Scheme 5.2 Synthesis of **101**. *Reagents and conditions*: (i) Acetyl chloride, NEt_3, THF, RT, 6.5 h, 55%.

The second interaction lost, the Glu-amine donating a hydrogen bond to D506, was addressed through the synthesis of the dideoxy *S-N-tert*-butylsuccinimido glutathione, **109**, (**^{t}BuSG** lacking the Glu-carboxylic acid). In addition to dideoxy **^{t}BuSG**, **109**, deamino *S-N-tert*-butylsuccinimido glutathione, **112**, (**^{t}BuSG** lacking the Glu-amine) was synthesised to investigate whether the Glu-carboxylate makes any binding contribution to *Sd*KefCTD.

The syntheses of the dideoxy, **109**, (Scheme 5.3) and deamino, **112**, (Scheme 5.4) **^{t}BuSG** analogues were less trivial, with both being obtained in five steps. Initial coupling of *N*-Fmoc-L-Cys(Trt)-OH, **104**, with glycine *tert*-butyl ester hydrochloride using the coupling reagent HBTU afforded the fully protected L-Cys-Gly scaffold, **105**, which was subjected to a chemoselective deprotection of the Fmoc group. The ɣ-linked dideoxy and deamino glutamate analogues were introduced to the L-Cys-Gly scaffold by coupling the unprotected amine, **106**, to either *N*-Boc-ɣ-aminobutyric acid or glutaric anhydride. A global deprotection and subsequent conjugate addition to *N*-*tert*-butylmaleimide was then performed to afford **109** and **112** in yields of 37% and 12%, respectively, over three steps.



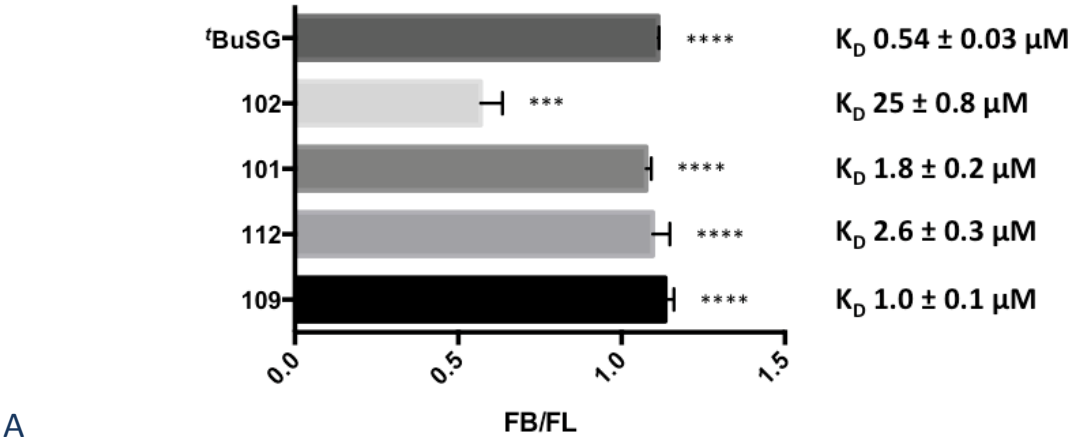
Scheme 5.3 Synthesis of **109**. *Reagents and conditions*: (i) Glycine *tert*-butyl ester hydrochloride, HBTU, DIPEA, DMF, RT, 24 h, 99%; (ii) 20% Piperidine in DMF, RT, 5 h, 74%; (iii) *N*-Boc-ɣ-aminobutyric acid, EDC·HCl, HBTU, DIPEA, CH_2Cl_2, RT, 24 h; (iv) Triethylsilane, 20% TFA in CH_2Cl_2, RT, 6.5 h; (v) *N*-*tert*-Butylmaleimide, H_2O, RT, 19.5 h, 37% (over three steps).



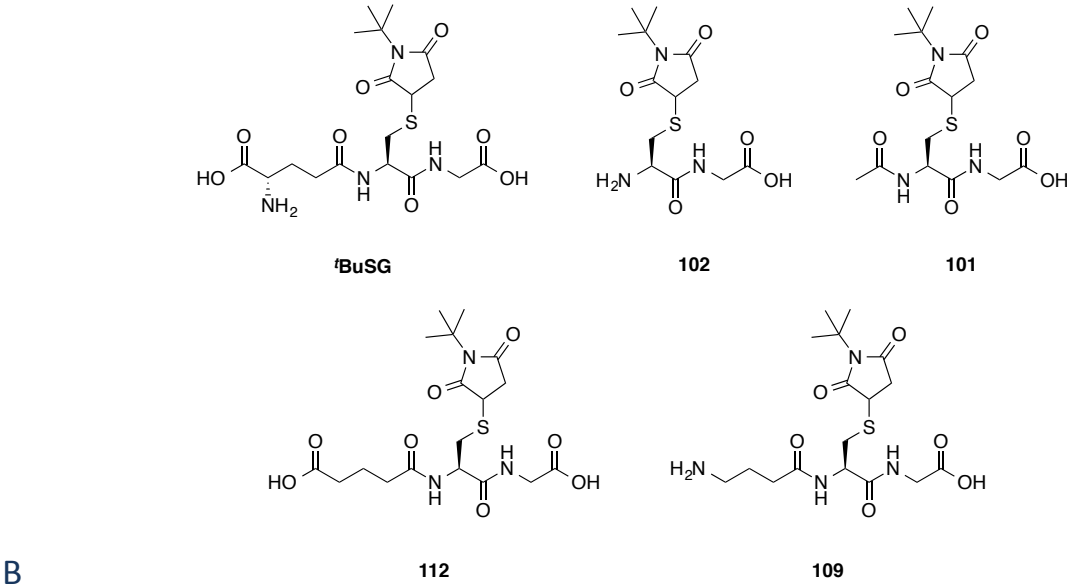
Scheme 5.4 Synthesis of **112**. *Reagents and conditions:* (i) Glycine *tert*-butyl ester hydrochloride, HBTU, DIPEA, DMF, RT, 24 h, 99%; (ii) 20% Piperidine in DMF, RT, 5 h, 74%; (iii) Glutaric anhydride, NEt_3, $CHCl_3$, RT, 15.5 h; (iv) Triethylsilane, 20% TFA in CH_2Cl_2, RT, 6.5 h; (v) *N-tert*-Butylmaleimide, H_2O, RT, 14 h, 12% (over three steps).

5.5 *In vitro* analysis of the glutamate segment of ^{t}BuSG

The resulting chemical probes from the dissection of the glutamate of **^{t}BuSG**, **101**, **102**, **109** and **112**, were subjected to a qualitative assessment of binding using the competition fluorescence assay[5] (described in Chapter 3; Section 3.2.1), and dissociation constants were determined for all of them using ^{1}H CPMG NMR (described in Chapter 4; Section 4.6). As a brief reminder of how the competition fluorescence assay results are interpreted: $(F_B/F_L - 1) = 0$ indicates no binding, and non-zero values of $(F_B/F_L - 1)$ indicate the strength of binding, with larger values implying stronger binding.



A



B

Figure 5.5 (A) Shows (F_B/F_L – 1) of the indicated compound (1 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). K_D values were obtained using ^{1}H CPMG NMR, assuming single site binding. K_D values are shown next to the (F_B/F_L – 1) bars corresponding to the compounds that they were obtained for. See the Appendix in Section 5.20 for Figures of the data obtained by ^{1}H CPMG NMR. (B) The structures of the compounds tested.

The competition fluorescence assay showed that completely removing the glutamate (compound **102**) only resulted in a partial loss in affinity when compared to full-length **^{t}BuSG** (Figure 5.5). The change in affinity caused by removing the glutamate was quantified by ^{1}H CPMG

NMR, which found that the dissociation constant increased from 0.54 ± 0.03 μM for full-length **^{t}BuSG** to 25 ± 0.8 μM for the truncated dipeptide, **102**. Given that roughly one third of the peptide backbone had been removed, this drop was not as substantial as might be expected.

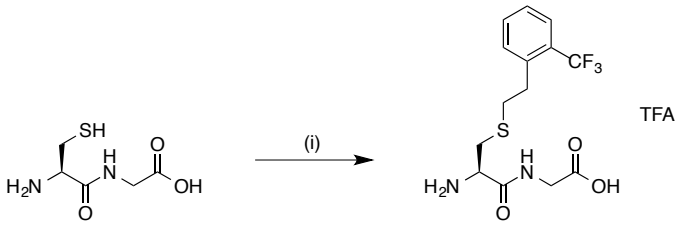
Acetylation of **102** restored the majority of the affinity, with ($F_B/F_L - 1$) of the acetylated dipeptide, **101**, becoming analogous to that of **^{t}BuSG** and the K_D of **101** returning to 1.8 ± 0.2 μM (Figure 5.5). This result fits with the prior mutational studies on Q419,[1] further demonstrating the importance of the binding contributions of Q419, which is proposed to donate a hydrogen bond to the carbonyl of the Glu-Cys amide bond of glutathione S-conjugates (Figure 5.3). Although the majority of the affinity was restored upon acetylating **102**, there was still a disparity between the K_D values of **^{t}BuSG** and the acetylated dipeptide, **101**: 0.54 ± 0.03 μM and 1.8 ± 0.2 μM respectively. As the published mutational studies on D499 had found that it was required for activation by the glutathione S-conjugate **ESG**, it was hypothesised that this difference could be due to the acetylated dipeptide, **101**, not being able to form an interaction with D506 (the equivalent residue to D499 in *Sd*KefCTD).[2] The dideoxy **^{t}BuSG** analogue, **109**, provided an insight into this hypothesis, as reintroduction of the amino group was predicted to improve the dissociation constant through picking up an interaction with D506. As expected, there was an increase in the affinity of **109** (K_D = 1.0 ± 0.1 μM) in comparison to **101** (K_D = 1.8 ± 0.2 μM). The difference in K_D between full-length **^{t}BuSG** (K_D = 0.54 ± 0.03 μM) and **109** (K_D = 1.0 ± 0.1 μM) may be accounted for by the lack of the inductively electron withdrawing Glu-carboxylate group, which could enhance the hydrogen bond donating abilities of the Glu-amine to D506 (Figure 5.5).

The redundancy of the Glu-carboxylate was next investigated through testing the deamino **^{t}BuSG** analogue, **112** (Figure 5.5). The K_D of **112** (2.6 ± 0.3 μM) was found to be higher when compared to its truncated analogue **101** (K_D = 1.8 ± 0.2 μM). This result suggests that, as hypothesised, the carboxylate does not make a direct contribution to binding. It is possible that

the K_D of the deamino **^{t}BuSG** analogue, **112**, is higher than **101** due to the cost incurred upon desolvating the deamino Glu-carboxylate from the bulk solvent not being compensated for by a binding interaction between the deamino Glu-carboxylate and *Sd*KefCTD.

5.6 Analysis of the glutamate segment of 98

To investigate whether glutathione S-conjugates, other than those formed *via* conjugate additions with maleimides, also retain a moderate affinity upon removal of the glutamate; the glutamate was removed from the fluorinated analogue, **98**. Compound **98** was chosen as its ability to bind to *Sd*KefCTD was quantified in Chapter 4, where it was developed as a fluorinated reporter ligand for use in the ^{19}F NMR assay. The desired truncated analogue of **98** was synthesised using a photochemical thiol-ene reaction[7] to couple L-Cys-Gly with 2-trifluoromethyl styrene to afford **113** in 19% yield (Scheme 5.5).



103 **113**

Scheme 5.5 Synthesis of 1**13**. *Reagents and conditions:* (i) 2-Trifluoromethyl styrene, 2,2-dimethoxyphenyl acetophenone, THF/H_2O, RT, hv, 5 h, 19%.

Compound **113** was subjected to a qualitative assessment of binding using the competition fluorescence assay, along with the full-length glutathione S-conjugates **98, ^{t}BuSG** and the truncated **^{t}BuSG** analogue **102**. The result of this analysis found that removing the glutamate from **98** did not cause a significant drop in affinity, with the value of (F_B/F_L – 1) for **113** being comparable to the truncated analogue of **^{t}BuSG**: **102** (Figure 5.6).

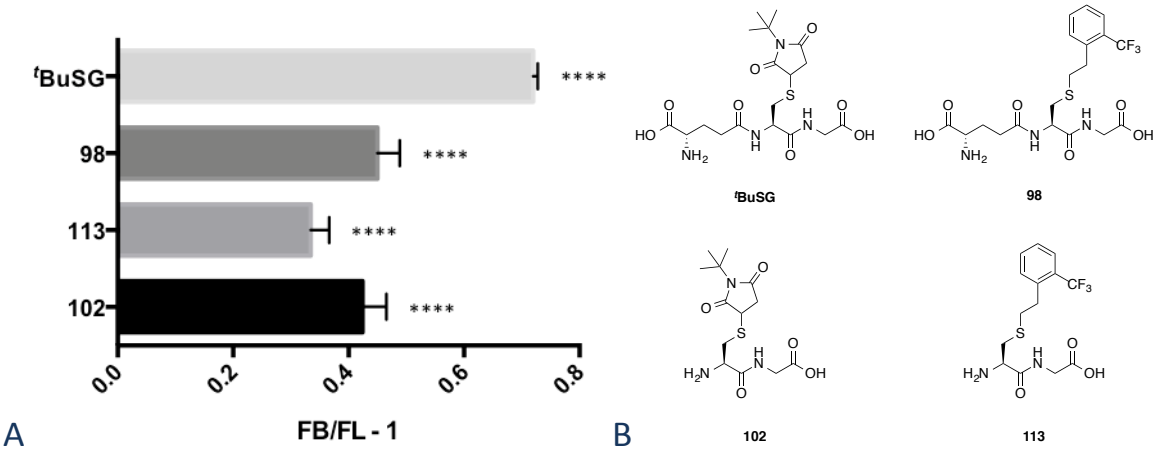


Figure 5.6 (A) Shows (F_B/F_L – 1) of the indicated compound (1 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) The structures of the compounds tested.

5.7 Analysis of the cysteine segment of ^{t}BuSG

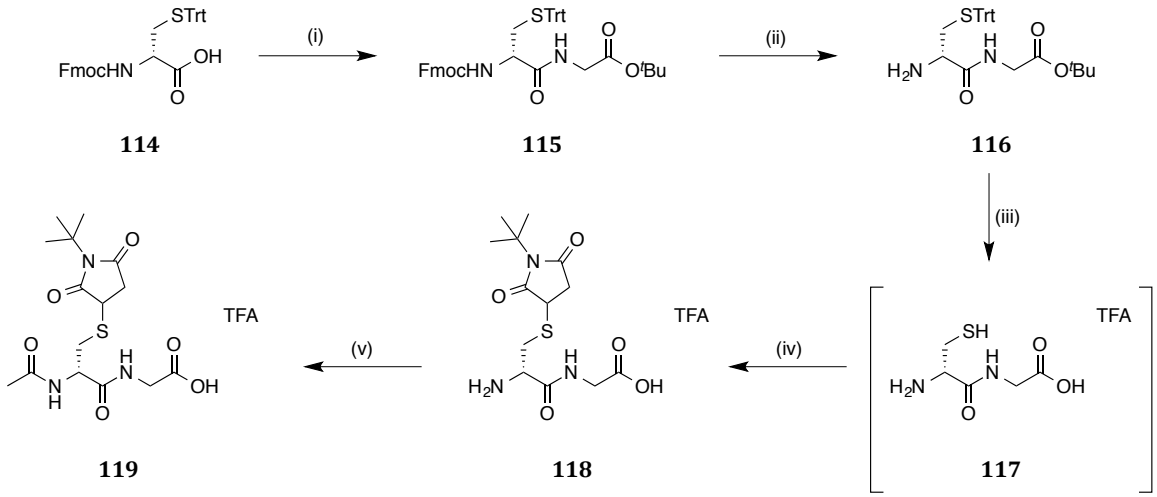
Having established the SAR of the glutamate section of the peptide backbone of glutathione S-conjugates, attention was shifted to understanding the binding contributions made by the central cysteine segment. With the knowledge that the acetylated dipeptide, **101**, has an affinity comparable to full-length **^{t}BuSG**, it was possible to explore the SAR of this section using the more synthetically tractable acetylated dipeptide, **101**, as a reference scaffold.

Three key interactions between the central cysteine segment and *Sd*KefCTD were identified for investigation. The first involved understanding the unknown role of the succinimide ring in the binding mode of glutathione S-conjugates. The succinimide clearly makes a vital contribution to the affinity of glutathione S-conjugates, as highlighted by the disparity in the reported K_D values of glutathione (K_D = 900 ± 200 µM; fluorescence emission spectra)[1] and **^{t}BuSG** (K_D = 0.40 ± 0.2 µM; fluorescence emission spectra)[1], however, it is not clear what the source of

this difference is. The second and third interactions proposed for investigation were the following polar contacts: 1) The Cys-Gly amide N*H* donating a hydrogen bond to Q419; 2) The Cys-Gly amide carbonyl oxygen accepting a hydrogen bond from the D506-R507 amide N*H*.

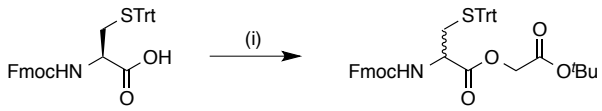
5.8 Synthesis of the probes for the cysteine segment of ^{t}BuSG

To explore the role of the succinimide ring in the binding of glutathione S-conjugates, it was decided to investigate the effect that inverting the stereochemistry of the cysteine from L- to D- would have on the affinity. This information would give an indication as to whether the vector of the succinimide ring is important for forming specific interactions with the protein. Synthesis of the D-analogue proceeded through a similar route to that used to make the dideoxy, **109**, and deamino, **112**, **^{t}BuSG** analogues (Schemes 5.3 & 5.4). The synthesis commenced with a coupling of *N*-Fmoc-D-Cys(Trt)-OH, **114**, to glycine *tert*-butyl ester hydrochloride using the coupling reagent HBTU. A global deprotection was then performed, with initial removal of the base labile Fmoc, followed by the acid labile trityl and *tert*-butyl groups. The deprotected thiol was then subjected to a conjugate addition to *N*-*tert*-butylmaleimide, followed by acetylation with acetyl chloride to afford **119** in a 46% yield (Scheme 5.6). The e.e. of the two major diasteroisomers (D-cysteinyl) over the two minor diasteroisomers (L-cysteinyl) was determined to be 97% by chiral HPLC.



Scheme 5.6 Synthesis of **119**. *Reagents and conditions:* (i) Glycine *tert*-butyl ester hydrochloride, HBTU, DIPEA, DMF, RT, 120 h, 88%; (ii) 20% Piperidine in DMF, RT, 4 h, 47%; (iii) Triethylsilane, 20% TFA in CH_2Cl_2, RT, 7 h; (iv) *N*-*tert*-Butylmaleimide, H_2O, RT, 16 h, 71% (over two steps); (v) Acetyl chloride, NEt_3, THF, RT, 4 h, 46%.

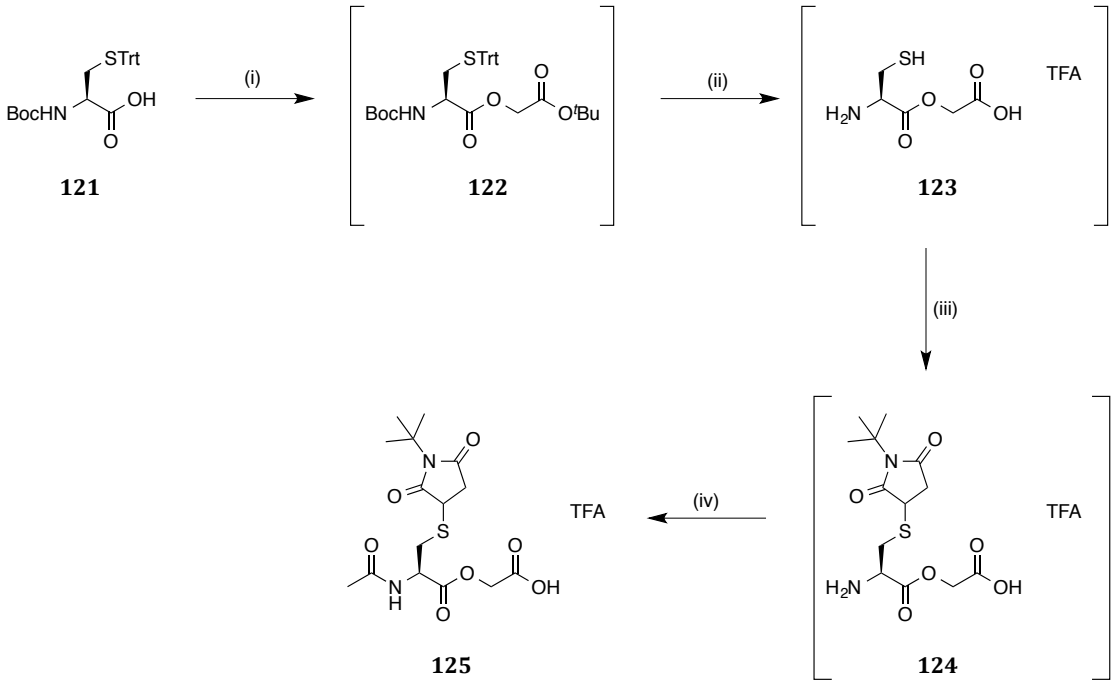
To investigate the importance of the Q419 interaction with the Cys-Gly amide N*H*, the hydrogen bond donating capability of the Cys-Gly amide on compound **101** was eliminated, by converting it to the ester, **125**. It was anticipated that **125** could be synthesised in a similar manner to the dideoxy, **109**, and deamino, **112**, **^{t}BuSG** analogues (Schemes 5.3 & 5.4) by replacing glycine *tert*-butyl ester hydrochloride with *tert*-butyl-2-hydroxyacetate. The first step was modified to involve the use of the coupling reagent PyBOP, which had had prior success for forming esters within the group, to couple *N*-Fmoc-L-Cys(Trt)-OH, **104**, to *tert*-butyl-2-hydroxyacetate in 81% yield (Scheme 5.7). However, upon characterising the resulting compound, **120**, an $[\alpha]_D^{20}$ = −0.6 (*c* 1.0, $CHCl_3$) was measured, and two peaks were observed in the chiral HPLC trace (Peak 1: Ret. Time = 23.945 min, 45.18%; Peak 2: Ret. Time = 30.247 min, 54.62%). As there was no literature data available for this compound, these results were taken as evidence that the L-cysteine had been racemised under the reaction conditions, with the two peaks in the chiral HPLC corresponding to the two enantiomers.



104 **120**

Scheme 5.7 Synthesis of **120**. *Reagents and conditions:* (i) *tert*-Butyl-2-hydroxyacetate, PyBOP, NEt_3, THF, RT, 20.5 h, 81%.

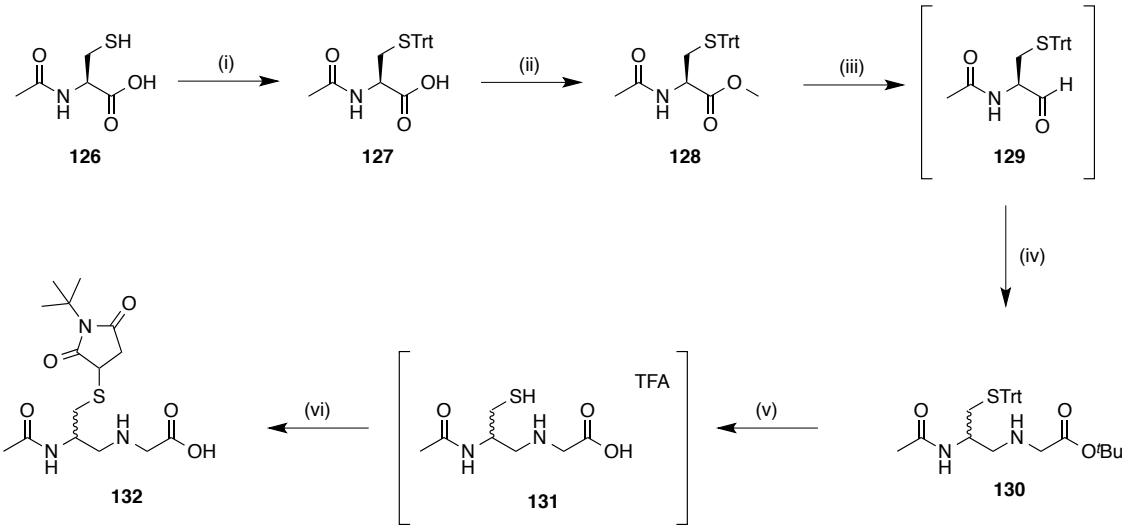
An alternative approach to the first step was therefore employed to circumvent the problem of racemisation. A literature search identified examples of using the coupling reagent *N,N'*-dicyclohexylcarbodiimide (DCC), in the absence of base, to form esters from amino acids without the issue of racemisation. The DCC coupling of *N*-Boc-L-Cys(Trt)-OH, **121**, to *tert*-butyl-2-hydroxyacetate was successful, however, it was not possible to remove all of the resulting DCC urea impurity by silica gel chromatography, so the product was taken on without further purification. Using *N*-Boc-L-Cys(Trt)-OH, **121**, instead of *N*-Fmoc-L-Cys(Trt)-OH, **104**, meant that deprotection could now be achieved in a single step, followed by conjugate addition to *N-tert*-butylmaleimide. A final acetylation step using acetyl chloride afforded **125** in a 11% yield over four steps (Scheme 5.8). The final product was chiral and non-racemic, with the $[\alpha]_D^{20}$ = −24.0 (*c* 0.75, H_2O) and two peaks observed by chiral HPLC (Peak 1: Ret. Time = 20.625 min, 50.48%; Peak 2: Ret. Time = 29.359 min, 49.32%) corresponding to the two diastereomers arising from the conjugation step to *N-tert*-butylmaleimide.



Scheme 5.8 Synthesis of **125**. *Reagents and conditions:* (i) *tert*-Butyl-2-hydroxyacetate, *N,N'*-dicyclohexylcarbodiimide, acetonitrile, 0 °C, 2 h, RT, 15.5 h; (ii) Triethylsilane, 20% TFA in CH_2Cl_2, RT, 2 h; (iii) *N-tert*-Butylmaleimide, H_2O, RT, 7 h; (iv) Acetyl chloride, NEt_3, THF, RT, 3.5 h, 11% over four steps.

To address whether the Cys-Gly amide carbonyl oxygen acts as a hydrogen bond acceptor to the D506-R507 amide N*H*, it was proposed to remove the carbonyl from the Cys-Gly amide bond by synthesising **132** using a reductive amination (Scheme 5.9). In the previous routes implemented, the acetylation of the Cys-amine was performed in the last step, however, this was not possible in this instance as acetylation might occur at two possible sites: the Cys-amine and the secondary amine arising from the removal of the Cys-Gly amide carbonyl. Given that it would be difficult to achieve a selective acetylation at the Cys-amine, it was decided to introduce the acetyl group at an earlier stage in the synthesis. The synthesis thus commenced with *N*-acetyl-L-cysteine, **126**, which was protected at the thiol position using trityl chloride, to afford **127**. The carboxylic acid of compound **127** was converted to the methyl ester **128** using thionyl chloride. The methyl ester, **128**, was then reduced to the aldehyde, **129**, using DIBAL, and coupled to glycine *tert*-butyl ester hydrochloride under the reductive amination conditions of sodium cyanoborohydride to afford **130**. Upon characterising **130**, however, it seemed likely that the compound had

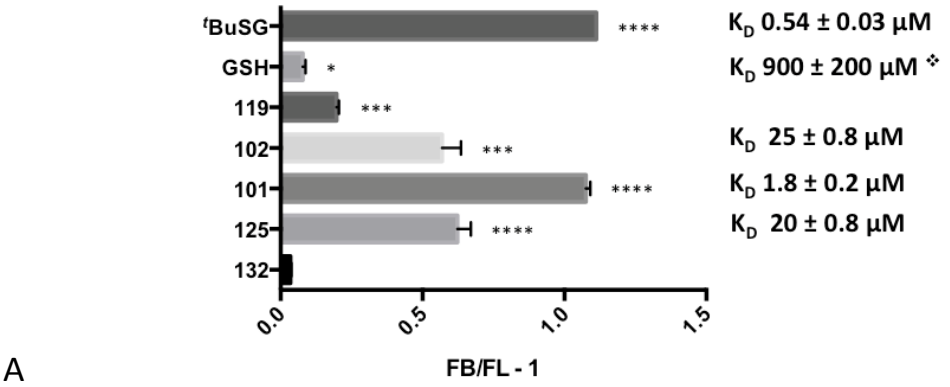
racemised, as although only one peak could be observed by chiral HPLC (Ret. Time = 16.576 min, 88.18%), the $[\alpha]_D^{20}$ = +0.7 (*c* 1.0, $CHCl_3$). Racemisation appeared to have occurred at the reductive amination step, as the methyl ester, **128**, and the slightly impure aldehyde intermediate, **129**, both had a single peak by chiral HPLC (Ret. Time = 17.780 min, 99.10%; Ret. Time = 17.659 min, 98.37%, respectively) and $[\alpha]_D^{20}$ values of +52.9 (*c* 1.0, $CHCl_3$) and +34.8 (*c* 1.0, $CHCl_3$), respectively. As there was not an obvious solution to this racemisation problem, the racemate was taken on to the final step by performing a global deprotection and a final conjugate addition to *N*-*tert*-butylmaleimide to afford **132** in 17% yield over two steps. The final product, **132**, exhibited two very broad peaks by chiral HPLC (Peak 1: Ret. Time = 19.990 min, 62.48%; Peak 2: Ret. Time = 26.366 min, 35.50%), and the $[\alpha]_D^{20}$ = +0.5 (*c* 0.59, H_2O), further suggesting that the cysteine had racemised. The product **132**, that was suspected to have racemised, was tested nonetheless, as the compounds were being dosed in the competition fluorescence assay at a final concentration of 1 mM, which is in large excess compared to the protein, *Sd*KefCTD (6 µM), and the fluorescent probe, **DNGSH** (5 µM). As the desired stereoisomers of **132** would be dosed at a final concentration of 0.5 mM, any affinity that they have for *Sd*KefCTD should still be observable.



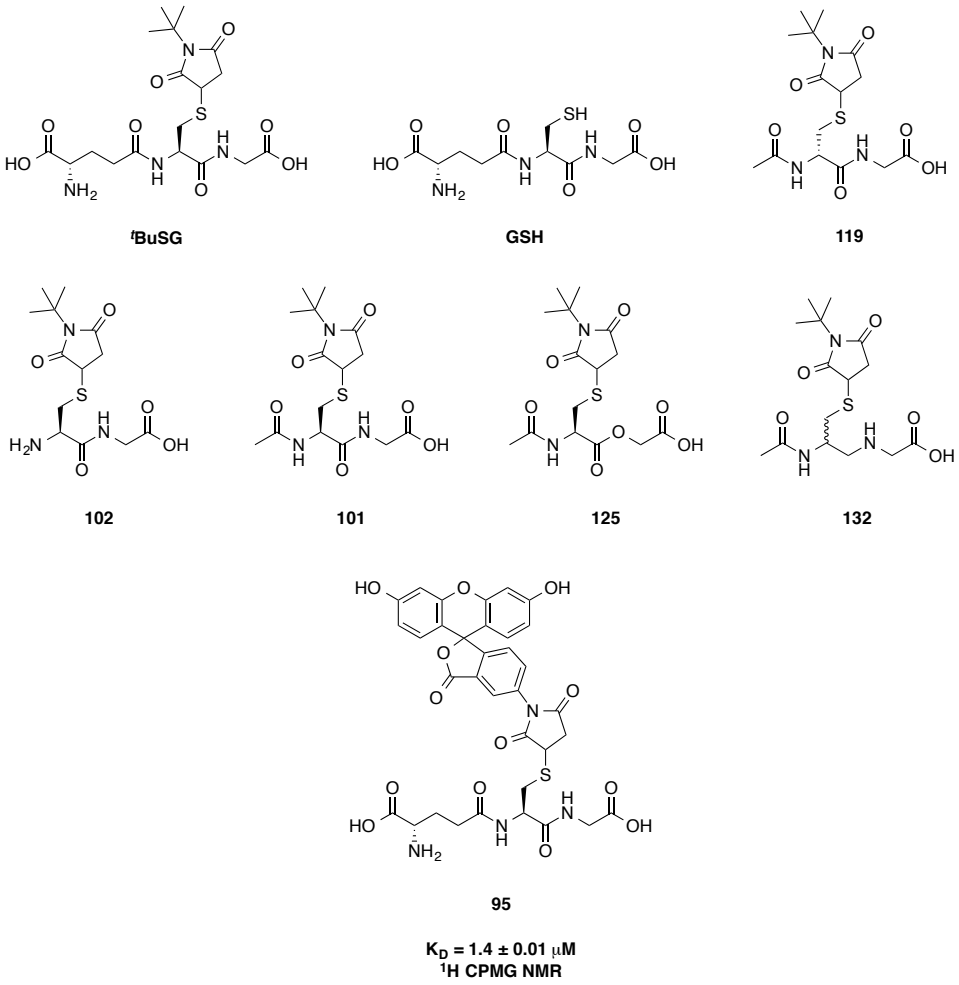
Scheme 5.9 Synthesis of **132**. *Reagents and conditions*: (i) Trityl chloride, DMF, RT, 20 h, 49%; (ii) Thionyl chloride, MeOH, RT, 3.5 h, 66%; (iii) DIBAL, toluene, −60 °C, 2 h; (vi) Glycine *tert*-butyl ester, sodium cyanoborohydride, MeOH, RT, 17 h, 8% (over two steps); (v) Triethylsilane, 20% TFA in CH_2Cl_2, RT, 2.5 h; (vi) *N*-*tert*-Butylmaleimide, H_2O, RT, 8 h, 17% (over two steps).

5.9 *In vitro* analysis of the cysteine segment of ^{t}BuSG

The resulting chemical probes from the dissection of the cysteine segment of **^{t}BuSG**, **119**, **125** and **132**, were subjected to a qualitative assessment of binding using the competition fluorescence assay, and a dissociation constant was determined for compound **125** using ^{1}H CPMG NMR (Figure 5.7).



A



B

Figure 5.7 (A) Shows (F_B/F_L – 1) of the indicated compound (1 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). K_D values were obtained using ^{1}H CPMG NMR, assuming single site binding. K_D values are shown next to the (F_B/F_L – 1) bars corresponding to the compounds that they were obtained for. See the Appendix in Section 5.20 for Figures of the data obtained by ^{1}H CPMG NMR. ❖The K_D for glutathione (**GSH**) was obtain using fluorescence emission spectra.[1] (B) The structures of the compounds tested.

As can be observed from the competition fluorescence data, removal of the succinimide ring from **^{t}BuSG**, to give glutathione results in a significant drop in affinity, which has been quantified as follows: glutathione has a K_D of 900 ± 200 μM (fluorescence emission spectra)[1] and **^{t}BuSG** has a K_D of 0.54 ± 0.03 μM (^{1}H CPMG NMR). Interestingly, inversion of the stereochemistry of the cysteine from L- to D- in compound **119** causes a drop in affinity of a similar magnitude to that observed when removing the succinimide ring, suggesting that the vector of the succinimide ring is important for maintaining affinity for *Sd*KefCTD.

A further lesson can be learnt about the role of the succinimide ring when considering the affinity of *S*-*N*-(5-fluoresceinyl)succinimido glutathione, **95**, which was synthesised in Chapter 4, Section 4.9.2. Compound **95** was found to have a dissociation constant of 1.4 ± 0.01 μM by ^{1}H CPMG NMR, and 9.1 ± 0.3 μM by microscale thermophoresis (Figure 5.7). Given that fluorescein has a molecular weight of 332 g/mol and is significantly more bulky than *tert*-butyl, the fact that **95** still binds with a moderate affinity suggests that *Sd*KefCTD is able to accommodate electrophiles conjugated to glutathione with a broad range of sizes. This finding makes biological sense given that Kef is a system for the detoxification of general electrophiles.

Probing the importance of the interaction of Q419 with the N*H* of the Cys-Gly amide bond showed that a loss in affinity was observed upon replacing the Cys-Gly amide of **101** (K_D = 1.8 ± 0.2 μM) with an ester to afford compound **125** (K_D = 20 ± 0.8 μM). This drop in affinity was comparable to that observed when the acetyl group is removed from the cysteine amine of **101** (K_D = 1.8 ± 0.2 μM) to give **102** (K_D = 25 ± 0.8 μM). This observation is further emphasised by the analogous (F_B/F_L – 1) values obtained for **102** and **125** in the competition fluorescence assay. This result suggests that both the carbonyl oxygen of the Glu-Cys amide and the N*H* of the Cys-Gly amide make an equal contribution to the Q419 interaction, consistent with each of them forming a hydrogen bonding interaction with Q419.

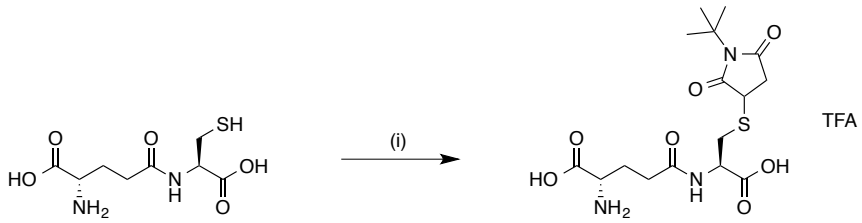
The final interaction of the central cysteine segment was the polar contact between the Cys-Gly amide carbonyl oxygen and the D506-R507 amide N*H*. As the racemate of compound **132** was being tested, the (F_B/F_L – 1) value was taken to be half of what would have been obtained had only the desired stereoisomers been tested. Removing the carbonyl of the Cys-Gly amide bond resulted in a complete loss in affinity. Even if the (F_B/F_L – 1) value is doubled, it is still only comparable to that of glutathione (K_D = 900 ± 200 μM; fluorescence emission spectra).[1] It is not clear whether this substantial drop of affinity is due to the loss of the interaction with the D506-R507 amide N*H*, or the increased conformational flexibility of **132** in the absence of the Cys-Gly amide bond.

5.10 Analysis of the glycine segment of ^{t}BuSG

Finally, attention was turned to the glycine of **^{t}BuSG**, which is proposed to make polar contacts through its carboxylate with R423 and R523 of *Sd*KefCTD. These residues have been shown to be important for affinity through the mutational studies performed by Miller *et al.*[4] on the equivalent residues of *Ec*KefC: R416 and R516. The X-ray crystal structure 3L9X (Figure 5.1) suggests that N551 in *Ec*KefC forms a water-mediated hydrogen bond with the Gly-carboxylate of the glutathione S-conjugate **ESG** *via* a structural water molecule. This proposed interaction is supported by mutational studies showing that N551 is important for glutathione binding.[5] It is possible that this water-mediated interaction is also present in *Sd*Kef with the residue D558 (the equivalent residue to N551 in *Sd*Kef). The first dissection therefore involved complete removal of the glycine from **^{t}BuSG**, a chemical mutation equivalent to biologically mutating R423 and R523.

5.11 Synthesis of the probes for the glycine segment of ^{t}BuSG

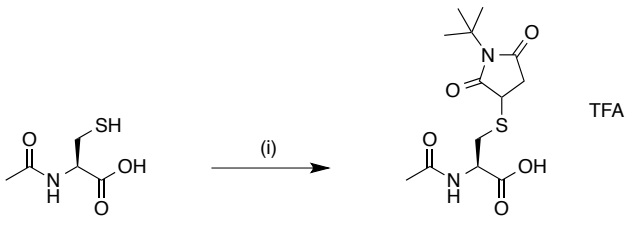
Complete removal of the glycine from **^{t}BuSG** was achieved through the synthesis of **134** using a conjugate addition of commercially available γ-L-Glu-L-Cys to *N*-*tert*-butylmaleimide, under the same conditions used to synthesise **^{t}BuSG**, to afford **134** in a moderate yield of 63% (Scheme 5.10).



133 **134**

Scheme 5.10 Synthesis of **134**. *Reagents and conditions*: (i) *N*-*tert*-Butyl-maleimide, NaOH, H_2O, RT, 4.5 h, 63%.

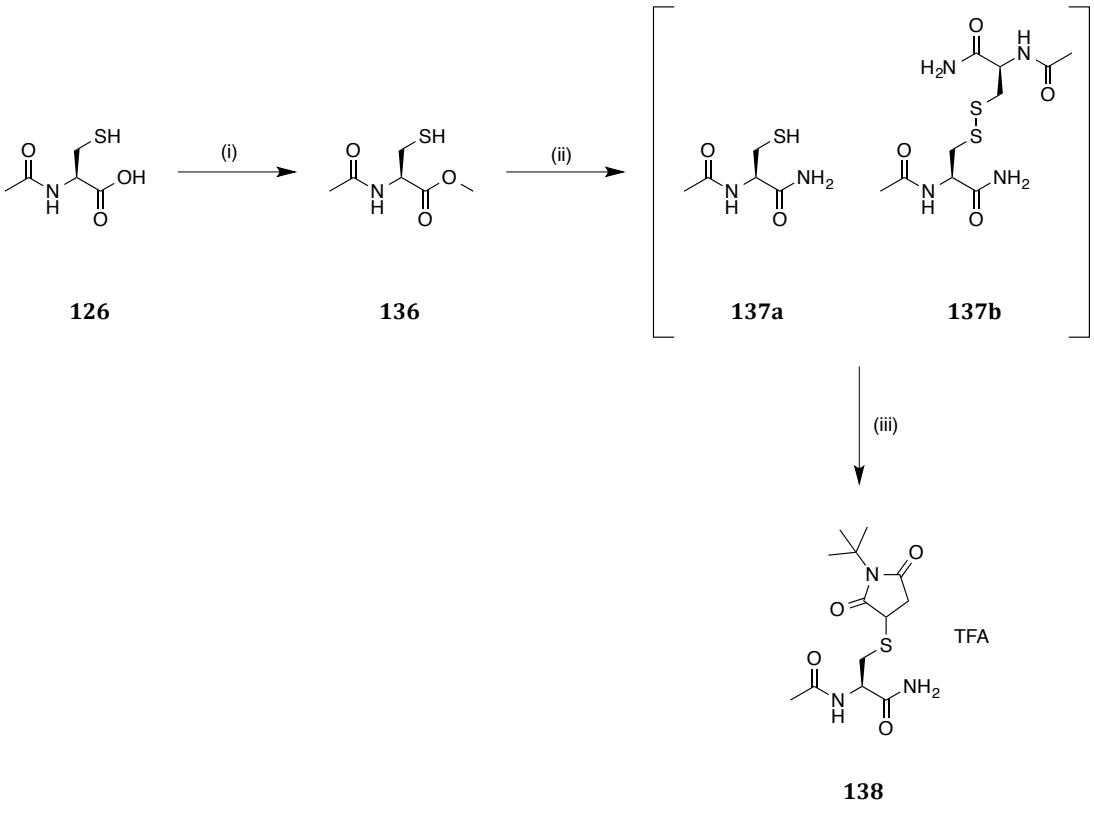
All of the remaining dissections of the glycine were performed using the more synthetically tractable acetylated dipeptide, **101**, as a reference scaffold. It was decided to remove the glycine from compound **101** to determine whether the central cysteine segment could retain some affinity through the interactions of the succinimide ring and hydrogen bonding to Q419. Two analogues were synthesised to investigate this, the first involved coupling commercially available *N*-acetyl-L-cysteine, **126**, to *N*-*tert*-butylmaleimide *via* a conjugate addition to afford the probe **135** in 67% yield (Scheme 5.11).



126 **135**

Scheme 5.11 Synthesis of **135**. *Reagents and conditions*: (i) *N*-*tert*-Butyl-maleimide, NaOH, H_2O, RT, 3.5 h, 67%.

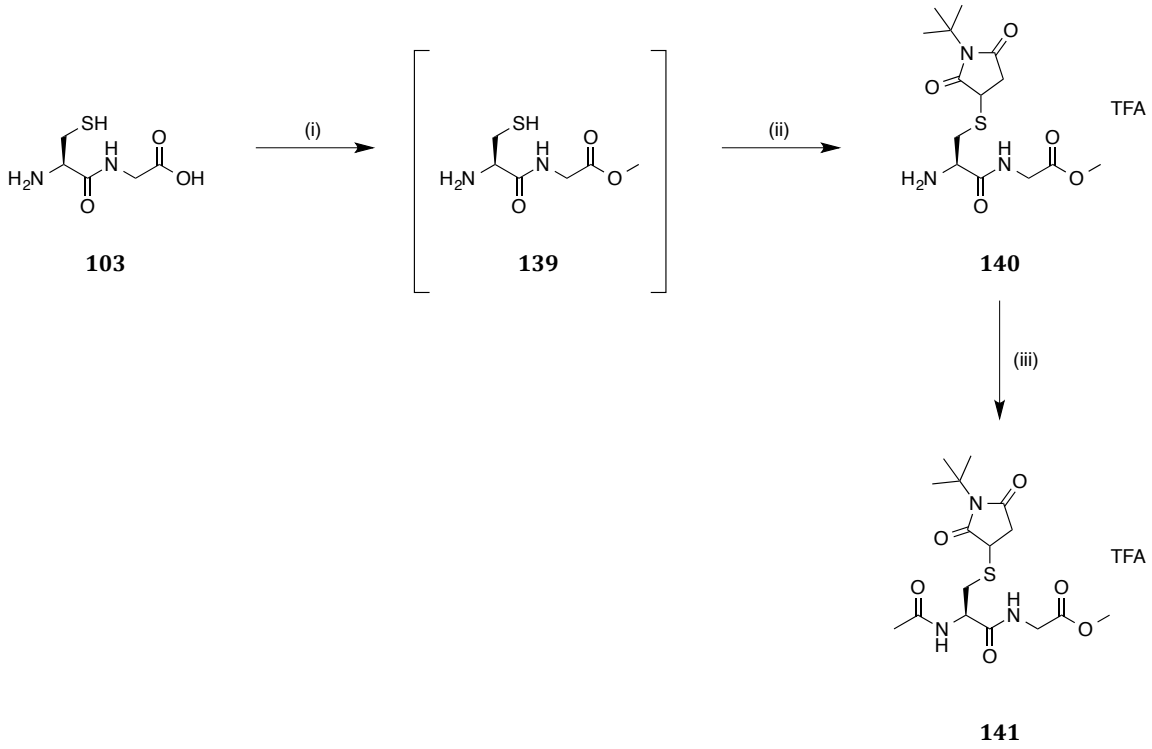
The second probe involved the reintroduction of the amide to **135** in case the absence of the NH was detrimental to the affinity due to one of the hydrogen bonding interactions with Q419 being removed. Synthesis of the second probe commenced by stirring *N*-acetyl-L-cysteine, **126**, in methanol in the presence of acetyl chloride to form the methyl ester, **136**. The resulting methyl ester, **136**, was then subjected to ammonium hydroxide to form *N*-acetyl-L-cysteinamide **137a**. However, it was found that the resulting product had partially oxidised to form the disulfide, **137b**. The following step was therefore modified to involve pre-stirring **137a** and the oxidised *N*-acetyl-L-cysteinamide, **137b**, with triphenyl-phosphine polymer-bound beads to form the reduced *N*-acetyl-L-cysteinamide **137a**. The resulting reduced thiol could then be coupled to *N*-*tert*-butyl-maleimide *via* conjugate addition to afford **138** in a 6% yield over two steps (Scheme 5.12).



126 **136** **137a** **137b** **138**

Scheme 5.12 Synthesis of **138**. *Reagents and conditions*: (i) Thionyl chloride, MeOH, RT, 1.5 h, 45%; (ii) Ammonium hydroxide, toluene, RT, 19 h; (iii) Triphenylphosphine, DMF, RT, 4 h, *N*-*tert*-Butyl-maleimide, NEt_3, DMF, RT, 144 h, 6% (over two steps).

To investigate whether there was scope for building off the Gly-carboxylate, the methyl ester of the acetylated dipeptide, **101**, was formed. It would be interesting to see whether the resulting ester could maintain interactions with R423 and R523 through the carbonyl oxygen and methoxyl group. The first step involved conversion of commercially available L-Cys-Gly, **103**, to the methyl ester, **139**, by stirring in methanol and thionyl chloride. The thiol of the resulting L-Cys-Gly methyl ester, **139**, was then coupled to *N*-*tert*-butylmaleimide *via* conjugate addition, before being acetylated with acetyl chloride to afford **141** in 64% yield (Scheme 5.13).

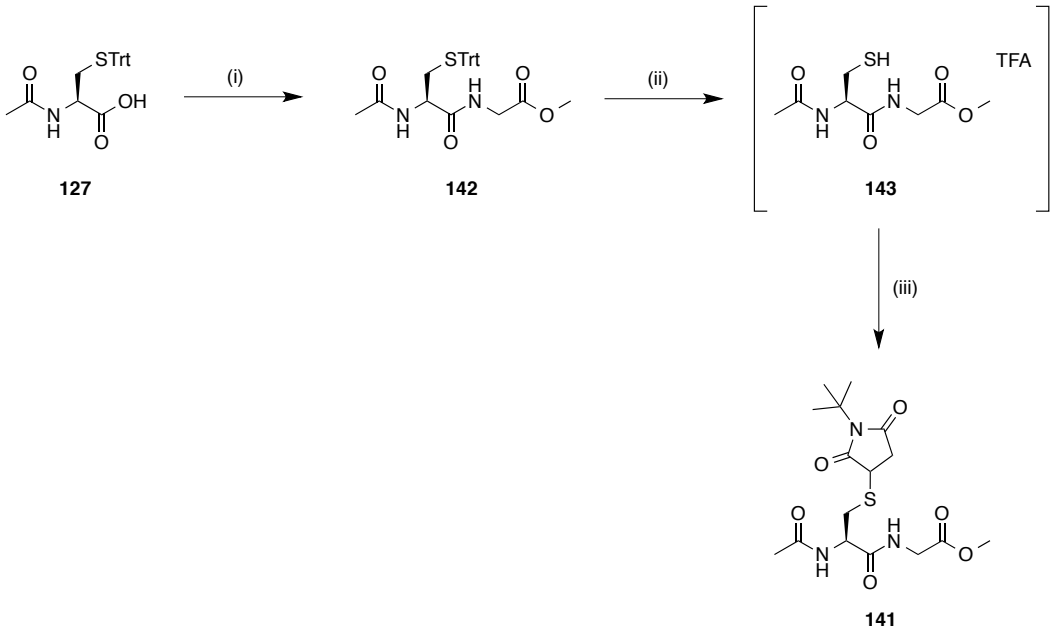


103 **139** **140** **141**

Scheme 5.13 Synthesis of **141**. *Reagents and conditions*: (i) Thionyl chloride, MeOH, RT, 2.5 h; (ii) *N*-*tert*-Butyl-maleimide, H_2O, RT, 8.5 h, 57% (over two steps); (iii) Acetyl chloride, NEt_3, THF, RT, 4 h, 64%.

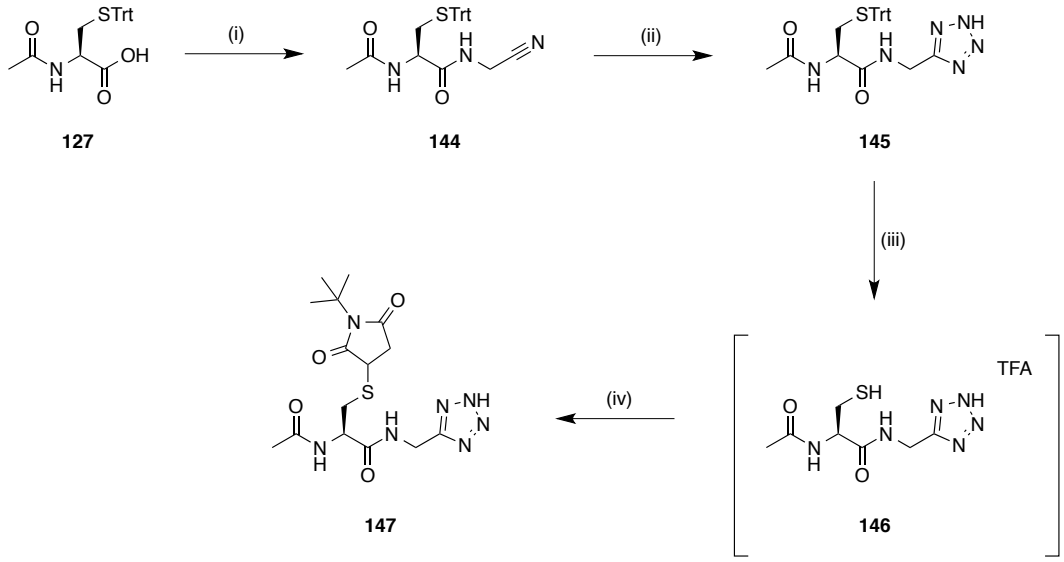
Compound **141** was found to be an important *in vivo* probe, as will be discussed later in this chapter. A route was therefore developed to allow for a scale-up synthesis of **141** to be employed. This route commenced with the coupling of *N*-acetyl-L-Cys(Trt)-OH, **127**, (synthesised in Scheme 5.9) to glycine methyl ester hydrochloride using the coupling reagent HBTU. The trityl

group was then deprotected to form the free thiol, **143**, which underwent a conjugate addition with *N*-*tert*-butylmaleimide to afford **141** in a 35% yield over two steps (Scheme 5.14).



Scheme 5.14 Synthesis of **141**. *Reagents and conditions*: (i) Glycine methyl ester hydrochloride, HBTU, DIPEA, DMF, RT, 24 h, 79%; (ii) Triethylsilane, 20% TFA in CH_2Cl_2, RT, 3 h; (iii) *N*-*tert*-Butylmaleimide, H_2O, RT, 9 h, 35% (over two steps).

To investigate whether there was scope for modifying the Gly-carboxylate to known carboxylic acid bioisosteres, a tetrazole analogue of **101** was synthesised. Synthesis of the tetrazole analogue, **147**, commenced with *N*-acetyl-L-Cys(Trt)-OH, **127**, (synthesised in Scheme 5.9) being coupled to aminoacetonitrile hydrochloride, using the coupling reagent HBTU, to afford **144** in 72% yield. Compound **144** was then subjected to a 1,3-dipolar cycloaddition with sodium azide in the presence of zinc bromide using microwave conditions, as reported by Shie *et al.*[8] Deprotection of the trityl group followed by coupling of *N*-*tert*-butylmaleimide to the free thiol, afforded the desired tetrazole probe **147** in a 29% yield over two steps (Scheme 5.15).



Scheme 5.15 Synthesis of **147**. *Reagents and conditions*: (i) Aminoacetonitrile hydrochloride, HBTU, DIPEA DMF, RT, 23 h, 72%; (ii) Sodium azide, zinc bromide, THF, microwave irradiation (80 W), 80 °C, 0.5 h, 25%; (iii) Triethylsilane, 20% TFA in CH_2Cl_2, RT, 5 h; (iii) *N-tert*-Butylmaleimide, H_2O, RT, 7 h, 29% (over two steps).

5.12 *In vitro* analysis of the glycine segment of ^{t}BuSG

The resulting chemical probes from the dissection of the glycine of **^{t}BuSG**, **134**, **135**, **138**, **141** and **147**, were subjected to a qualitative assessment of binding using the competition fluorescence assay, and a dissociation constant was determined for compound **147** using ^{1}H CPMG NMR (Figure 5.8). As can be seen from the (F_B/F_L – 1) values for compounds **134** and **135**, removal of the glycine completely abrogates binding to *Sd*KefCTD. Furthermore, it was not possible to restore the affinity through the reintroduction of the amide that interacts with Q419 (compound **138**). The (F_B/F_L – 1) value of compound **141** shows that the methyl ester is unable to maintain interactions with R423 and R523 through the carbonyl oxygen and methoxyl group and that the carboxylic acid is required for full affinity. The tetrazole analogue, **147**, on the other hand, was found to bind with a K_D of 7.6 ± 0.5 μM and (F_B/F_L – 1) value similar to that of the carboxylic acid, **101**, (K_D = 1.8 ± 0.2 μM). This result is promising for developing non-peptidic analogues of **101**, demonstrating that there is scope for introducing carboxylic acid bioisosteres.

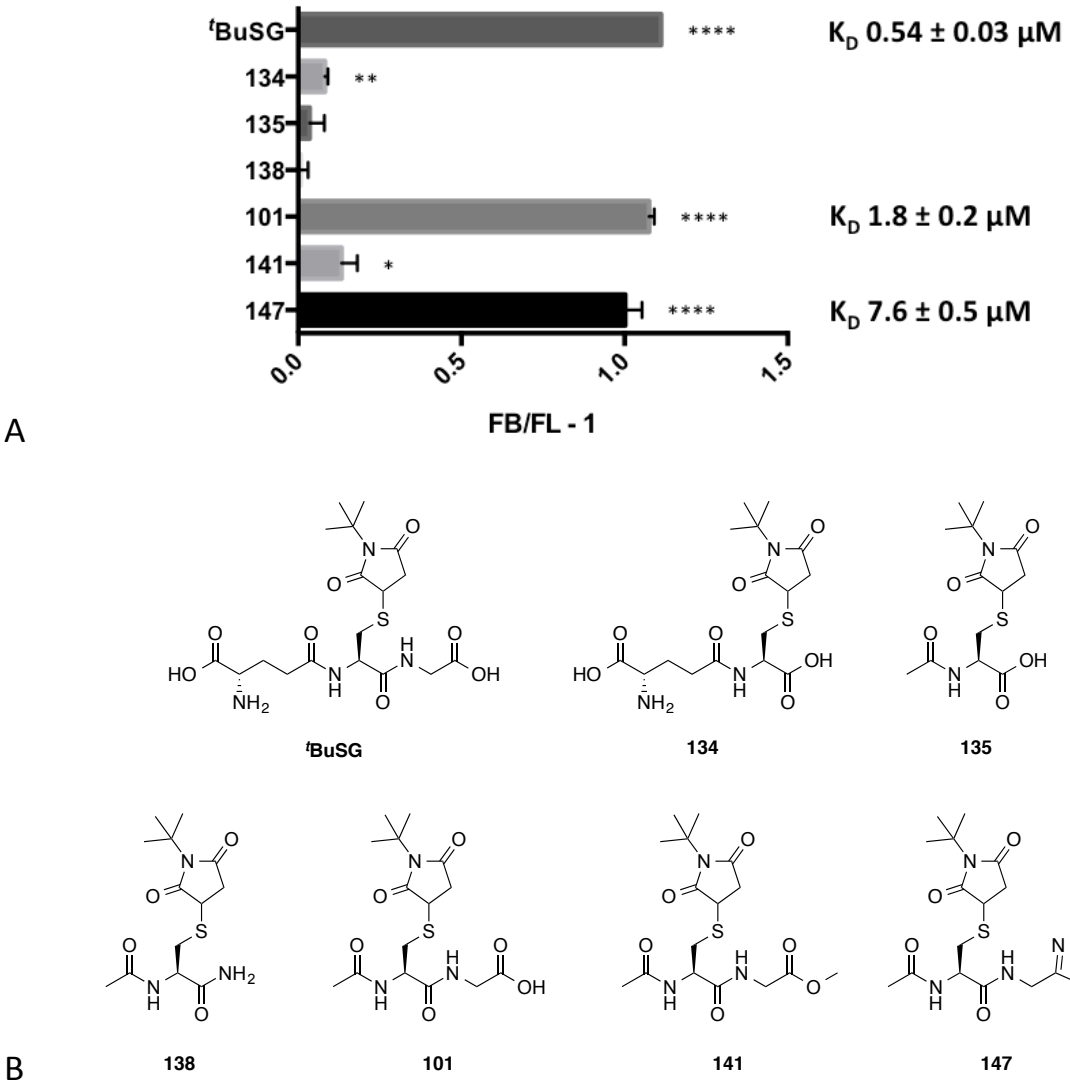


Figure 5.8 (A) Shows ($F_B/F_L - 1$) of the indicated compound (1 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). K_D values were obtained using ^{1}H CPMG NMR, assuming single site binding. K_D values are shown next to the ($F_B/F_L - 1$) bars corresponding to the compounds that they were obtained for. See the Appendix in Section 5.20 for Figures of the data obtained by ^{1}H CPMG NMR. (B) The structures of the compounds tested.

5.13 Group efficiencies of ^{t}BuSG and overall dissection conclusions

The group efficiency (GE), which is defined as the binding energy per heavy atom of the group, of each section of **^{t}BuSG** was determined using the quantitative data obtained from its dissection (Figure 5.9).

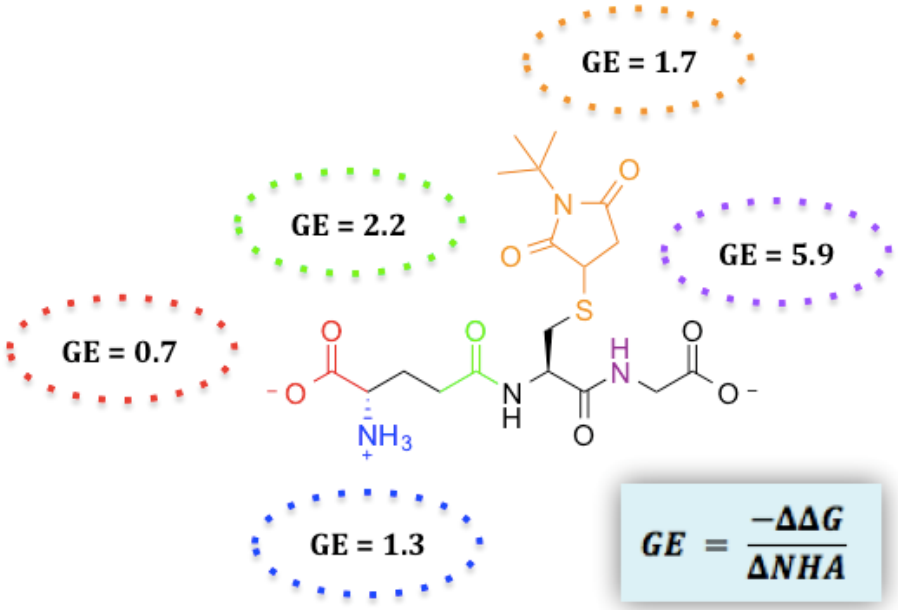


Figure 5.9 Group efficiencies (GE) calculated using K_D values obtained from the dissection of the glutathione S-conjugate **^{t}BuSG**. GEs are shown in circles colour-coded according to the group investigated. $-\Delta\Delta G$ is the change in Gibbs free energy caused by introducing the group and ΔNHA is the number of heavy atoms in the group introduced. GE values are calculated at 298K and have the following units: kJ mol^{-1} NHA^{-1}.

An example of how the GEs were calculated has been outlined in Equation 5.1 for the Glu-carboxylate, which was calculated to have a GE = 0.7 kJ mol^{-1} NHA^{-1}. The two compounds involved in this calculation are **102** and **112**. The Gibbs free energy (ΔG) of each of these compounds can be calculated from their K_D values using the equation on line (1) of Equation 5.1. By subtracting the Gibbs free energy of compound **102** from **112** it is possible to work out the change in Gibbs free energy ($-\Delta\Delta G$) upon introducing the deamino Glu-segment to compound **102** to give **112**. This information can be inserted into the numerator of the equation for GE:

lines (2), (3) and (4) of Equation 5.1. Finally the number of heavy atoms that are in the group introduced (ΔNHA), 8 in this instance, can be inserted into the denominator of the equation for GE: line (5) of Equation 5.1.

(1) $$\Delta G = -RT \ln K_D$$

(2) $$GE = \frac{-\Delta\Delta G}{\Delta NHA}$$

(3) $$GE = \frac{-\left(\Delta G_{\boldsymbol{compound\ 112}} - \Delta G_{\boldsymbol{compound\ 102}}\right)}{\Delta NHA}$$

(4) $$GE = \frac{-\left(\left[-RT \ln K_{D\ \boldsymbol{compound\ 112}}\right] - \left[-RT \ln K_{D\ \boldsymbol{compound 102}}\right]\right)}{\Delta NHA}$$

(5) $$GE = \frac{-\left(\left[-RT \ln\left(2.6 \times 10^{-6}\right)\right] - \left[-RT \ln\left(25 \times 10^{-6}\right)\right]\right)}{8}$$

$$= 0.7 \text{ kJ mol}^{-1} \text{ NHA}^{-1}$$

Equation 5.1 Worked through example of how the GE values were calculated using K_D values obtained by ^{1}H CPMG NMR for the dissected fragments **102** and **112** of the glutathione S-conjugate **^{t}BuSG**.

Removal of the glycine unit caused a complete loss in affinity for *Sd*KefCTD, such that it was not possible to determine a GE for the Gly-carboxylate. If it had been possible to obtain affinities for the weak fragments that lacked the Gly-carboxylate, it is likely that this part of the molecule would have been found to be the most efficient.

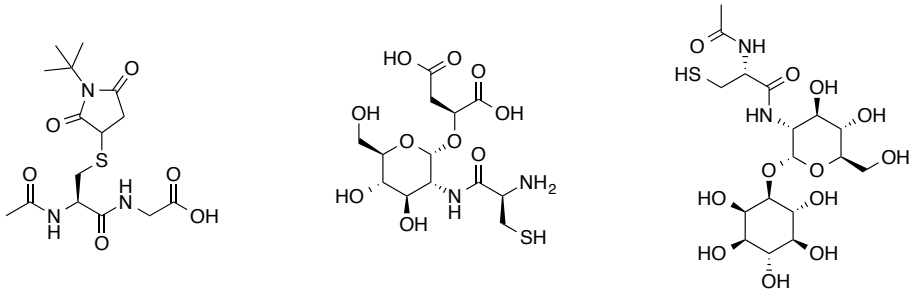
The next most efficient groups were the Cys-Gly amide N*H* (GE = 5.9), and the Glu-Cys amide carbonyl oxygen (GE = 2.2), which both interact with Q419. It is worth noting that the contribution to affinity from both of these groups is approximately the same, however, the group efficiency of the N*H* of the Cys-Gly amide bond is higher due to only one heavy atom changing upon forming the ester analogue. It would be interesting to see what effect removing both the N*H* of the Cys-Gly amide bond and the carbonyl of the Glu-Cys amide bond has on affinity, as you would expect to see complete abrogation, equivalent to the mutation Q419K.[1]

Analysis of the glutamate segment found that the most significant contribution came from the Glu-Cys amide carbonyl oxygen (GE = 2.2), followed by the Glu-amine (GE = 1.3). The Glu-amine is likely to be contributing to affinity through interactions with D506 and the drop in affinity observed upon removing it corresponds with the mutational studies performed on D499 in *Ec*KefC.[2] The Glu-carboxylate had the lowest group efficiency (GE = 0.7), which is in fitting with the R498A mutation data (unpublished data; Figure 5.4), and the fact that the equivalent residue of R498 in *Sd*KefCTD is I505. This finding supports the hypothesis that the Glu-carboxylate is redundant, making minimal contribution to the affinity of glutathione S-conjugates in *Sd*KefCTD.

In conclusion, the chemical mutations performed on **^{t}BuSG** correspond very well with the reported biological mutations of *Ec*KefC. A truncated dipeptide, **101**, has been synthesised that retains the majority of the affinity of full-length **^{t}BuSG**. Further dissection of this truncated peptide has found that three key interactions are made to *Sd*KefCTD: 1) the hydrogen bonds formed by the Glu-Cys amide carbonyl oxygen and the Cys-Gly amide N*H* with Q419; 2) a directional hydrophobic interaction of the succinimide ring with *Sd*KefCTD; 3) the polar contacts of the Gly-carboxylate with R423, R523 and potentially D558 *via* a structural water molecule. Furthermore, it was found that all three of these interactions are required to be intact simultaneously in order to maintain affinity. In addition to these three interactions, deletion of the Cys-Gly amide carbonyl oxygen was found to cause a complete loss in affinity. It is unclear as to whether this is due to a loss of interaction with the D506-R507 amide NH or a result of increased conformational flexibility. The information provided by this dissection will aid the overall goal of developing membrane permeant probes and progressing towards non-peptidic mimetics of glutathione and glutathione S-conjugates.

5.14 Synthesis and testing of ^{t}Bu-bacillithiol

The finding that the majority of the glutamate in **^{t}BuSG** is redundant raised questions as to why it was present and whether there was an evolutionary link between glutathione and alternative low-molecular-weight thiols, such as bacillithiol, **148**, and mycothiol, **149**, (Figure 5.10). These thiols are found in Gram-positive bacteria, indicating that the Kef mechanism of activation could be universal across Gram-negative and Gram-positive bacteria, potentially presenting a unified approach to targeting bacterial drug resistance. This hypothesis is supported by the findings of Chandrangsu *et al.* who demonstrated that upon the introduction of the toxic electrophile methyl glyoxal to *Bacillus subtilis*, *S*-lactoyl-bacillithiol is formed, which activates the KhtSTU K^+ effux pump.[9]

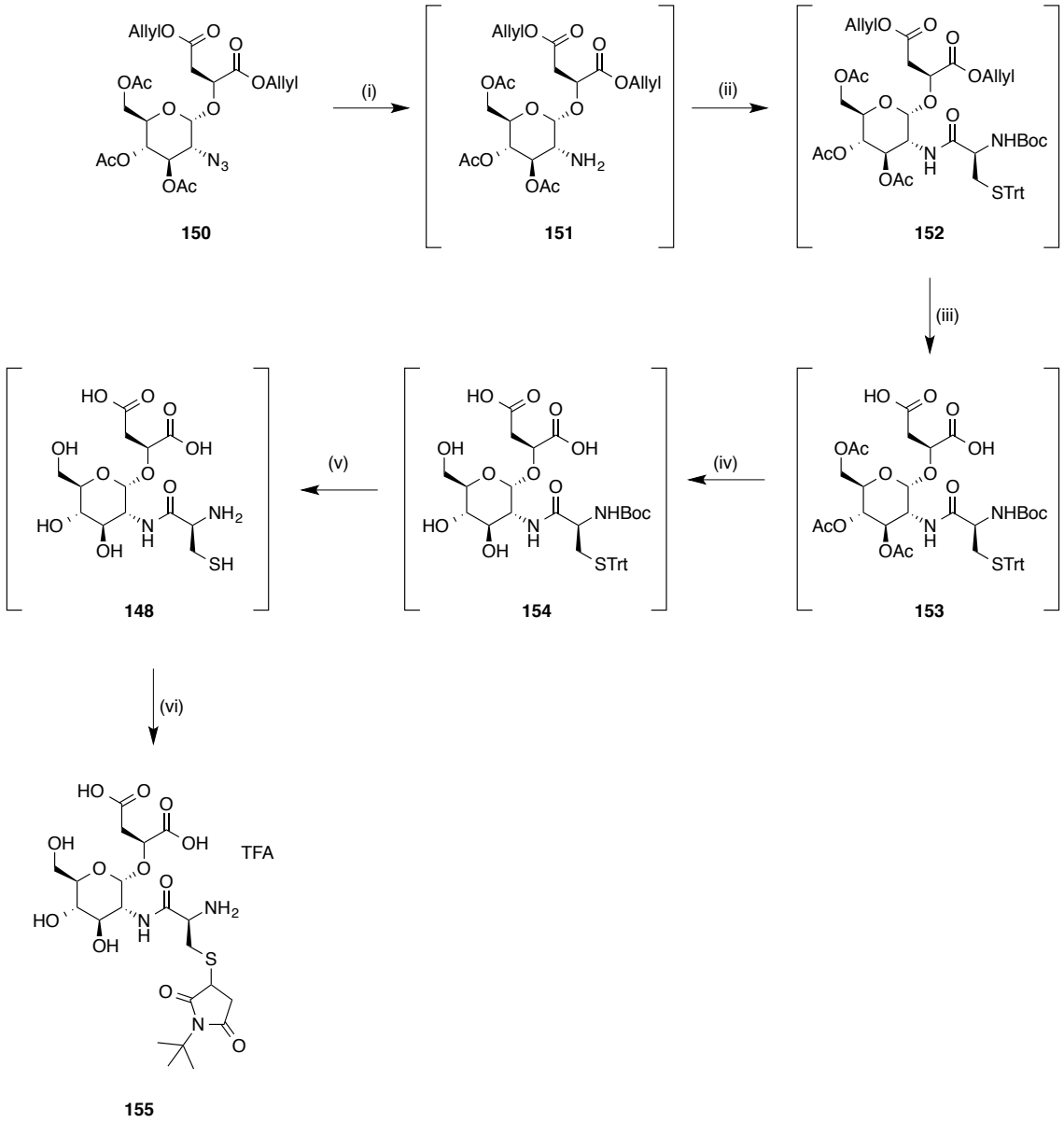


101 **148** **149**

Figure 5.10 Structures of the acetylated dipeptide **x**, bacillithiol **x**, and mycothiol **x**.

Comparison of the structure of bacillithiol, **148**, to the acetylated dipeptide, **101**, shows that bacillithiol has two carboxylic acids attached to the 1 position of the pyranose ring, which could potentially mimic the interactions of the Gly-carboxylate of **101**. The cysteine of bacillithiol, **148**, at the 2 position of the pyranose ring, also has no glutamate-like structure built off it. Mycothiol, **149**, on the other hand, has no obvious mimics of the Gly-carboxylate on **101**, but the cysteine amine is acetylated. It was proposed to test whether an S-conjugate of bacillithiol, **148**, and mycothiol, **149**, bound to *Sd*KefCTD given some of the structural similarities they have with **101**.

As the most synthetically tractable, bacillithiol, **148**, was selected for initial investigation, with Dr. Chris Hamilton (University of East Anglia) kindly providing the procedure and precursor, **150**, for synthesising bacillithiol. The *S*-*N*-*tert*-butylsuccinimido S-conjugate of bacillithiol, **155**, was synthesised as a direct comparison with the acetylated dipeptide, **101**. The synthesis commenced with an azide reduction using zinc dust, followed by coupling of the resulting amine, **151**, to *N*-Boc-L-(Trt)Cys-OH using the coupling reagent PyBOP. Global deprotection initially involved deallylation of the carboxylic acids using tetrakis(triphenylphosphine)palladium(0), followed by deacetylation of the hydroxyl groups (at positions 3, 4 and 6) using sodium methoxide. Finally, removal of the Boc-group was achieved by stirring in the presence of TFA to form bacillithiol, **148**. Bacillithiol, **148**, was coupled to *N*-*tert*-butylmaleimide *via* a conjugate addition reaction to afford the desired product **155** (Scheme 5.16).



Scheme 5.16 Synthesis of **155**. *Reagents and conditions*: (i) Zn, glacial acetic acid, RT, 4h; (ii) *N*-Boc-L-(Trt)Cys-OH, PyBOP, DIPEA, DMF, RT, 15.5 h; (iii) Imidazole, tetrakis(triphenylphosphine)palladium(0), triphenylphosphine, CH_2Cl_2, RT, 6 h; (iv) Sodium methoxide, MeOH, 0 °C, 3 h; (v) Triethylsilane, TFA in CH_2Cl_2, H_2O, 0 °C, 0.75 h; (vi) *N*-*tert*-Butyl-maleimide, H_2O, RT, 7 h.

To confirm that the α-anomer had been synthesised, the $^{1}J_{C1\text{-}H1}$ coupling constant was determined from the CLIP-HSQC spectrum of compound **155**. The value for $^{1}J_{C1\text{-}H1}$ was found to be 173 Hz, which is indicative of an α-pyranose: $^{1}J_{C1\text{-}H1}$ ≈ 160 Hz for β-pyranoses and ≈ 170 Hz for α-pyranoses (Figure 5.11).[10]

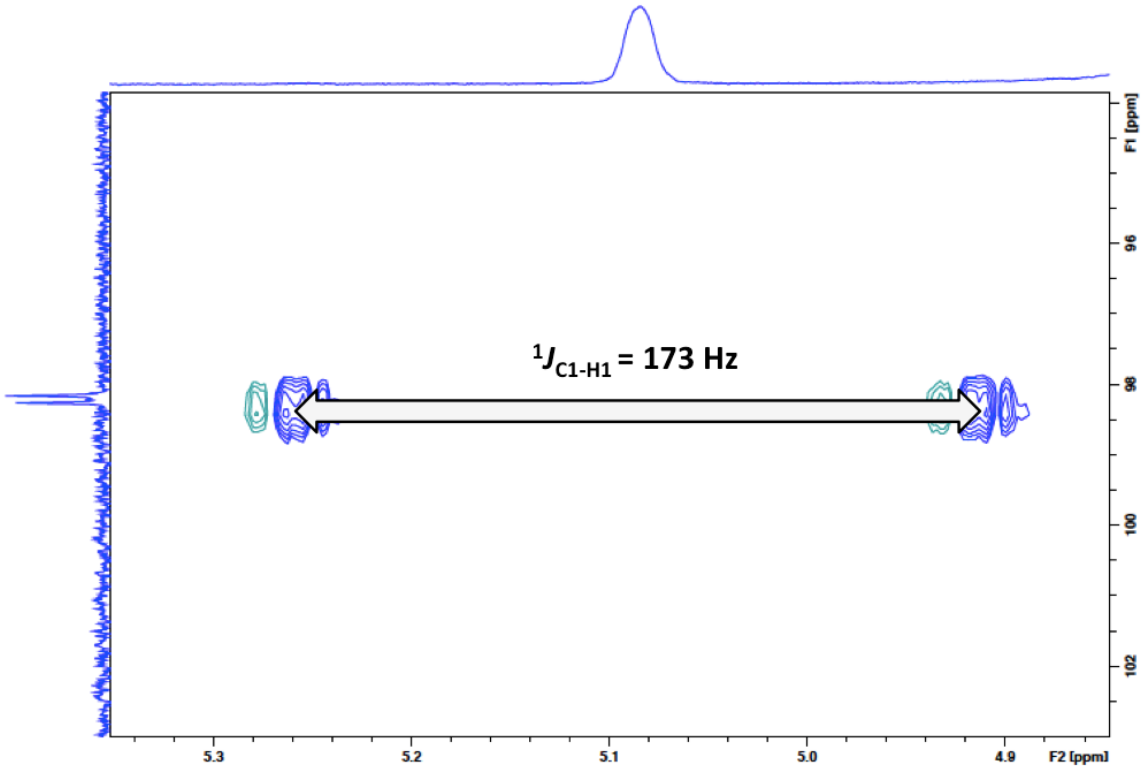
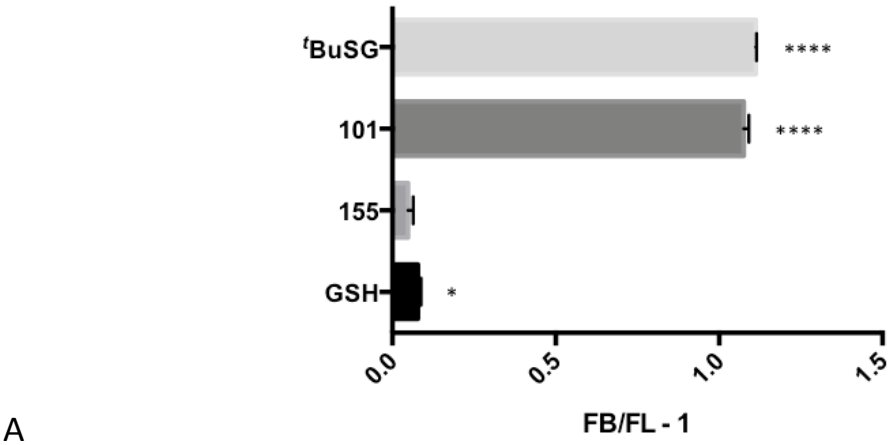
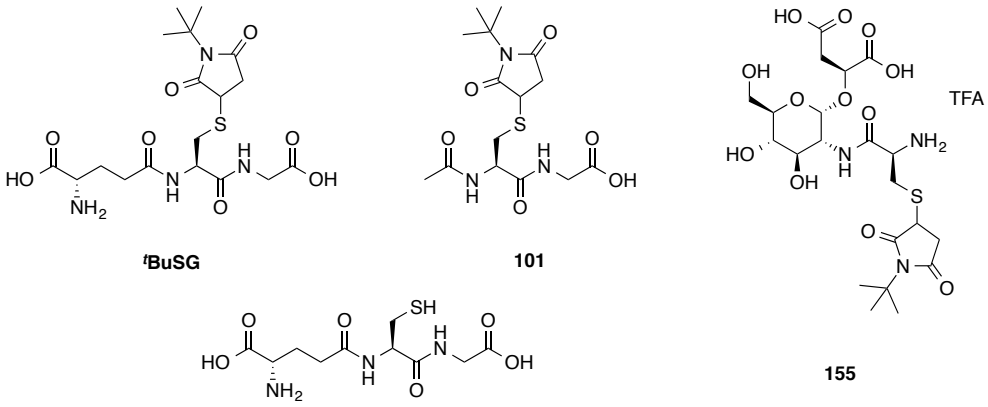


Figure 5.11 CLIP-HSQC spectrum highlighting the $^{1}J_{C1-H1}$ coupling constant of compound **155**. $^{1}J_{C1-H1}$ ≈ 160 Hz for β-pyranoses and ≈ 170 Hz for α-pyranoses.[10]

The ability of **155** to bind to *Sd*KefCTD was assessed using the competition fluorescence assay and ^{1}H CPMG NMR (Figure 5.12). Compound **155** did not appear to show any statistically significant binding to *Sd*KefCTD by either of these approaches and so this line of investigation was suspended (Figure 5.12).



A



B

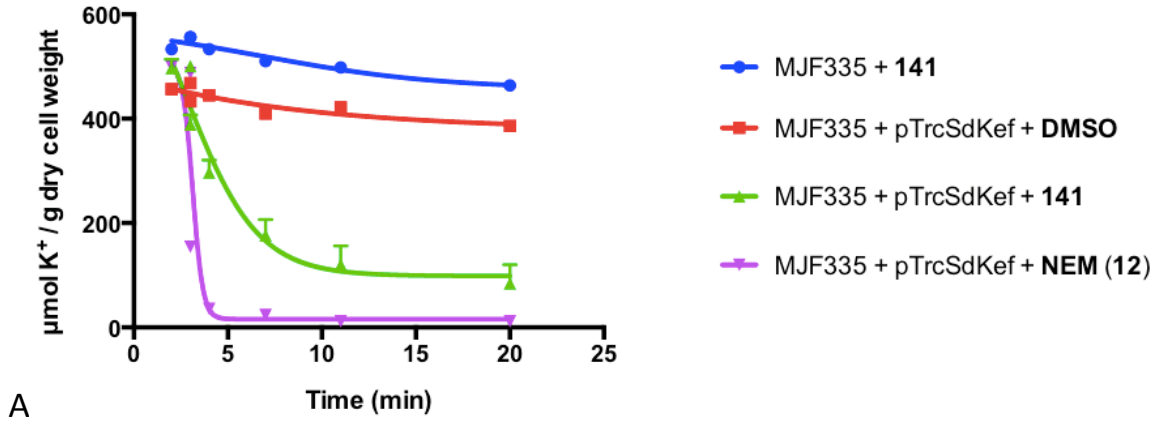
Figure 5.12 (A) Shows (F_B/F_L – 1) of the indicated compound (1 mM) at 525.5 nm emission in the presence of *Sd*KefCTD (6 µM) and **DNGSH** (5 µM). Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$). (B) The structures of the compounds tested.

5.15 K^+ efflux studies

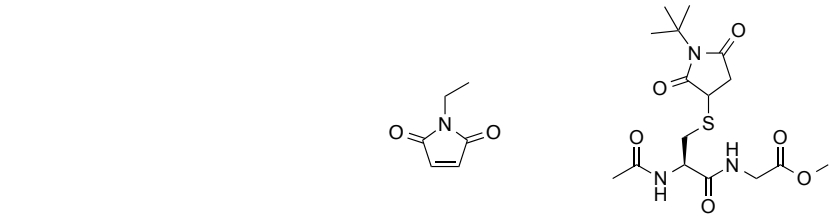
Having established that removing most of the glutamate component from **^{t}BuSG** has negligible effect on affinity *in vitro*, the effect of removing the glutamate on **^{t}BuSG**'s ability to activate full-length *Sd*KefCTD was investigated *in vivo*. In order to test the contribution of the glutamate *in vivo*, an analogue of **^{t}BuSG** was required that lacked the glutamate and was membrane permeant. Dr. Jess Healy showed in her PhD dissertation that the glutathione S-conjugate **ESG** is membrane impermeant, but that once it was converted to the ethyl ester it became membrane

permeant.[11] It was therefore speculated that the methyl ester, **141**, is membrane permeant, hence, Silvia Ekkerman in the Miller Group, University of Aberdeen, tested **141** in K^+ efflux studies. K^+ efflux studies look to measure the internal K^+ concentration of cells at time intervals after dosing with a compound of interest using flame photometry. If a compound is activating Kef, a drop in the internal K^+ concentration will be observed over time due to the efflux of K^+.

Although **141** had been found to lack affinity for *Sd*KefCTD in the competition fluorescence assay due to the loss of the interactions made by the Gly-carboxylate, it was thought that once inside the bacterium, it would be hydrolysed by cellular esterases to the active dipeptide **101**. The efflux experiments were performed in MJF335[4] *E. coli* cells, which had been transformed with pTrcSdKefH$_6$[1] (a wild-type, full-length *Sd*Kef high copy number plasmid). The MJF335 cell strain has its endogenous K^+ uptake (*kdpABC*), K^+ efflux (*kefB* and *kefC*) and glutathione synthesis (*gshA*) genes knocked out, so any K^+ efflux observed should be a result of activation of full-length *Sd*Kef. The first experiments run were the negative controls of MJF335 + **141** without pTrcSdKefH$_6$ (blue; Figure 5.13) and MJF335 + pTrcSdKefH$_6$ + **DMSO** (red; Figure 5.13), which both showed no efflux of K^+. The next experiment was the positive control of MJF335 + pTrcSdKefH$_6$ + the electrophile *N*-ethylmaleimide (**NEM**; **12**), which reacts with glutathione introduced into MJF335 to form **ESG** (purple; Figure 5.13). The positive control showed rapid efflux of K^+, within the space of 5 minutes. Finally, the experiment was run with MJF335 + pTrcSdKefH$_6$ + **141**, which showed that **141** was able to elicit efflux of K^+, albeit not as rapidly or extensively as the positive control **NEM** (green; Figure 5.13).



A



B

Figure 5.13 K^+ efflux experiments performed by Silvia Ekkerman using *N*-ethylmaleimide (**NEM**; **12**) and compound **141** to elicit efflux in the MJF335 strain transformed with pTrcSdKefH$_6$. The graph shows K^+ content (µmol K^+ / g dry cell weight) of the cells at various time points after dosing with the methyl ester, **141**, (0.5 mM) or **NEM**, **12**, (0.5 mM). Error bars indicate one standard deviation of uncertainty (n = 3 for MJF335 + pTrcSdKefH$_6$ + **141**; n = 1 for all other measurements).

The results of these studies were two-fold: firstly they showed that **141** is cell membrane permeable and presumably hydrolyses to the active compound, **101**; secondly that the glutamate segment of glutathione S-conjugates is not essential for their ability to active Kef, as K^+ efflux was observed upon dosing of the truncated dipeptide, **141**. The slight delay in K^+ efflux upon administration of **141** in comparison to **NEM**, **12**, could be attributed to the process of hydrolysing the methyl ester, **141**, to the carboxylic acid, **101**.

5.16 Kirby-Bauer Disc Diffusion Assay

Having demonstrated that **141** was cell permeant, and that dosing it into cells containing the plasmid for full-length *Sd*Kef elicited K^+ efflux, it was decided to use **141** to validate Kef as a novel antibacterial target. The assay of choice for demonstrating this was a Kirby-Bauer disc diffusion assay,[6] which was performed by Dr. Anthony Chan in the Conway Group, University of Oxford. This assay involves streaking an agar plate with bacteria, and then placing a disc dosed with a compound onto the plate. A zone of inhibition of growth / cell death should be observed around the disc if the compound has antibacterial activity. The diameter of the zone formed can provide a qualitative indication of the efficacy of the compound for eliciting inhibition of growth / cell death. Dr. Chan used the *E. coli* MJF335[4] cell strain for this study, dosing the discs with a negative control of Milli-Q™ water, and increasing concentrations of the methyl ester **141**: 100 μg; 200 μg; 400 μg.

The first experiment run involved dosing the methyl ester, **141**, onto the MJF335 cell strain lacking the pTrcSdKefH$_6$ plasmid (Figure 5.14 A). This was to show that **141** did not cause inhibition of growth / cell death through off-target effects. No zones were observed around the dosed discs for this experiment (Figure 5.14 A). The second experiment involved dosing **141** onto the MJF335 cell strain transformed with the pTrcSdKefH$_6$ plasmid, but no IPTG to induce expression of the *sdkef* gene (Figure 5.14 B). No zones were observed around the dosed discs for this experiment (Figure 5.14 B). The final experiment involved dosing **141** onto the MJF335 cell strain with the pTrcSdKefH$_6$ plasmid and IPTG to induce expression of the *sdkef* gene (Figure 5.14 C). As can be observed from Figure 5.14 C, zones of inhibition of cell growth / cell death were observed around the dosed discs for this experiment. The fact that the zones of inhibition of cell growth / cell death are only observed once the cells are induced with IPTG

suggests that the effect of the methyl ester, **141**, to elicit inhibition of growth / cell death is target specific: through the activation of overexpressed full-length *Sd*Kef.

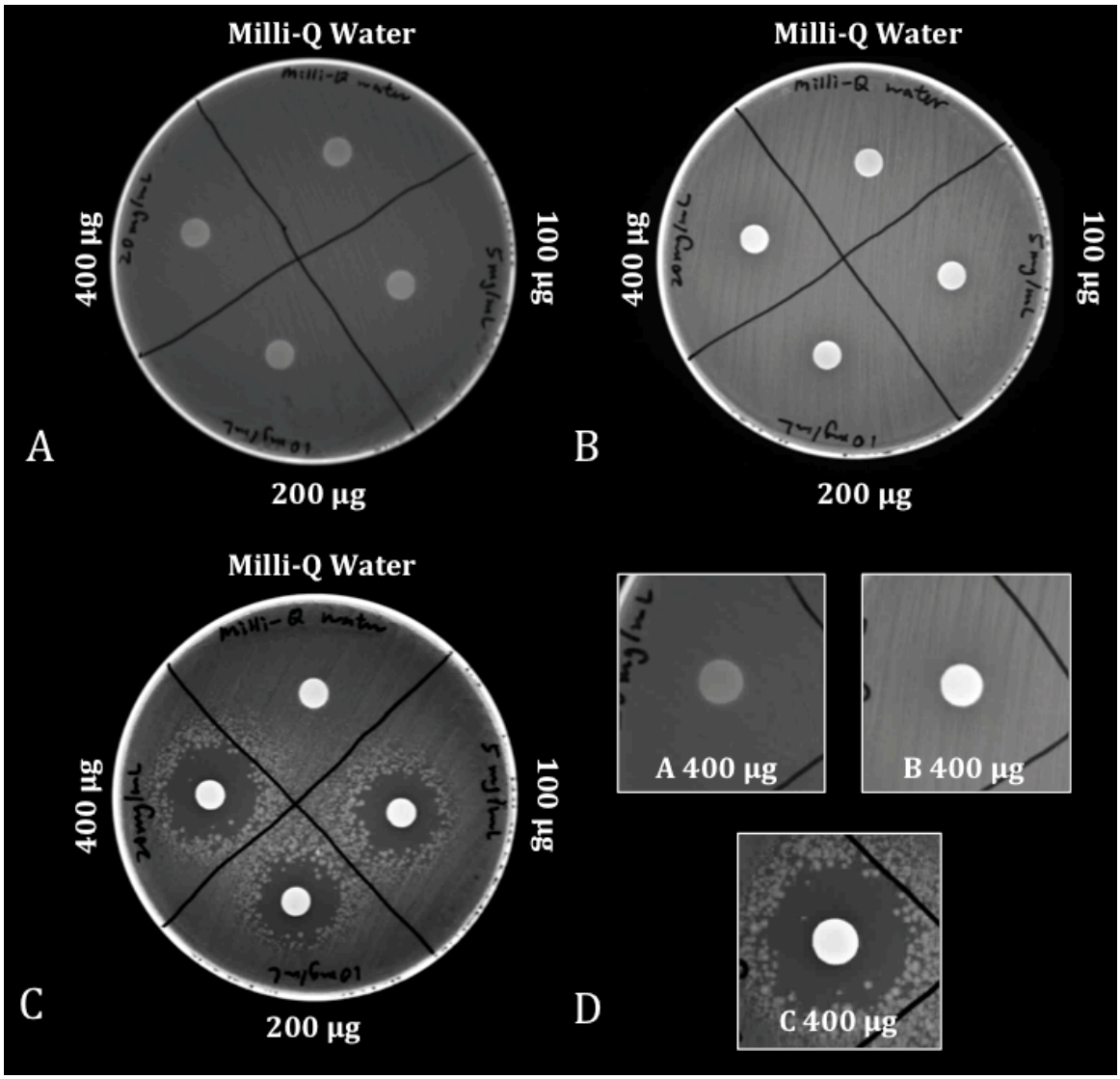


Figure 5.14 Kirby-Bauer disc diffusion assay assessing the ability of the methyl ester, **141**, to elicit inhibition of growth / cell death of *E. coli* MJF335 cells containing or lacking the plasmid pTrcSdKefH$_6$ for full-length *Sd*Kef. (A) *Escherichia coli* MJF335 without pTrcSdKefH$_6$ dosed with **141**; (B) *Escherichia coli* MJF335 with pTrcSdKefH$_6$, without IPTG dosed with **141**; (C) *Escherichia coli* MJF335 with pTrcSdKefH$_6$, with IPTG dosed with compound **141**. (A)-(C) have been dosed with 20 µL of the negative control Milli-Q™ water and the following concentrations of **141**: 5 mg/mL; 10 mg/mL; 20 mg/mL. (D) Magnified images of the discs dosed with 20 µL of **141** at a concentration of 20 mg/mL from experiments (A)-(C).

To investigate the dose dependence of the methyl ester, **141**, on the diameter of the zone of inhibition of growth / cell death around the disc, the methyl ester, **141**, was dosed at a range of concentrations (400-6.25 μg; serial dilution of factor 0.5; Figure 5.15). As can be observed from Figure 5.15, increasing the concentration of the methyl ester, **141**, resulted in a corresponding increase in the diameter of the zone of inhibition. This experiment further demonstrates that the observed inhibition of growth / cell death is being instigated by the methyl ester, **141**.

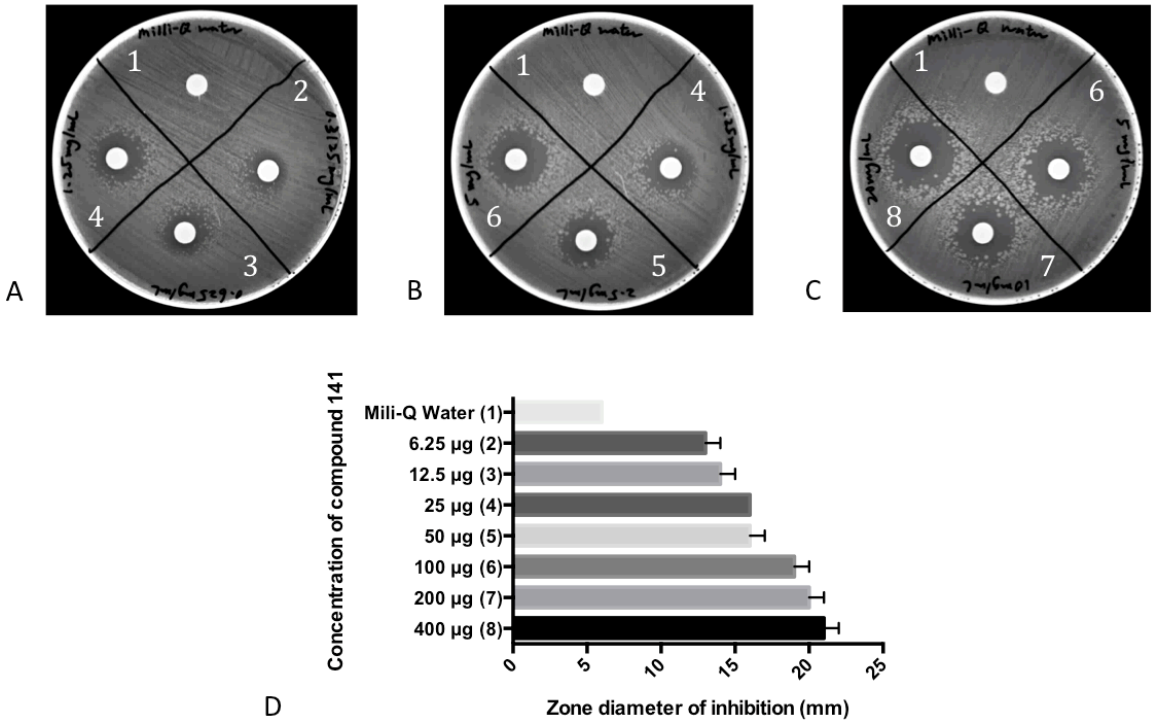
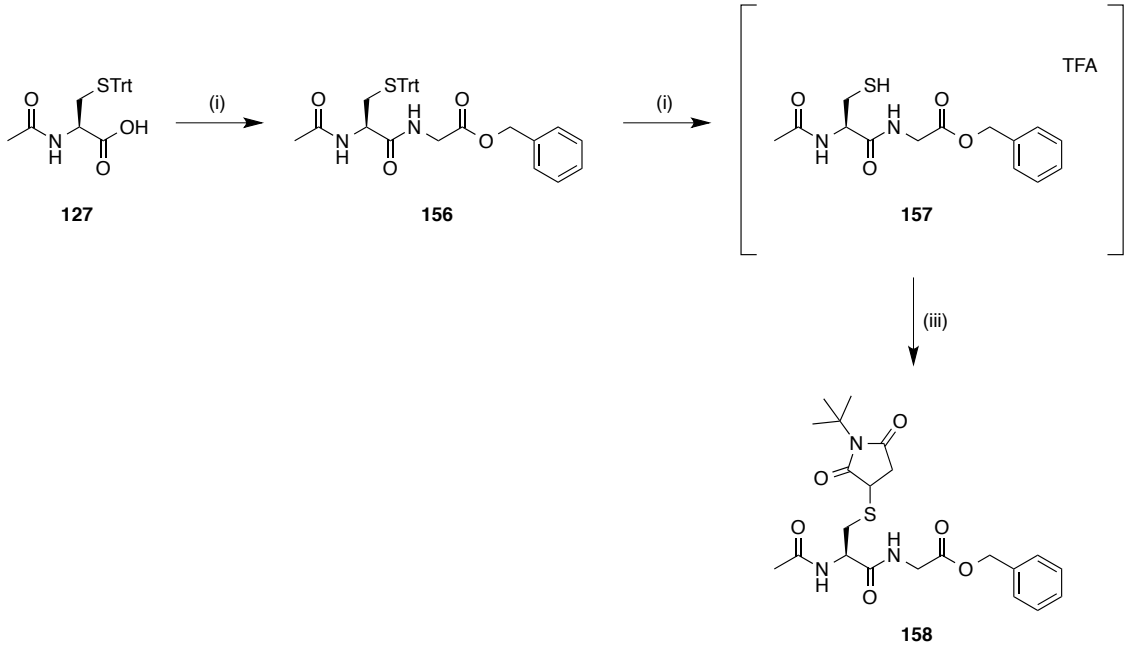


Figure 5.15 Kirby-Bauer disc diffusion assay analysing the dose dependence of the methyl ester, **141**, to elicit inhibition of growth / cell death of *E. coli* MJF335 cells containing the plasmid pTrcSdKefH_6 for full-length *Sd*Kef induced with IPTG. (A)-(C) Images of the plates for the dose dependence of the methyl ester, **141**, which was dosed at the following concentrations: (2) 6.25 μg; (3); 12.5 μg; (4) 25 μg; (5) 50 μg; (6) 100 μg; (7) 200 μg; (8) 400 μg. (D) A graph plotting the concentration that the methyl ester, **141**, was dosed against the diameter of the zone of inhibition (mm).

To investigate whether there was scope for modifying the methyl ester without affecting the properties of permeability and ability to undergo hydrolysis, the benzyl ester, **158**, was synthesised and tested by the Kirby-Bauer disc diffusion assay.[6] In addition to the benzyl ester, **158**, the carboxylic acid, **101**, and tetrazole, **147**, were tested to see if they could permeate the

membrane and activate full-length *Sd*Kef. The benzyl ester, **158**, was synthesised using exactly the same approach as the methyl ester, **101**, by using the glycine benzyl ester hydrochloride in the first step (Scheme 5.17). Purification of the final product, **158**, was found to be non trivial, with several runs by preparative HPLC being required to isolate it.



Scheme 5.17 Synthesis of **158**. *Reagents and conditions*: (i) Glycine benzyl ester hydrochloride, HBTU, DIPEA, DMF, RT, 25.5 h, 82%; (ii) Triethylsilane, 20% TFA in CH_2Cl_2, RT, 5 h; (iii) *N-tert*-Butylmaleimide, H_2O / MeOH, RT, 3 h, 24% (over two steps).

Testing the carboxylic acid, **101**, and tetrazole, **147**, using the Kirby-Bauer disc diffusion assay[6] found that they did not cause inhibition of growth / cell death of the MJF335 cell strain with the pTrcSdKefH_6 plasmid and IPTG to induce expression of the *sdkef* gene (Figure 5.16 A). This is presumably due to the fact that although the carboxylic acid, **101**, and tetrazole, **147**, have been shown to bind to *Sd*KefCTD *in vitro*, they are not able to permeate the cells to exert their effect *in vivo*. The benzyl ester, **158**, was, however, found to cause inhibition of growth / cell death of the MJF335 cell strain with the pTrcSdKefH_6 plasmid and IPTG to induce expression of the *sdkef* gene (Figure 5.16 B). This result suggests that the nature of the ester does not affect the resulting

activity and provides more support that compound **141** undergoes hydrolysis in the cell to form the active carboxylic acid, **101**, as it is unlikely that the glutathione-binding pocket would be able to tolerate the bulky benzyl group on the Gly-carboxylate.

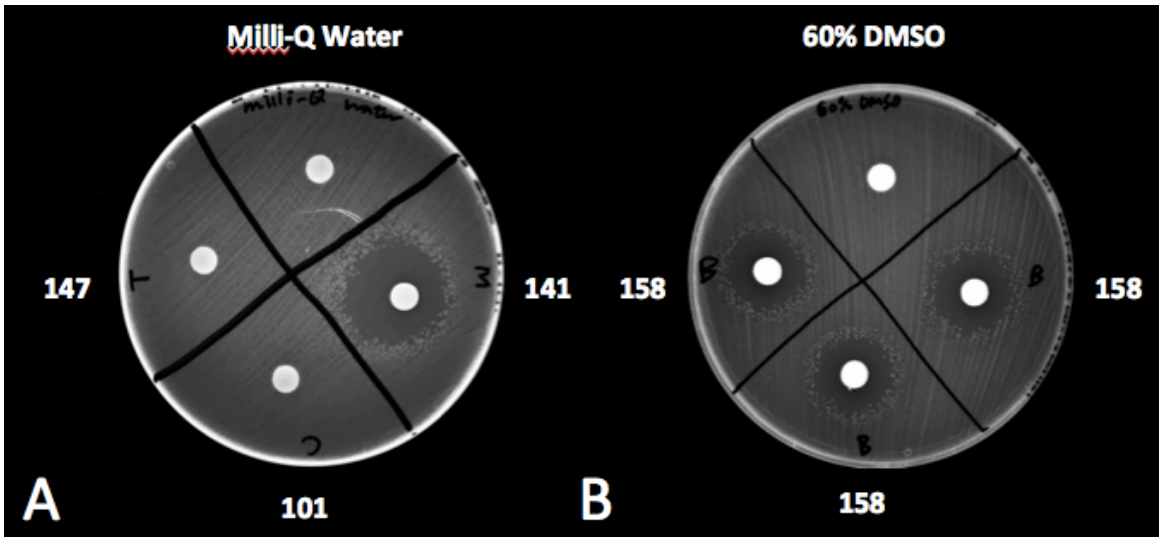


Figure 5.16 Kirby-Bauer disc diffusion assay analysing the ability of compounds **101**, **141**, **147** and **158** to elicit inhibition of growth / cell death of *E. coli* MJF335 cells containing the plasmid for full-length *S. denitrifications* Kef (pTrcSdKefH$_6$). (A) *Escherichia coli* MJF335 with pTrcSdKefH$_6$ and IPTG dosed with: the methyl ester, **141**, (M); the carboxylic acid, **101**, (C); the tetrazole, **147**, (T); (B) *Escherichia coli* MJF335 with pTrcSdKefH$_6$ and IPTG dosed with the benzyl ester, **158**, (B); (A) has been dosed with 20 µL of the negative control Milli-Q™ water and 20 mg/mL of **101**, **141** and **147**. (B) has been dosed with 20 µL of the negative control 60% **DMSO** in Milli-Q™ water and 16.7 mg/mL of the benzyl ester **158**.

5.17 A further negative control for the Kirby-Bauer Disc Diffusion Assay

As a further negative control, Alexander Axer synthesised the D-form of the methyl ester **141** (compound **159**). The rationale behind this was that the D-form of the free carboxylic acid **101** (compound **119**) had negligible affinity for *Sd*KefCTD in the competition fluorescence assay, and so by extension should be inactive in the Kirby-Bauer disc diffusion assay.[6] Testing compound **159** would provide further evidence that the inhibition of growth / cell death is attributed to **141**

activating *Sd*Kef, and not due to compound toxicity or off target effects caused by other functionalities in **141**. Compound **159** was dosed onto the MJF335 cell strain with the pTrcSdKefH_6 plasmid and IPTG to induce expression of the *sdkef* gene (Figure 5.17 A). As can be observed from Figure 5.17 A, zones of inhibition of cell growth / cell death were not observed and so compound **159** was effective as a negative control.

5.18 Assessing other homologues of Kef using the Kirby-Bauer Disc Diffusion Assay

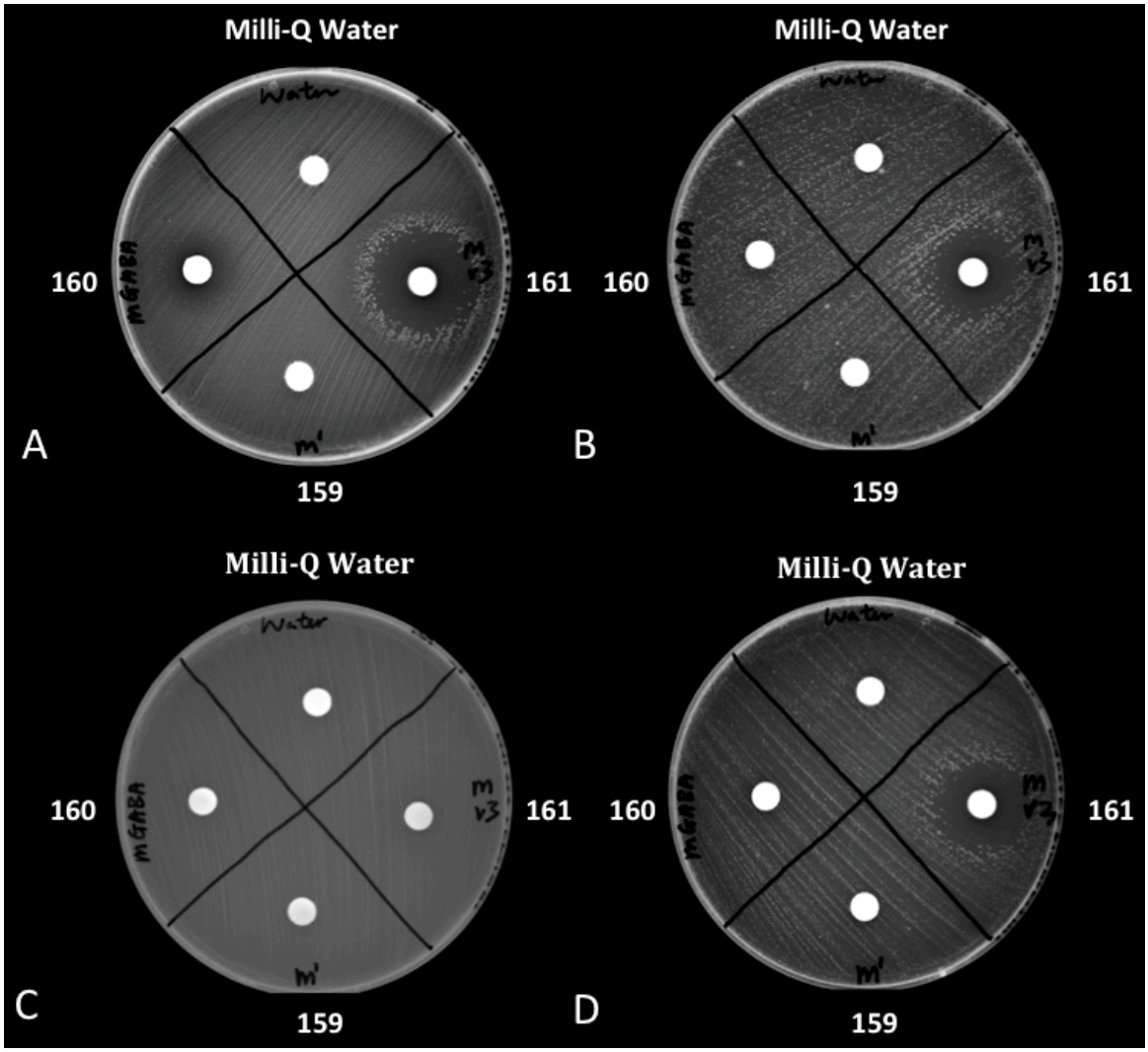
Having obtained a proof of concept for Kef as a target for the development of novel antibacterial agents using *Sd*Kef, it was decided to see if the validation could be extended to other, more clinically relevant homologues of Kef. The homologues selected for testing were *Ec*KefF-C, *Ec*KefG-B and a homologue of Kef from the ESKAPE pathogen *Acinetobacter baumanii* (*Ab*Kef). These homologues were tested by dosing the methyl ester, **141**, onto the MJF335 cell strain with plasmids containing the genes corresponding to each of these homologues and IPTG to induce expression. Unfortunately, the methyl ester, **141**, was unable to cause inhibition of growth / cell death in the presence of the other homologues of Kef (data not shown).

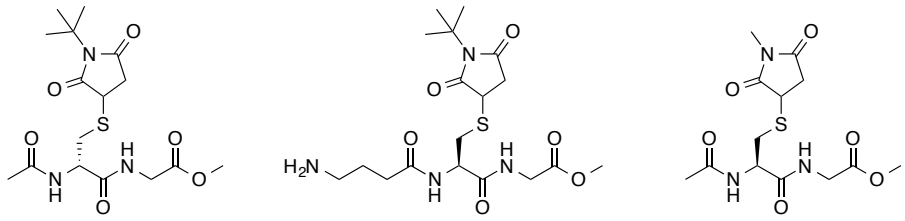
In order to understand why **141** was unable to elicit inhibition of growth / cell death in the other homologues of Kef, Alexander Axer and Dr. Anthony Chan synthesised and tested analogues of **141**. It was proposed that the glutamate may be more important for activation in these homologues and so the first analogue synthesised involved reintroducing the interaction with the residue D506. The resulting compound was the methyl ester of compound **109**, which has the NH_2 group of the glutamate present (compound **160**). The active form of this analogue should also be more potent than that of the methyl ester, **141**, as a slight improvement in the affinity was observed for compound **109** when compared to **101** in the ^{1}H CPMG NMR studies in Figure 5.5. Testing compound **160** using the Kirby-Bauer disc diffusion assay[6] found that it only

caused inhibition of growth / cell death in the strain containing the plasmid for *Sd*Kef (Figure 5.17 A-D).

The next proposal was that the other homologues might be more sensitive to the nature of the electrophile conjugated to the thiol, as there are examples in the literature where this degree of subtlety exists. For example, *S*-lactoylglutathione, **16**, the glutathione S-conjugate that arises from methylglyoxal, **13**, detoxification strongly activates *Ec*KefB, but only weakly activates *Ec*KefC.[12]

Analogues of the methyl ester, **141**, were therefore synthesised with different maleimides conjugated to the thiol. This approach led to a positive result, whereby the analogue **161**, which has a methyl group in place of the *tert*-butyl group on the succinimido ring was able to elicit inhibition of growth / cell death in MJF335[4] cell strains with plasmids containing the genes corresponding to each of the homologues, *Sd*Kef, *Ec*KefF-C and *Ab*Kef and IPTG to induce expression (Figure 5.17 A & B & D). Compound **161** was, however, still unable to cause inhibition of growth / cell death in MJF335 cell strains with plasmids containing the gene corresponding to *Ec*KefG-B and IPTG to induce expression (Figure 5.17 C). This phenomenon requires further investigation to provide an explanation but the negative result in *Ec*KefG-B is promising nonetheless as it shows that the results for *Ec*KefF-C and *Ab*Kef were unlikely to be related to compound toxicity or off-target effects.





159 **160** **161**

Figure 5.17 Kirby-Bauer disc diffusion assay assessing the ability of compounds, **159**, **160** and **161** to elicit inhibition of growth / cell death of *E. coli* MJF335 cells containing the plasmids: (A) pTrcSdKefH$_6$ for full-length *Sd*Kef; (B) pkC11 for full-length *Ec*KefF-C; (C) pKefGB for full-length *Ec*KefG-B; (D) pTrcAbKefH$_6$ for full-length *Acinetobacter baumanii* Kef (*Ab*Kef). (A) *Escherichia coli* MJF335 with pTrcSdKefH$_6$, with IPTG dosed with compounds **159**, **160** and **161**; (B) *Escherichia coli* MJF335 with pkC11, with IPTG dosed with compounds **159**, **160** and **161**; (C) *Escherichia coli* MJF335 with pKefGB, with IPTG dosed with compounds **159**, **160** and **161**; (D) *Escherichia coli* MJF335 with *Ab*Kef, with IPTG dosed with compounds **159**, **160** and **161**. (A)-(D) have been dosed with 20 µL of the negative control Milli-Q™ water and 20 mg/mL of compounds, **159**, **160** and **161**.

5.19 Conclusions

Dissection of a known glutathione S-conjugate, **^{t}BuSG**, has established that the key binding interactions to *Sd*KefCTD are made through the Gly-carboxylate, and the central cysteine segment through the Glu-Cys amide carbonyl oxygen, the Cys-Gly amide N*H* and a directional hydrophobic interaction of the succinimide ring. Knowledge that the majority of the glutamate makes minimal contribution to the affinity has allowed the removal of a carboxylic acid and primary amine from the peptide backbone to give a truncated peptide, **101**. Esterification of the Gly-carboxylic acid of the truncated peptide, **101**, has resulted in a membrane permeant analogue, **141**, of **^{t}BuSG** for use in *in vivo* studies. The *in vivo* K^+ efflux studies performed using **141** showed that the glutamate groups removed from the peptide backbone are not required to activate full-length *Sd*Kef. Finally, application of the methyl ester, **141**, to *E. coli* cells expressing the *sdkef* gene in a Kirby-Bauer disc diffusion assay[6] has demonstrated for the first time that endogenously applied small molecules targeting *Sd*Kef can elicit inhibition of growth / cell death. Furthermore, synthesis and testing of compound **141** has shown that other more clinically relevant homologues of Kef can be targeted, including *Ec*KefF-C and *Ab*Kef. Kef thus shows promise as a novel therapeutic target for the development of antibacterial agents.

5.20 Chapter 5 Appendix

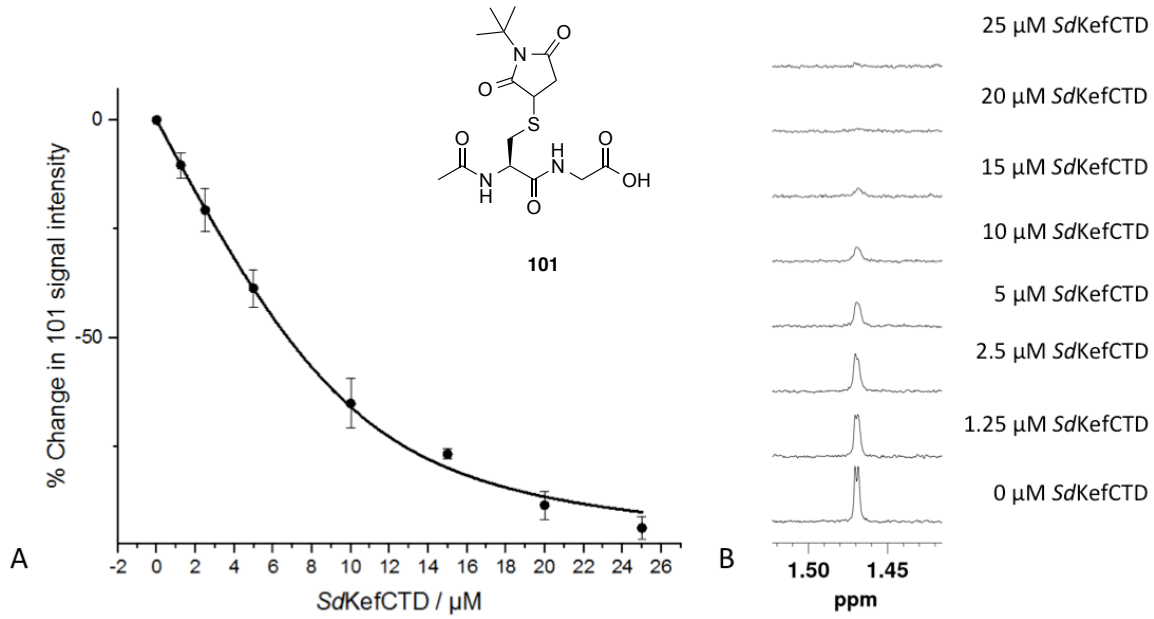


Figure 5.18 Analysis of the binding of **101** by ^{1}H CPMG NMR. (A) Shows the saturation-binding curve of **101** by ^{1}H CPMG NMR. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1; (B) highlights the attenuation of the **101** signal intensity upon titration of *Sd*KefCTD. It is worth noting that the ^{t}Bu signal in (B) appears as two singlets due to **101** being tested as a mixture of diastereomers.

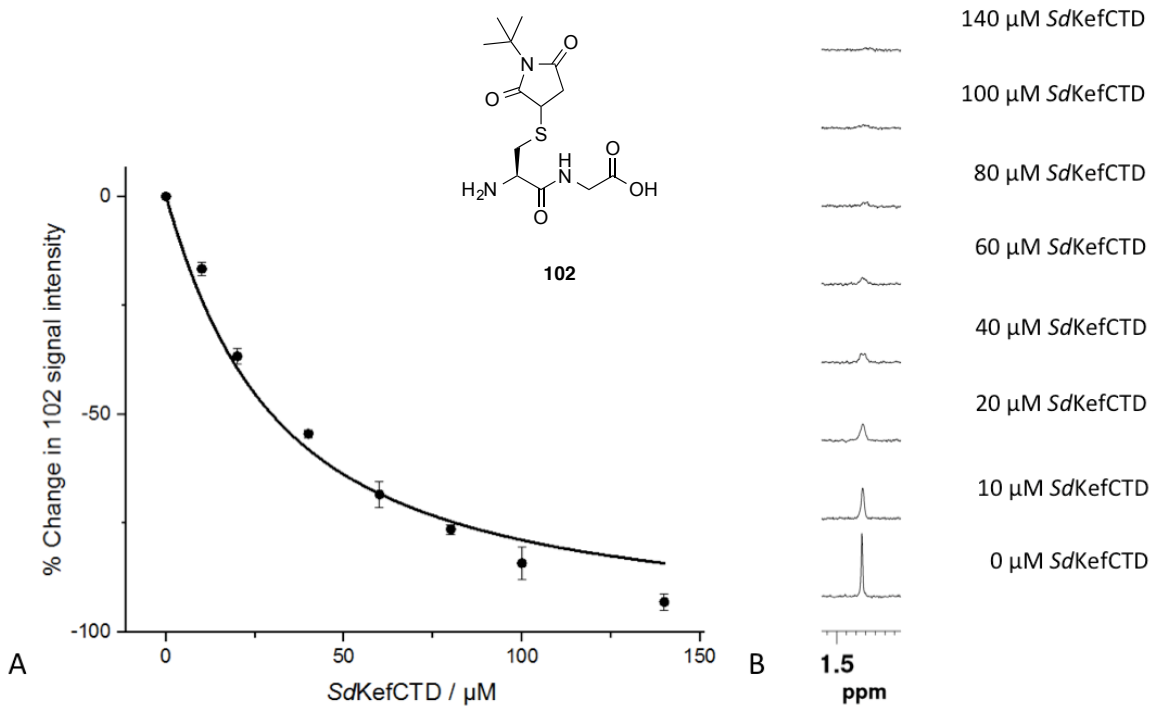


Figure 5.19 Analysis of the binding of **102** by ^{1}H CPMG NMR. (A) Shows the saturation-binding curve of **102** by ^{1}H CPMG NMR. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1; (B) highlights the attenuation of the **102** signal intensity upon titration of *Sd*KefCTD.

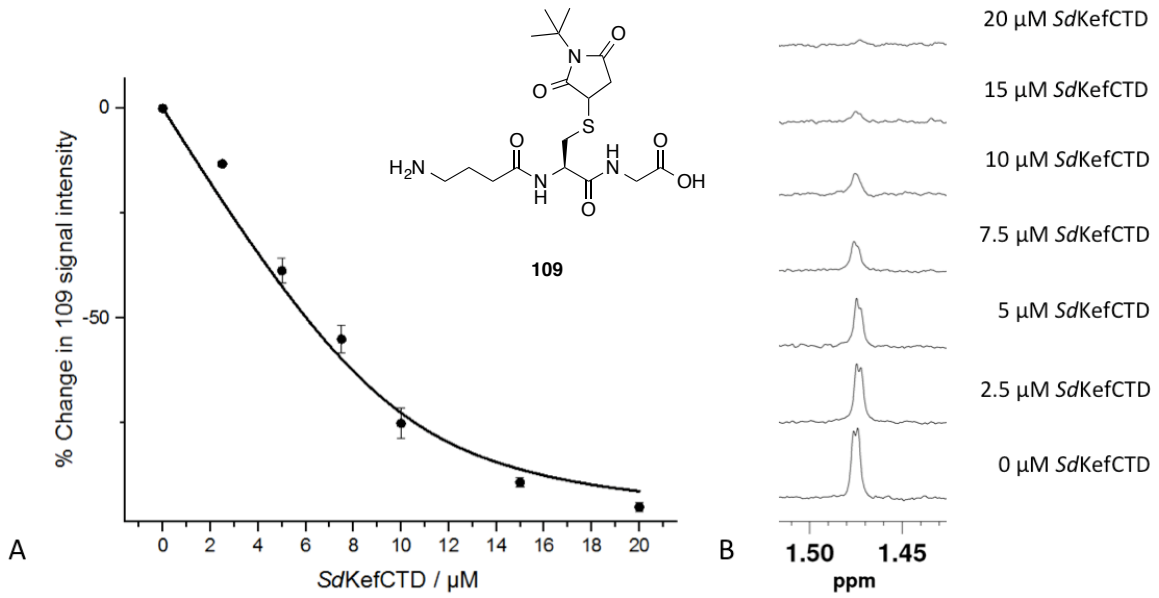


Figure 5.20 Analysis of the binding of **109** by ^{1}H CPMG NMR. (A) Shows the saturation-binding curve of **109** by ^{1}H CPMG NMR. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1; (B) highlights the attenuation of the **109** signal intensity upon titration of *Sd*KefCTD. It is worth noting that the ^{t}Bu signal in (B) appears as two singlets due to **109** being tested as a mixture of diastereomers.

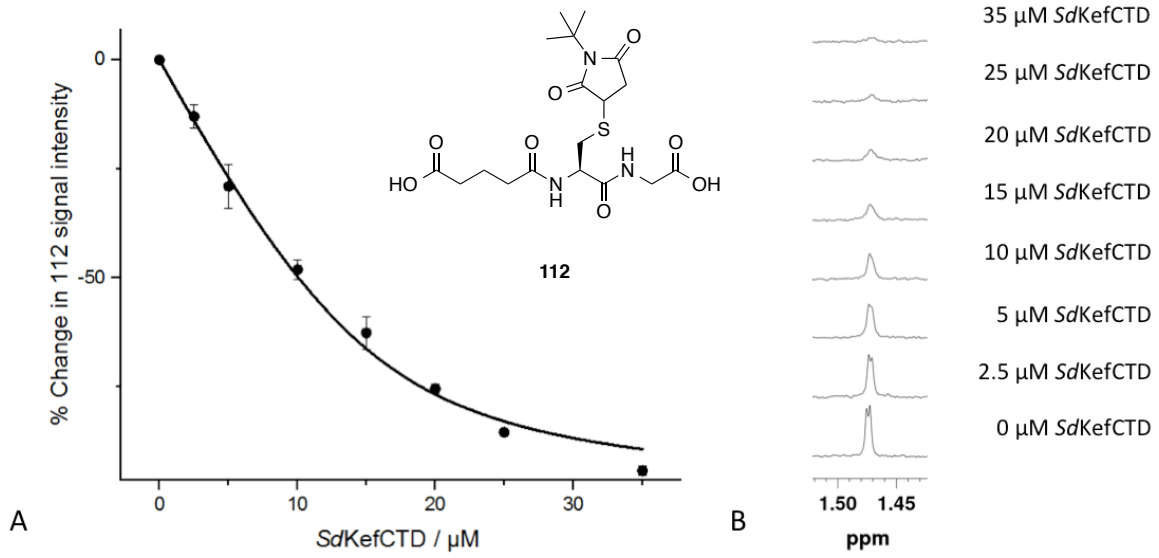


Figure 5.21 Analysis of the binding of **112** by ^{1}H CPMG NMR. (A) Shows the saturation-binding curve of **112** by ^{1}H CPMG NMR. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1; (B) highlights the attenuation of the **112** signal intensity upon titration of *Sd*KefCTD. It is worth noting that the ^{t}Bu signal in (B) appears as two singlets due to **112** being tested as a mixture of diastereomers.

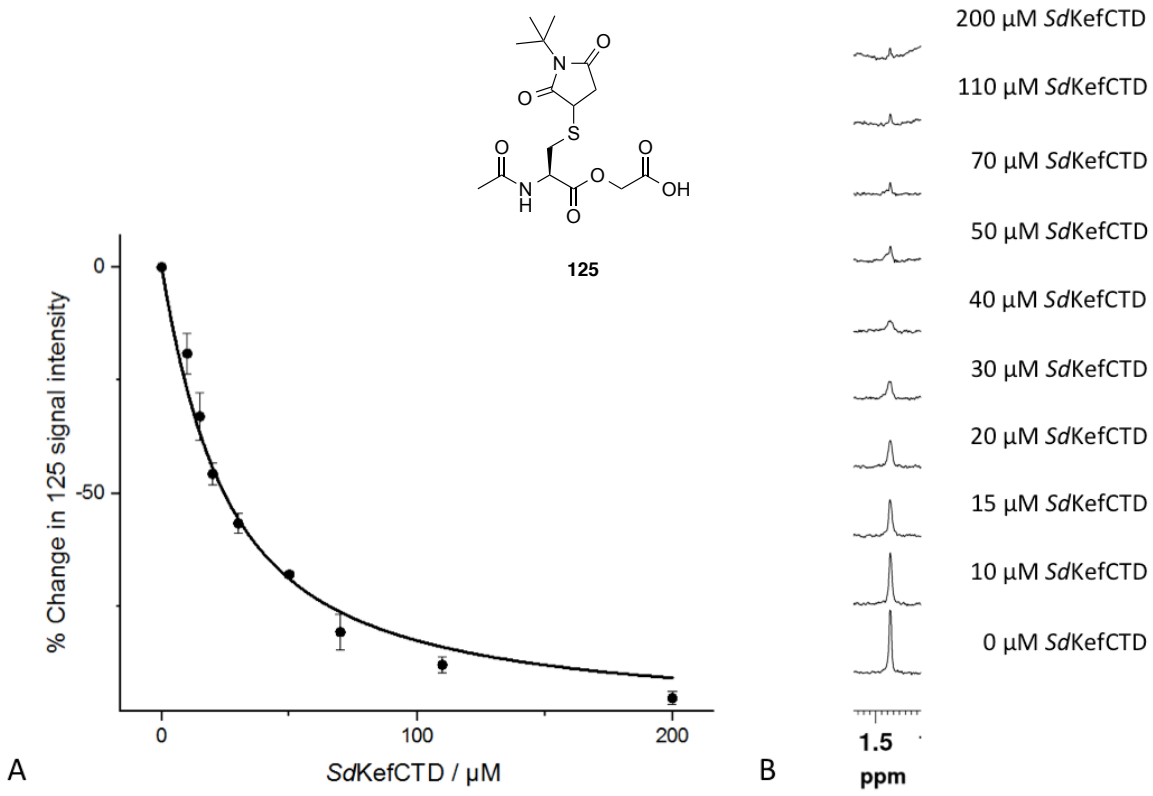


Figure 5.22 Analysis of the binding of **125** by ^{1}H CPMG NMR. (A) Shows the saturation-binding curve of **125** by ^{1}H CPMG NMR. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1; (B) highlights the attenuation of the **125** signal intensity upon titration of *Sd*KefCTD.

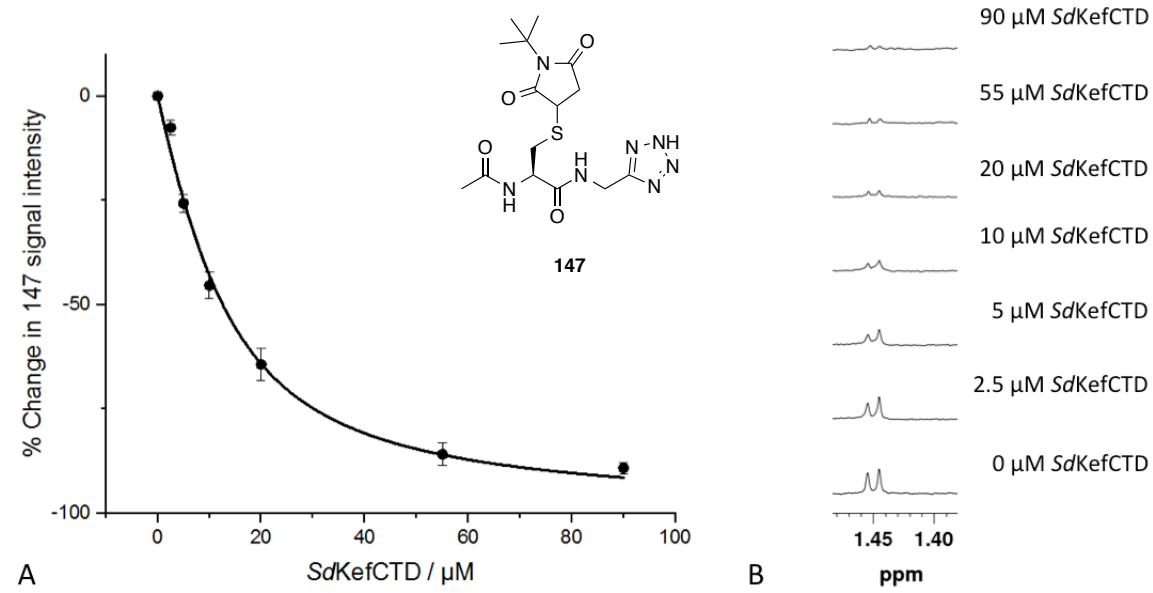


Figure 5.23 Analysis of the binding of **147** by ^{1}H CPMG NMR. (A) Shows the saturation-binding curve of **147** by ^{1}H CPMG NMR. Error bars indicate one standard deviation of uncertainty (n = 3). Curve fitted using Origin 9.1; (B) highlights the attenuation of the **147** signal intensity upon titration of *Sd*KefCTD. It is worth noting that the ^{t}Bu signal in (B) appears as two singlets due to **147** being tested as a mixture of diastereomers.

5.21 Chapter 5 References:

[1] J. Healy, S. Ekkerman, C. Pliotas, M. Richard, W. Bartlett, S. C. Grayer, G. M. Morris, S. Miller, I. R. Booth, S. J. Conway, et al., *Biochemistry* **2014**, *53*, 1982–1992.

[2] T. P. Roosild, S. Castronovo, J. Healy, S. Miller, C. Pliotas, T. Rasmussen, W. Bartlett, S. J. Conway, I. R. Booth, *PNAS* **2010**, *107*, 19784–19789.

[3] The PyMOL Molecular Graphics System, Version 1.7.4 Schrodinger, LLC, **2010**.

[4] S. Miller, R. M. Douglas, P. Carter, I. R. Booth, *J. Biol. Chem.* **1997**, *272*, 24942–24947.

[5] T. P. Roosild, S. Castronovo, S. Miller, C. Li, T. Rasmussen, W. Bartlett, B. Gunasekera, S. Choe, I. R. Booth, *Structure* **2009**, *17*, 893–903.

[6] A. W. Bauer, W. M. Kirby, J. C. Sherris, M. Turck, *Am. J. Clin. Pathol.* **1966**, *45*, 493–496.

[7] C. E. Hoyle, C. N. Bowman, *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 1540–1573.

[8] J. J. Shie, J. M. Fang, *J. Org. Chem.* **2007**, *72*, 3141–3144.

[9] P. Chandrangsu, R. Dusi, C. J. Hamilton, J. D. Helmann, *Molecular Microbiology* **2014**, *91*, 706–715.

[10] K. Bock, C. Pedersen, *J. Chem. Soc., Perkin Trans. 2* **1974**, 293–297.

[11] J. Healy, Studies on the Bacterial Potassium Efflux System KefC, and Its Ancillary Protein KefF, University of Aberdeen, **2010.**

[12] G. P. Ferguson, A. W. Munro, R. M. Douglas, D. McLaggan, I. R. Booth, *Molecular Microbiology* **1993**, *9*, 1297–1303.

Chapter 6:

Conclusions and Future Work

6. Conclusions and future work

6.1 Conclusions

A construct of the C-terminal domain of *Shewanella denitrificans* Kef (*Sd*KefCTD) has successfully been expressed and purified. A 2.9 Å resolution X-ray crystal structure of *Sd*KefCTD was resolved by Dr. Christos Pliotas, which has additional electron density observed in the Rossmann fold that has been attributed to the nucleotide **AMP**. DSF analysis, CPMG edited ^{1}H NMR and analytical HPLC experiments were performed to confirm the identity of the nucleotide to be **AMP**. It was concluded that **AMP** is likely to play a key role in stabilising *Sd*KefCTD, with the majority of the stability coming through interactions of the **AMP** phosphate group.

In silico protein-ligand docking and ligand-based virtual screening were performed in collaboration with industrial sponsor InhibOx to identify novel ligands of *Sd*KefCTD. Compounds predicted to bind by these approaches were tested using a competition fluorescence assay and STD NMR. A number of compounds were found to bind using both of these biophysical assays and series of analogues were synthesised based on them. However, after implementing more rigorous negative controls it became apparent that the compounds were in fact false positives. Alternative assays were therefore developed to obtain more reliable, quantitative data for ligands of *Sd*KefCTD.

Four quantitative assays were successfully developed and applied to *Sd*KefCTD. Two of these were fluorescence based, microscale thermophoresis and fluorescence polarisation, and two were NMR based, ^{1}H CPMG NMR and ^{19}F NMR. Each of these assays can be used in future screening efforts to identify novel ligands of *Sd*KefCTD. Employing the assays in tandem should allow genuine hits to be distinguished from false positives more reliably.

The competition fluorescence assay and ^{1}H CPMG NMR were subsequently implemented in the dissection of a known glutathione S-conjugate, **^{t}BuSG**. The key binding interactions of **^{t}BuSG** to *Sd*KefCTD are established through the Gly-carboxylate, and the central cysteine segment through the Glu-Cys amide carbonyl oxygen, the Cys-Gly amide N*H* and a directional hydrophobic interaction of the succinimide ring. The information obtained from this dissection allowed the development of a membrane permeant truncated analogue of **^{t}BuSG**, compound **141**, for use in *in vivo* studies (Figure 6.1). Application of **141**, to *E. coli* cells expressing the *sdkef* gene in a Kirby-Bauer disc diffusion assay has demonstrated, for the first time that small molecules targeting *Sd*Kef can elicit inhibition of growth / cell death. Furthermore, an analogue of **141**, compound **161**, has shown that other more clinically relevant homologues of Kef can also be targeted, including *Ec*KefF-C and *Ab*Kef. Kef thus shows promise as a novel therapeutic target for the development of antibacterial agents.

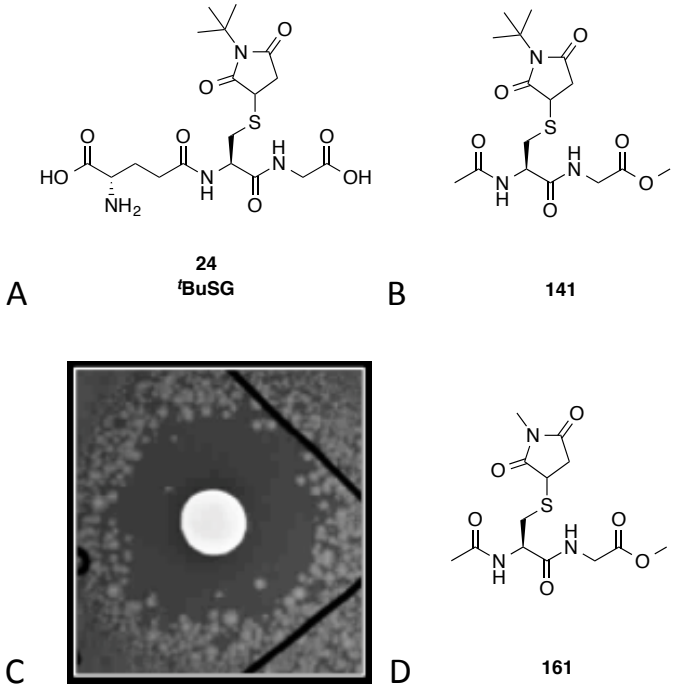


Figure 6.1 (A) Structure of **^{t}BuSG**, **24**; (B) Structure of compound **141**; (C) Kirby-Bauer disc diffusion assay assessing the ability of compound **141** (dosed at 400 µg) to elicit inhibition of growth / cell death of *E. coli* MJF335 cells containing or lacking the plasmid pTrcSdKefH$_6$ for full-length *Sd*Kef; (D) Structure of compound **161**.

6.2 Future work

A proof of concept for Kef as a target for the development of novel antibacterial agents has been obtained using a recombinant system that has its endogenous K^+ uptake (*kdpABC*), K^+ efflux (*kefB* and *kefC*) and glutathione synthesis (*gshA*) genes knocked out. Although the recombinant system acts as a useful model for prescreening homologues of Kef from pathogenic bacteria, as was demonstrated with *Ab*Kef, it is important that Kef is validated in wild-type strains that have these genes intact. Furthermore, the recombinant system has only one Kef gene present that is being expressed at high levels. Wild-type bacteria are likely to have more than one Kef system, for example KefB and KefC in *E. coli*, which are being expressed at lower basal levels.

To extend the proof of concept to wild-type strains it is likely that each of the challenges posed by the presence of these genes and the lower expression levels of Kef will need to be overcome. Future work to address the presence of these genes in wild type strains is as follows:

The K^+ uptake systems and high glutathione background could be overcome by identifying more potent ligands of Kef that are able to outcompete the low affinity glutathione and cause more permanent activation / inhibition of Kef. More potent ligands could be obtained through modifications to the dipeptide **101**, or by performing a high throughput screen using the assays developed in Chapter 4 to identify ligands of Kef with novel binding modes.

Modifications of the dipeptide **101** could commence at four key locations (Figure 6.2). The first is the Gly-carboxylate, which could be replaced by carboxylic acid bioisosteres, for example. A tetrazole (compound **147**) has already been implemented at this position with minimal impact to the affinity, but there are many other bioisosteres that could be tried. It is worth noting that the tetrazole analogue, **147**, was not membrane permeant and so if a potent bioisostere is identified

it may need to be modified to improve membrane permeability, e.g. by forming a prodrug. The second site is the conjugate at the Cys-thiol position, as different homologues of Kef have been found to be sensitive towards the nature of the conjugate. More *in vivo* studies therefore need to be performed to investigate the impact that this group has on both the potency and ability to activate various homologues of Kef. The third location is the acetylated cysteinylamine. It is possible that building into this region of the glutathione binding-site could help with gaining affinity. Finally it may be possible to make changes at the Cys-Gly amide bond by replacing it with peptide mimetics such as a carbon-carbon double bond. This may help to progress towards more non-peptidic ligands that are more metabolically stable.

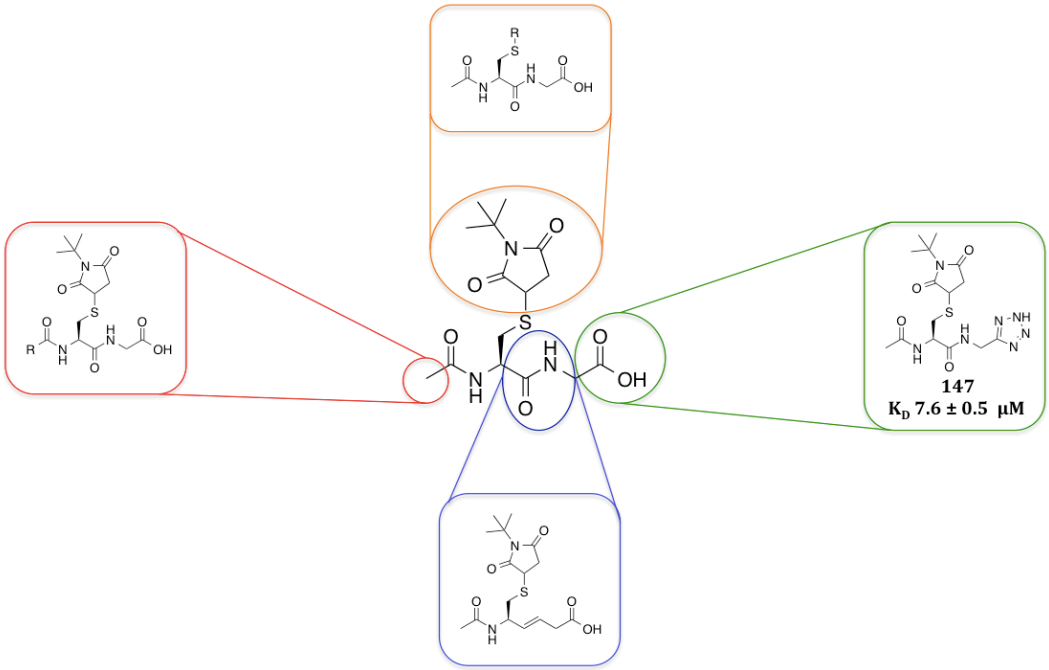


Figure 6.2 Examples of modifications that could be made to compound **101** to help improve its affinity and develop non-peptidic ligands.

Future work to address the lower expression levels in wild type strains could involve developing one or more ligands that are able to target each of the Kef systems present in a given wild-type strain. The extent of activation elicited by a given glutathione S-conjugate is known to be subject to the following degree of subtlety: when the conjugation adduct is large and hydrophobic, all

Kef transporters are activated; when the adduct is small and hydrophilic, only a subset of Kef are activated. For example, *S*-lactoylglutathione, **16**, the glutathione S-conjugate that arises from methylglyoxal, **13**, detoxification is a small adduct that strongly activates *Ec*KefB, but only weakly activates *Ec*KefC. Different analogues of compound **141** may therefore need to be developed that can be dosed together to simultaneously activate multiple Kef systems in a given wild-type strain. The intention of this approach is that hopefully the number of Kef activated will increase to similar levels present in the recombinant model system.

Chapter 7:

Experimental

7. Experimental

7.1 Chapter 2 experimental

7.1.1 Expression and purification of *Sd*KefCTD

7.1.1.1 Strains and plasmids

The Miller Group, University of Aberdeen, provided all strains and plasmids (Tables 7.1 & 7.2).

Table 7.1 *E. coli* strains used.

***E. coli* strain**	**Genotype**	**Reference**
Frag1	F$^-$, *lacZ82*(Am), *λ*$^-$, *rha-4*, *thiE1*, *gal-33*	9
Frag5	Frag1, *Δ(kdpC-kdpA)18*	10
MJF335	Frag5, *lacI*, *kefB*, *kefC::*Tn*10*, *gshA::*Tn*10*(Kan)	1
MJF276	F- Δ*kdpABC5 thi rha lacI lacZ trkD1kefB157 kefC::Tn10*	2
MJF373	MJF276 Δ*kefFC::kan,* Δ*crp kefB::Tn10*	3

Table 7.2 Plasmids used.

Plasmid	**Description**	**Reference**
pTrcSdKefH$_6$	C-terminal hexa-His tagged (LEH$_6$) full-length *Sdkef* (codon-optimised) expressed from *trc* promoter. GenBank ID of *SdKef*: **ABE53663.1**	3
pTrcSdKefQCTD	C-terminal hexa-His tagged (LEH$_6$) truncated *Sdkef* (i.e. *Sd*KefCTD) expressed from *trc* promoter.	3
pkC11	Full-length *EckefF-C* expressed under own promoter	4
pKefGB	Full-length *EckefG-B* expressed under own promoter	(unpublished)
pTrcAbKefH$_6$	C-terminal hexa-His tagged (LEH$_6$) full-length AbKef (codon-optimised) expressed from trc promoter. GenBank ID of *AbKef*: **KDF03402.1**	(unpublished)

The expression construct used was pTrcSdKefQCTD, which consists of the C-terminal KTN domain of *Sd*Kef, as well as, the Q-linker (GHELEVDIEPK) encoding for a total of 237 amino acids.[3] The Q-linker contains the regulatory loop from the membrane domain (HELEVDIEP).

The sequence of pTrcSdKefQCTD was verified by submitting it to Source BioScience (http://www.sourcebioscience.com/). pTrcSdKefQCTD was sequenced using the PBADR (5' – GATTTAATCTGTATCAGG – 3') primer. Below is the sequencing results obtained from Source BioScience aligned with the pTrcSdKefQCTD plasmid reference sequence, using CLC Main Workbench 7.6.2. (Figure 7.1).

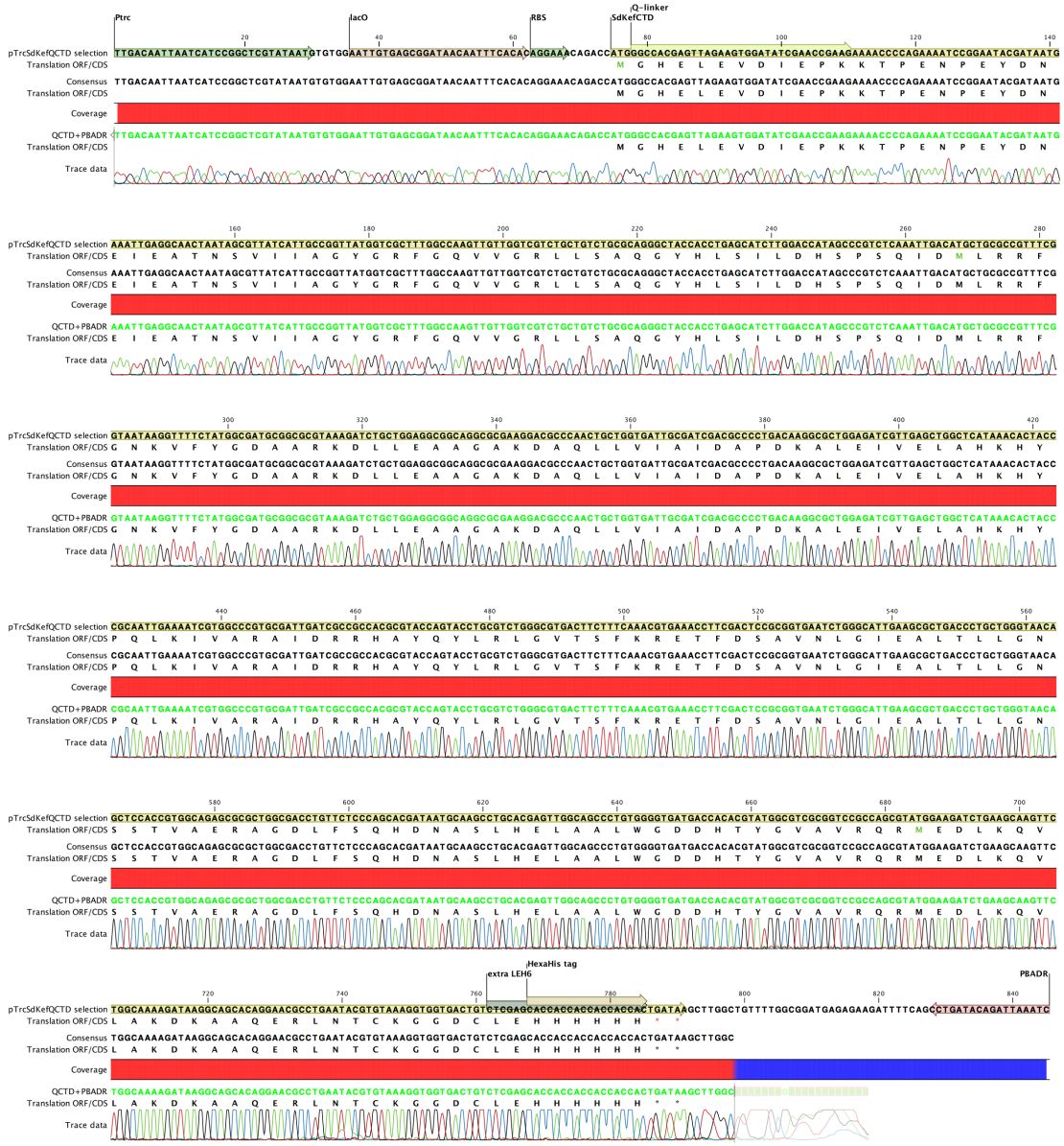


Figure 7.1 The sequencing results obtained from Source BioScience aligned with the pTrcSdKefQCTD plasmid reference sequence, using CLC Main Workbench 7.6.2.

7.1.1.2 Protein expression and purification

Ampicillin was used at a final concentration of 100 µg/mL unless stated otherwise.

7.1.2.1 Transformation - pTrcSdKefQCTD was transformed for expression into the *E. coli* strain MJF373,[3] which lacks the gene for the catabolite activator protein (CAP), a common contaminant in immobilised metal ion affinity chromatography.

The MJF373 competent cells were thawed on ice for 20 min. After this time, 2 µL of pTrcSdKefQCTD was added to 20 µL of the MJF373 competent cells in a pre-chilled 1.5 mL microcentrifuge tube. The tube was left on ice for 30 min before being heat shocked at 42 °C in a water bath for 30 s. The tube was returned to the ice for 5 min, 300 µL of 2× TY media (16 g L^{-1} Oxoid™ Tryptone, 10 g L^{-1} Oxoid™ Yeast Extract, 5 g L^{-1} NaCl) was added gently to each tube, which was then incubated at 37 °C for 1 h. After this time, 50 µL of the transformation mixture was plated onto agar plates containing ampicillin and 0.2% *w/v* glucose and incubated overnight at 37 °C.

Starter culture - A single colony was picked from the agar plate and used to inoculate 100 mL of 2× TY media containing ampicillin and 0.2% *w/v* glucose. The resulting starter culture was incubated overnight at 37 °C under foil.

Protein expression - 6 × 6 mL of starter culture was used to inoculate 6 × 600 mL of 2× TY media containing ampicillin and 0.2% *w/v* glucose, which were grown at 37 °C until an OD_{600} of 0.8 was reached. After cooling to 30 °C, expression was induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG; final concentration of 0.8 mM) for 4 h. Cells were harvested by centrifugation and the resulting bacterial pellet was frozen in a sealable plastic bag at −80 °C until further processing.

7.1.2.2 Protein purification - The buffers used in the purification of *Sd*KefCTD are shown in Table 7.3.

Table 7.3 Buffers used in the purification of *Sd*KefCTD

Reagents	Nickel affinity column buffers (pH 7.6)				
	Charge	**Binding**	**Wash**	**Elution**	**Strip**
HEPES	–	50 mM	50 mM	50 mM	50 mM
NaCl	–	0.5 M	0.5 M	0.5 M	0.5 M
Imidazole	–	5 mM	30 mM	500 mM	–
$NiSO_4$	100 mM	–	–	–	–
EDTA	–	–	–	–	100 mM

Cell pellets were dissolved in binding buffer (5× mass of cell pellet *w/v*), and lysed by sonication (Ultrasonic Processor, Sonics & Materials Inc.), set at 40%, applying ten 30-second bursts interrupted by 40-second pauses. The lysate was clarified by centrifugation and the supernatant decanted off, filtered through a 0.45 μm filter before being applied to a nickel affinity purification column. The nickel affinity column contained NOVAGEN® His·Bind® Resin charged with 20 mL of 100 mM $NiSO_4$ under gravity. The nickel affinity column was equilibrated under gravity with 50 mL of binding buffer prior to loading the cell lysate. The column was then washed under gravity with approximately 100 mL wash buffer. When all residual products of bacterial fermentation were washed, the protein of interest was batch eluted with 30 mL elution buffer under gravity. Eluted protein was collected in 1 mL fractions. The peak fractions as determined by UV/Vis spectroscopy and SDS-PAGE were subjected to buffer exchange using a PD-10 column (GE Healthcare) into buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4. Protein concentration was estimated by measuring the absorbance at 280 nm using a Nanodrop®

ND-1000 spectrophotometer (Nanodrop® Technologies Inc.) with the 'Protein A280' program module, according to the manufacturer's instructions, and was found to be 17.8 mg/mL.

The purity of *Sd*KefCTD was analysed by SDS-PAGE. For denaturing SDS-PAGE analysis, samples were prepared by mixing with 2× Laemmli Sample Buffer and denatured at 100 °C for 3 min. Gels were run on a Mini-PROTEAN 3 apparatus (Bio-Rad Laboratories, Inc.) at a constant potential of 200 V in the Laemmli Buffer System, according to manufacturer's specifications (Bio-Rad Laboratories, Inc.). Following electrophoresis, the gels were stained for 10 min with InstantBlue™ Coomasie® staining solution [0.5% (*w/v*) Brilliant Blue G, 50% (*v/v*) methanol, 10% (*v/v*) glacial acetic acid, in Milli-Q water (Merck Millipore Corp.)]. The stained gels were then gradually destained with Destaining Solution [40% (*v/v*) methanol, 10% (*v/v*) glacial acetic acid, in Milli-Q water (Merck Millipore Corp.)] until protein bands on the gels become visible. All chemicals were purchased from Sigma-Aldrich Co., except 2× Laemmli Sample Buffer (Bio-Rad Laboratories, Inc.), β-mercaptoethanol (Bio-Rad Laboratories, Inc), glacial acetic acid (Thermo Fisher Scientific Inc.), and glycine (Thermo Fisher Scientific Inc).

7.1.2 Differential scanning fluorimetry - Differential scanning fluorimetry was performed using a Stratagene Mx3005P qPCR (Experimental filter set, excitation 492 nm, emission 568 nm). The initial temperature was set to 25 °C and held for 5 min. The temperature was then increased in increments of 1 °C for 55 cycles (held for 1.5 min per cycle). Stock solutions of the ligands under examination were prepared to a concentration of 100 mM in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4 (with the exception of adenosine, which was prepared in **DMSO**). The 100 mM stock solutions were then diluted to a concentration of 10 mM in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4. A protein master mix was prepared containing *Sd*KefCTD (13.3 µM) and SYPRO® Orange (2.2×, Invitrogen) in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4. 96 well plates (Axygen) were prepared using the

protein master mix (22.5 µL per well; 12 µM final concentration of *Sd*KefCTD and 2× final concentration of SYPRO® Orange) and the appropriate ligand (2.5 µL per well; 1 mM final concentration). The plate was centrifuged at 1000 rpm for 3 min before being run. Controls were performed with dye alone, ligand and dye, and *Sd*KefCTD alone. The T_m (melting temperature) was identified by fitting to the Boltzmann equation using Prism 6.[5] The change in unfolding temperature (ΔT_m) was calculated as the T_m of *Sd*KefCTD + ligand minus the T_m of *Sd*KefCTD + 2.5 µL buffer, or *Sd*KefCTD + 2.5 µL **DMSO** in the case of adenosine. A Student's *t*-test was performed to determine which ΔT_m values obtained were statistically significant.

For the DSF studies performed on the *Sd*KefCTD mutant R416E, the same experimental protocol was employed, with the exception of the following: Protein master mixes were prepared containing wild-type *Sd*KefCTD or the *Sd*KefCTD mutant R416E (8.7 µM) and Sypro Orange (2.2×, Invitrogen) in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4. Ninety-six well plates (Axygen) were prepared using the protein master mix (22.5 µL per well; 7.8 µM final concentration of protein and 2× final concentration of dye) and **AMP** (2.5 µL per well; 1 mM final concentration).

7.1.3 Establishing the identity of the nucleotide natively bound to *Sd*KefCTD using ^{1}H CPMG NMR - All ^{1}H CPMG NMR data were obtained in collaboration with Amjad Khan in the Claridge Group, University of Oxford. All the NMR experiments were carried out at a ^{1}H frequency of 700 MHz using a Bruker Avance III spectrometer equipped with a TCI inverse cryoprobe. The spectra were recorded in Wilmad® NMR tubes (5 mm diameter, Bruker). The NMR experimental parameters used were: number of transients, 16; relaxation delay, 2 s; total echo time, 96 ms (in CPMG edited ^{1}H experiments). The PROJECT-CPMG sequence ($90°_x - [\tau - 180°_y - \tau - 90°_y - \tau - 180°_y - \tau]_n$ – acq) was applied as described by Aguilar *et al.*[6] Water suppression was achieved by

presaturation and the residual water peak was further attenuated by removing low frequency components from the FID (Bruker qfil routine with window of 0.1000 ppm). The pulse tip angle calibration was done for all the samples. All the experiments were conducted at 278 K.

*Sd*KefCTD was prepared to a final concentration of 330 µM in deuterated buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4 as follows: The purified *Sd*KefCTD solution was exchanged with deuterated buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4 (which was prepared in D_2O) by using PD MiniTrap G-25 columns (GH Healthcare). The exchange was done by the spin protocol, according to manufacturer's instruction. *Sd*KefCTD was denatured in an Eppendorf® tube on a heat block at 80 °C for 3 h. After this time, the resulting protein precipitates were removed by centrifugation and the supernatant was directly used for ^{1}H CPMG NMR studies. The experiment was repeated without removing the samples from NMR tubes to confirm **AMP** was not observed from contamination.

The reference spectra of each of the ligands **AMP**, **ADP** and **NADH** were also were run at a final concentration of 330 µM in the deuterated buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4 (which was prepared in D_2O). The ligands were also heat treated in an Eppendorf® tube on a heat block at 80 °C for 3 h.

For the ^{1}H CPMG NMR studies performed on the *Sd*KefCTD mutant R416E, the same experimental protocol was employed, with the exception of the following: The spectra were recorded in MATCH NMR tubes (3 mm diameter, Bruker). The NMR experimental parameters used were: number of transients, 11780; relaxation delay, 2 s; acquisition time, 11 h and 35 min; total echo time, 40 ms (in CPMG edited ^{1}H experiments). Wild-type and mutant R416E *Sd*KefCTD was prepared to a final concentration of 5.2 µM in deuterated buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4 (which was prepared in D_2O).

7.1.4 Verifying the identity of the nucleotide using analytical HPLC - The protocol for performing the analytical HPLC studies on the natively bound nucleotide of *Sd*KefCTD was modified from the one described by Chen *et al.*[7] Analytical HPLC was performed on a PerkinElmer Flexar system with a Binary LC Pump and UV/VIS LC Detector. A Dionex Acclaim® 120 column (C18, 5 µm, 120 Å, 4.6 × 150 mm) was used for all experiments. The method employed for the HPLC analysis of the nucleotides is as follows:

Step	Step Time / min	Flow (mL/min)	%A	%B
0	0.1	1.00	100.0	0.0
1	1.0	1.00	100.0	0.0
2	10.0	1.00	0.0	100.0
3	3.0	1.00	0.0	100.0
4	1.0	1.00	100.0	0.0
5	5.0	1.00	100.0	0.0

Moblie Phases: A = 99.9:0.1 H_2O : Formic acid; B = 99.9:0.1. MeCN : Formic acid.

A 20 µL aliquot of a solution of **AMP** (adenosine 5'-monophosphate disodium salt bought from Sigma Aldrich) at a final concentration of 50 µM was injected into the analytical HPLC machine to provide a reference spectrum of know concentration. This procedure was repeated for **ADP** and **NADH**.

A 20 µL aliquot of the supernatant of denatured *Sd*KefCTD (preparation follows) at a final concentration of 50 µM was injected into the analytical HPLC machine to provide a spectrum of the ligand natively bound to *Sd*KefCTD.

*Sd*KefCTD was purified, concentrated, and stored in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4 at a concentration of 320.8 µM, as determined by the NanoDrop 1000 Spectrophotometer. A 10 µL of solution containing 3.1 µL of purified *Sd*KefCTD and 6.9 µL of the above buffer was denatured at 95 °C for 5 min, and then subjected to centrifugation at 13,000 rpm (Mikro 20 centrifuge; Hettich) for 10 min to form a pellet of the denatured protein precipitate. The resulting pellet was suspended in 10 µL of the above buffer, and centrifuged at 13,000 rpm for 10 min to further remove weakly associated nucleotide from the denatured protein sample. The supernatants from the two centrifugation steps were combined to obtain the 20 µL sample of denatured *Sd*KefCTD that was injected into the analytical HPLC machine.

7.1.5 Nanoelectrospray mass spectrometry - Shane Chandler in the Benesch Group, University of Oxford, performed the nanoelectrospray mass spectrometry experiments. The experimental procedure provided by Shane Chandler is below:

Nanolectrospray (nESI) MS experiments were carried out on a QToF mass spectrometer (Waters Corp., Wilmslow, UK), according to previously described protocols,[8] and with instrument settings optimised for the transmission of intact noncovalent protein complexes. The sample was buffer exchanged into 200 mM ammonium acetate, and sprayed at a final concentration of 7.5 µM with respect to the dimer. Experiments were conducted in positive polarity with the following instrument settings: capillary voltage 1.4 kV, sample cone 40 V, extraction cone 25 V, backing pressure 3.5 mbar and collision cell pressure 1.75 MPa.

7.2 Chapter 3 experimental

7.2.1 Competition fluorescence assay - All steady state fluorescence measurements were performed using a Perkin Elmer LS-50B Spectrometer. A protein master mix was prepared containing *Sd*KefCTD (6 μM) and **DNGSH** (5 μM) in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4. Stock solutions of the ligands under examination were prepared to a concentration of 100 mM in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4, or **DMSO**. When the competing ligands were tested at a final concentration of 5 mM, 142.5 μL of the protein master mix and 7.5 μL of the 100 mM competing ligand stock were added to the cuvette. When the competing ligands were tested at a final concentration of 1 mM, 148.5 μL of the protein master mix and 1.5 μL of the 100 mM competing ligand stock were added to the cuvette. Samples were excited with light of 340 nm and the emission spectra measured from 375-670 nm (the step was 1 nm, the dwell time 0.3 s, 3 repeats, the excitation slit was 1 nm, the emission slit 2 nm and the temperature was held at 20 °C). A decrease in the fluorescence intensity after the addition of the desired competing ligand, at a final concentration of 1 mM or 5 mM, was interpreted as an indication of binding. The data are represented as ratios of the fluorescence intensities before (F_B) and after the addition of the competing ligand (F_L) at 525.5 nm (F_B/F_L-1).

For the negative control of performing the experiment in the absence of *Sd*KefCTD, a solution of **DNGSH** (5 μM) in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4 was prepared. When the competing ligands were tested at a final concentration of 5 mM, 142.5 μL of the **DNGSH** (5 μM) solution, and 7.5 μL of the 100 mM competing ligand stock were added to the cuvette. When the competing ligands were tested at a final concentration of 1 mM, 148.5 μL of the **DNGSH** (5 μM) solution, and 1.5 μL of the 100 mM competing ligand stock were added to the cuvette.

7.2.2 Saturation transfer difference NMR – Dr. Jess Healy and Dr. Sonia Diab in the Conway Group performed the saturation transfer difference NMR experiments. The experimental procedure provided by Dr. Sonia Diab is below:

A 100 mM stock solution of the compound being tested was prepared in **DMSO**. To each NMR tube was added the appropriate volume of *Sd*KefCTD, compound stock solution and deuterated 50 mM sodium phosphate buffer pH 7.4, to ensure that there was a 100 fold excess in the concentration of the compound, compared to the concentration of the protein (the protein had a final concentration in the range 15-30 μM).

^{1}H NMR saturation transfer difference experiments were performed on a Bruker AVII 500 (500 MHz) spectrometer equipped with a ^{1}H {^{19}F/^{13}C} triple resonance probe regulated at 298 K. Protein saturation was performed by irradiating the protein with a train of Gaussian-shaped pulses (50 ms). Water suppression was performed using an excitation sculpting method employing a water-selective sinc profile 180° pulse. Initial experiments to determine whether binding could be detected were performed with 3 s total protein saturation time and data were acquired for 160 transients per experiment.

7.2.3 *In silico* screening - Individual *in silico* protein-ligand docking was performed locally on an Intel® Core™2 Duo CPU E6550 @ 2.33GHz × 2 running Ubuntu 12.04 LTS. High-throughput *in silico* protein-ligand docking was performed on a 96-core cluster with the following hardware specifications: number of cores per CPU: 4 physical cores / 8 effective threads; clock speed: 2.27 GHz; RAM on each node: 16 GB; traditional rotating magnetic drives, 200 GB; Queuing system on the cluster: Sun Grid Engine; OS that the cluster nodes were running: ROCKS Linux.

7.2.3.1 Preparation of the receptors for AutoDock 4 and Vina - The receptors were initially prepared by adding polar hydrogen atoms at pH 7.4 using Protonate3D in Molecular Operating Environment (MOE) 2011.10 molecular modeling software. In the case of the validation dockings, MOE was used to assign Gasteiger charges to the receptor, but in the case of all other dockings Gasteiger charges were assigned using AutoDock's graphical user interface AutoDockTools (ADT; version 1.5.4). Following addition of Gasteiger charges, ADT built a "united atom model" of the receptor by merging non-polar hydrogens and adding their partial charges to their parent carbon atoms.

7.2.3.2 Preparation of the ligands for docking using AutoDock 4, Vina and GOLD - The ligands were prepared by adding polar hydrogens at pH 7.4 using Protonate3D in MOE. MOE was then used to assign Gasteiger charges, before performing an energy minimisation using an MMFF94x forcefield with a 0.05 gradient. In the case of GOLD no further processing took place and the ligands were used in mol2 format. In the case of AutoDock 4 and Vina, ADT was used to build a "united atom model" of the ligands.

7.2.3.3 Individual docking with AutoDock 4 (version 4.2) - The number of Genetic Algorithm runs was set to 100, with the maximum number of evaluations (evals) set to 2,500,000. All other parameters were set at the default values provided by ADT. Conformational searching was performed using the Lamarckian Genetic Algorithm. 3D desolvation, electrostatic, and atomic affinity grid maps of the glutathione-binding site for all of the atom types in the ligand being screened were generated using AutoGrid 4. The dimensions of the grid box used to generate these maps were 18.75 Å × 18.75 Å × 18.75 Å, with a grid point spacing of 0.375 Å and centre point: x = −25.961, y = 15.04, z = −4.422. The receptor was set as rigid for all dockings.

7.2.3.4 Individual docking with AutoDock Vina (version 1.1.2) - The search space was defined as a cubic box with dimensions 18.75 Å × 18.75 Å × 18.75 Å and centre x = −25.961, y = 15.04, z = −4.422. The exhaustiveness and number of modes were both set to 20 and all other parameters were kept at their default values.

7.2.3.5 Individual docking with GOLD (version 5.1) - The binding site was defined as having a radius of 11.6 Å with the origin set to (−25.961, 15.04, −4.422). The scoring function implemented was ChemPLP, with each molecule being docked a total of 10 times. The runs were configured with all the population and genetic operators set to auto (autoscale = 1, 100% search efficiency), allowing GOLD to determine the optimal number of operations for each ligand. All other parameters were set to their default settings, with the exception of the following: floodfill_center = file and initial_virtual_pt_match_max = 3.

7.2.3.6 Construction the virtual library of compound 74 and commercially available carboxylic acids - InhibOx's proprietary program ReactiOx was used to construct a virtual library of compound **74** coupled to 1,017,567 million commercially available carboxylic acids extracted from InhibOx's in-house database Scopius-C-Space 6. The first step involved specifying the reaction transformation and describing the corresponding atoms in the reactants and products. ReactiOx then used the reaction transformation and reagent files as an input to output product files. InhibOx's in-house program Conform was then used to generate conformers of the products by converting the products file from 2D smiles format to 3D sdf format.

7.2.3.7 High-throughput virtual screen with GOLD - The binding site was defined as having a radius of 11.6 Å with the origin set to (−25.961, 15.04, −4.422). 330 configuration files, containing between them 661 diversity subsets of InhibOx's Scopius-CSpace 6, were submitted as an array job of GOLD dockings on a 96-core cluster. The diversity subsets contained 1000 molecules in

mol2 format (or sdf format in the case of docking the virtual library constructed from compound **74** and the commercially available carboxylic acids), which were each docked into Hm1*Sd*KefCTD a total of 10 times. The scoring function implemented was ChemPLP and the runs were configured with all the population and genetic operators set to auto (autoscale = 0.3, 30% search efficiency). All other parameters were set to their default values with the exception of the following: rotate_carboxylic_oh = flip; initial_virtual_pt_match_max = 3. In the case of docking the virtual library constructed from compound **74** and the commercially available carboxylic acids, *protein_h_bond* constraints were imposed on the atoms in residues R423, R507, R523 so that GOLD favoured solutions involving hydrogen bonds between the docked ligand and these residues. The *protein_h_bond* constraints consisted of: the <*constraint weight*> or penalty for not forming the specified hydrogen bonds, which was set to 10.0000, and the <*min. geometry weight*> that determines the quality of a hydrogen bonding interaction, which was kept at its default value of 0.005. The "best-docked" solution was kept for each ligand and the top 500 ligands outputted by each configuration file were retained based on their fitness scores. At the end of docking all the configuration files, the results were filtered to give the top 100 solutions based on GOLD ChemPLP fitness score.

7.2.3.8 High-throughput virtual screen with Vina - The search space was defined as a cubic box with dimensions 18.75 Å × 18.75 Å × 18.75 Å and centre x = −25.961, y = 15.04, z = −4.422. The exhaustiveness was set to 2 and all other parameters were kept at their default values. Ligands were submitted as a Vina array job to a 96-core cluster. At the end of the screen the results were filtered to identify ligands with an affinity < −38.9 kJ/mol.

7.2.3.9 Water prediction and classification with WaterDock - WaterDock used AutoDock Vina to dock a single water molecule into Hm1*Sd*KefCTD's binding site 3 times to generate a total of 60 potential water sites. The search space for this was defined as a cubic box with dimensions 18.75 Å × 18.75 Å × 18.75 Å and centre $x = -25.961$, $y = 15.04$, $z = -4.422$. The exhaustiveness was set to 20, the energy_range to 100, the number of modes to 20 and with all other parameters kept at their default values. Once the water molecules had been docked, WaterDock then filtered and clustered all the binding modes using the statistical package R. WaterClassifier was then run to determine the probabilities of predicted waters being displaced. This was achieved by supplying WaterClassifier with the water positions in a pdb file format and the *apo* protein structure in a pdbqt file format. The output of the above process gave the water displacement probabilities in the order that the water molecules appeared in.

7.3 Chapter 4 experimental

7.3.1 Surface Plasmon Resonance (SPR) experimental - SPR experiments were performed on a Biacore™ T200, under the supervision of Dr. David Staunton from the Molecular Biophysics Suite, Department of Biochemistry, University of Oxford. HBS-EP buffer refers to pH 7.4 buffer containing 10 mM HEPES, 150 mM NaCl, 0.05% Tween in Milli-Q® water. All buffer was filtered and degassed before use.

The analysis cycle used for SPR with a Ni Affinity chip is as follows:

1) $NiCl_2$ (0.5 mM in HBS-EP buffer) was passed over the Series S Biacore Sensor Chip NTA at a flow rate of 10 µL/min for 60 s.
2) EDTA (3 mM in HBS-EP buffer) was passed over the Series S Biacore Sensor Chip NTA at a flow rate of 10 µL/min.
3) *Sd*KefCTD (0.1 µM stock in pH 7.4 buffer containing 50 mM Na-phosphate, 150 mM NaCl) was passed over the nickel-activated Series S Biacore Sensor Chip NTA at a flow rate of 10 µL/min for 60 s.
4) **^{t}BuSG** (200-0.05 µM in HBS-EP buffer) was passed over the Series S Biacore Sensor Chip NTA with *Sd*KefCTD captured on at a flow rate of 30 µL/min for 60 s.
5) EDTA (350 mM in HBS-EP buffer) was passed over the Series S Biacore Sensor Chip NTA at a flow rate of 30 µL/min for 60 s.

SPR with biotinylated glutathione S-conjugates:

1) Equilibrate CM5 chip in HBS-EP buffer.
2) Inject 2 × 8 µL of 50 mM NaOH over all four flow cells at a flow rate of 5 µL/min.
3) Inject NHS/EDC over the surface for 7 min.
4) Inject 100 µg/mL streptavidin (Sigma S-4762 streptavidin 1 g, made up to 0.5 mg/mL in 10 mM sodium acetate pH 4.0) in 10 mM sodium acetate pH 4.5 over all four flow cells for 14 min at a flow rate of 5 µL/min.
5) Block the surfaces with a 7 min injection of 1 M ethanolamine hydrochloride pH 8.5.
6) Wash with 3 × 200 µL of 100 mM glycine-HCl pH 2.0 over 30 s at flow rate of 30 µL/min.
7) Inject the desired biotinylated probe in HBS-EPS buffer at a final concentration of 100 nM repeatedly until the desired increase in response is attained at flow rate of 30 µL/min.
8) Wash with 2 × 200 µL of 100 mM glycine-HCl pH 2.0 over 30 s at flow rate of 30 µL/min.
9) Run method to titrate *Sd*KefCTD across the flow cells using a serial dilution of factor 0.5 from 100-0.10 µM of *Sd*KefCTD for 120 s at 30 µL/min.

7.3.2 AlphaScreen® experimental - AlphaScreen® studies were performed by a Synergy™ 2 MultiMode Microplate Reader using ProxiPlate™-384 Plus 6008280, white 384-shallow well Microplates. Each well was loaded with samples to give a final volume of 20 µL. The assay buffer (10 mg BSA, 5 mg CHAPS, 60 mg HEPES, 58 mg NaCl, pH 7.8 in 10 mL of Milli-Q® water; filter sterilised through a 0.22 µm filter) was prepared on the day of testing and stored at 4 °C. The experimental procedure for setting up the plates for the assay is as follows:

1) Prepare protein-probe master mixes in assay buffer containing the following: *Sd*KefCTD protein (500 nM final concentration) and the biotinylated probes **96** or **97** (500 nM final concentration); *Sd*KefCTD protein (100 nM final concentration) and the biotinylated probes **96** or **97** (100 nM final concentration).
2) Pipette 5 µL of glutathione (**GSH**) or **^{t}BuSG** (250 µM stock in assay buffer) into the appropriate wells (62.5 µM final concentrations of **GSH** / **^{t}BuSG**) or 5 µL of assay buffer in the case of wells containing Kef + buffer or Kef + biotinylates probes **96** / **97**.
3) Pipette 8 µL of each of the master mixes prepared in step 1 into the relevant wells. (Final concentrations of *Sd*KefCTD and the biotinylated peptides **96** and **97** of 200 nM from the 500 nM master mixes and 40 nM from the 100 nM master mixes)
4) Seal the plate and shake it at 300 rpm for 30 min at RT using an orbital shaker.
5) Prepare an AlphaScreen® acceptor-donor bead mix containing 6 µL Ni chelate acceptor beads (stock 5 mg/mL), 6 µL streptavidin donor beads (stock 5 mg/mL) and 588 µL of assay buffer (final concentration 50 µg/mL of the acceptor and donor beads).
6) Pipette 7 µL of the AlphaScreen® bead mix containing 50 µg/mL of acceptor and donor beads to each well in a dark room (final concentration 17.5 µg/mL of the acceptor and donor beads)
7) Seal the plate with a sheet of foil and shake it at 300 rpm for 1 h at RT using an orbital shaker.

7.3.3 "Pull-down" experimental

Dynabeads® MyOne™ Streptavidin C1 magnetic beads contain 10 mg/mL of magnetic beads in phosphate buffered saline pH 7.4, with 0.01% Tween®-20 and 0.09% sodium azide. SDS-PAGE was performed using Mini-PROTEAN 3 apparatus (Bio-Rad Laboratories, Inc.). The SDS-PAGE employed the Laemmli Buffer System, according to manufacturer's specifications

(Bio-Rad Laboratories, Inc.). All samples were prepared containing Dynabeads® and each of the following: (A) *Sd*KefCTD + biotinylated probe **96**/**97** + (B) *Sd*KefCTD; (C) *Sd*KefCTD + biotinylated probe **96**/**97** + 10 eq. **^{t}BuSG**; (D) *Sd*KefCTD + biotinylated maleimides used to make **96**/**97**. Samples were made in pH 7.4 buffer containing 50 mM Na-phosphate, 150 mM NaCl. The procedure for the pull-down study is as follows:

1) Add 5 µL of Dynabeads® MyOne™ Streptavidin C1 magnetic beads to each sample (final concentration 0.9 mg/mL).
2) Add 39 µL of *Sd*KefCTD (71.6 µM stock in buffer) to each sample (final concentration 50.8 µM).
3) Add 1 µL of the biotinylated probe **96**/**97** (100 mM stock in **DMSO**) to samples A and C (final concentration 1.8 mM). 1 µL of the biotinylated maleimides used to make **96**/**97** (100 mM stock in **DMSO**) to sample D (final concentration 1.8 mM).
4) Add 10 µL of buffer to samples A and D; 11 µL of buffer to sample B; 10 µL of **^{t}BuSG** (100 mM stock solutions in buffer) to sample C (final concentration 18 mM).
5) Incubate the samples at 0 °C for 2 h.
6) Place the samples next to a magnet for 3 min.
7) Carefully remove the supernatant, without disturbing the beads and keep the supernatant so that it can be loaded onto the SDS-PAGE.
8) Wash the beads by suspending them in 50 µL of buffer.
9) Place the samples next to a magnet for 3 min.
10) Carefully remove the buffer, without disturbing the beads.
11) Repeat steps 5-9 twice.
12) Heat the beads at 100 °C for 5 min in 20 µL of 2× Laemmli Sample Buffer.
13) Place the samples next to a magnet for 3 min and remove the Laemmli Sample Buffer so that it can be loaded onto the SDS-PAGE.

14) Mix the supernatant collected in step 6 in 2× Laemmli Sample Buffer and heat at 100 °C for 5 min.
15) Load denatured samples from steps 12 and 13 into the SDS-PAGE gel and run the gels at a constant potential of 200 V.
16) Stain the SDS-PAGE gels in Coomassie Staining Solution [0.5% (*w/v*) Brilliant Blue G, 50% (*v/v*) methanol, 10% (*v/v*) glacial acetic acid, in Milli-Q® water for 15 min. The stained gels were destained with Destaining Solution [40% (*v/v*) methanol, 10% (*v/v*) glacial acetic acid, in Milli-Q® water] until protein bands on the gels become clearly visible. The images of the destained SDS-PAGE gels were captured using an image scanner.

7.3.4 ^{1}H CPMG NMR experimental - All ^{1}H CPMG NMR data were obtained in collaboration with Amjad Khan in the Claridge Group, University of Oxford. ^{1}H CPMG NMR experiments were performed using Bruker, MATCH (3 mm diameter) NMR tubes on a Bruker Avance III 700 MHz spectrometer equipped with a TCI inverse cryoprobe at 298 K and were processed and analysed using Bruker Topspin™ 3.2. The PROJECT-CPMG pulse sequence ($90°_x$-[τ-$180°_y$-τ-$90°_y$-τ-$180°_y$- τ]$_n$-acq) as described by Aguilar *et al*.[9] was used to remove the broad resonances of the protein. All spectra were processed with a Lorentzian line broadening of 0.3 Hz. The CPMG edited ^{1}H NMR experiments were recorded with a total filter time of 88 ms. Water suppression was achieved by presaturation. For each sample the pulse tip-angle calibration was carried out using the single pulse nutation method.[10]

Samples of 160 µL volume were made up in pH 7.4 buffer containing 50 mM Na-phosphate, 150 mM NaCl in D_2O, with a fixed concentration of the ligand (typically 10 µM) and a range of *Sd*KefCTD concentrations (typically 0-30 µM; the concentration was increased until complete saturation of the ligand signal was achieved). The % decrease in signal intensity of the ligand upon the addition of *Sd*KefCTD was calculated using the equation shown in Equation 7.1.

$$\%\ \text{Decrease in signal intensity} = \frac{I_P - I_B}{I_B} \times 100$$

Equation 7.1 The equation used to calculate the % decrease in signal intensity of the ligand upon the addition of *Sd*KefCTD, where I_B is the signal intensity of the ligand in the absence of *Sd*KefCTD and I_P is the signal intensity of the ligand in the presence of *Sd*KefCTD at a given concentration.

The titration data (% decrease in ligand signal intensity plotted against *Sd*KefCTD concentration) was fitted using OriginPro 9.1 (Origin lab, Northampton, MA, USA) and K_D values were extracted using the equation described by Dalvit[3] (Equation 7.2).

$$y = A \times \frac{[P_T] + [L_T] + K_D - \sqrt{([P_T] + [L_T] + K_D)^2 - 4[P_T][L_T]}}{2[L_T]}$$

Equation 7.2 The equation used to extract K_D values from titration data, where y is the % decrease in ligand signal intensity; A is the maximum loss in the signal intensity; $[P_T]$ is the total protein concentration, $[L_T]$ is the total ligand concentration; K_D is the dissociation constant.

7.3.5 ^{19}F NMR assay experimental - Amjad Khan in the Claridge Group, University of Oxford, obtained all of the ^{19}F NMR assay data. The experimental procedure provided by Amjad Khan is below:

^{19}F NMR experiments were performed using Bruker, 5 mm (diameter) NMR tubes on a Bruker Avance III 600 (600 MHz) spectrometer equipped with a 5 mm BB-F/^{1}H Prodigy N_2 cryoprobe operating at 298K. All spectra were recorded with 1312 scans (acquisition time 1 h, 26 min) and analysed by Bruker Topspin™ 3.2. The spectra were processed with a Lorentzian line broadening of 10 Hz. The areas of the peaks were measured using Lorentzian shape deconvolution. Stock solutions of the ligands were prepared in pH 7.4 buffer containing 50 mM Na-phosphate, 150 mM NaCl in D_2O, with the exception of **98**, which was prepared in pH 7.4 buffer containing

50 mM Tris-D_{11} in H_2O. Samples of 500 µL volume were made up in pH 7.4 buffer containing 50 mM Na-phosphate, 150 mM NaCl in D_2O, with varying concentrations of the competing ligand (final concentrations in the range 0.04-30 mM), and fixed concentrations of **98** (final concentration 60 µM) and *Sd*KefCTD (final concentration 35 µM).

7.3.6 Microscale thermophoresis experimental - The microscale thermophoresis study was performed in triplicate on a Monolith NT.115 using standard treated capillaries at a temperature of 25 °C. 3 × 16 samples (30 µL) were prepared containing the fluorescein probe, **95**, at a final concentration of 25 nM and serial dilutions of *Sd*KefCTD with final concentrations in the range 300-0.024 µM. Each sample was loaded into a standard treated capillary and the experiment was run with 80% MST Power and 20% LED Power. K_D values were extracted by fitting the curve according to the law of mass action.

7.3.7 Fluorescence polarisation - All fluorescence polarisation studies were performed on a POLARstar OPTIMA Microplate Fluorometer using 96 well flat bottom polystyrene non sterile solid black 190 µL well volume Corning® Microplates. Each well was loaded with samples to give a final volume of 40 µL. The plate mode (equidistant) settings were as follows: Number of cycles: 3; Cycle time (s): 214; Number of flashes per well: 200. Samples were excited with plane polarised light in a single plane at 485 nm, and emission was simultaneously measured at 520 nm in the parallel and perpendicular planes, using the following optic settings: Excitation: 485; Emission A: 520-P; Emission B: 520-P. The general settings were as follows: Settling time (s): 1; Reading direction: unidirectional, horizontal left to right, top to bottom; target temperature (°C): 25. All fluorescence polarisation studies were performed in pH 7.4 buffer, containing 50 mM Na-phosphate, 150 mM NaCl in Milli-Q® water, which was filter sterilised through a 0.22 µm filter before use. The gain adjustment was performed on a well containing 25 nM of the fluorescein probe, **95**, in buffer for all fluorescence polarisation studies.

Controls were performed for all fluorescence polarisation studies with the fluorescein probe, **95**, alone and the ligand and fluorescein probe, **95**, in the absence of protein.

7.3.8 Fluorescence polarisation equilibrium saturation binding assay - The equilibrium saturation binding assay studies were performed in triplicate. The experimental procedure for setting up the plates for the assay with the fluorescein probe, **95**, at a final concentration of 25 nM is outlined below:

1) Pipette 30 µL of *Sd*KefCTD protein (342.3 µM stock in buffer) into wells: A1; B1; C1 (final *Sd*KefCTD protein concentration 257 µM).
2) Pipette 23.4 µL of *Sd*KefCTD protein (342.3 µM stock in buffer) + 6.6 µL of buffer into wells: A2; B2; C2 (final *Sd*KefCTD protein concentration 200 µM).
3) Pipette 21.0 µL of *Sd*KefCTD protein (342.3 µM stock in buffer) + 9.0 µL of buffer into wells: A3; B3; C3 (final *Sd*KefCTD protein concentration 180 µM).
4) Pipette 30 µL of buffer into wells: A4-12; B4-12; C4-12.
5) Pipette 30 µL of *Sd*KefCTD protein (342.3 µM stock in buffer) into wells: A4; B4; C4. Pipette up and down in each of these wells to mix the *Sd*KefCTD protein with the buffer thoroughly.
6) Transfer 30 µL of the diluted *Sd*KefCTD protein from A4, B4 and C4 to the wells A5, B5 and C5 respectively. Pipette up and down in each of these wells to mix thoroughly.
7) Repeat step 6 for each of the subsequent wells from A5-12, B5-12, C5-12, such that a serial dilution of the *Sd*KefCTD protein of factor 0.5 has been performed across the wells: A4-12; B4-12; C4-12 (final *Sd*KefCTD protein concentrations of 128-0.25 µM).
8) Pipette 10 µL of the fluorescein probe (100 nM stock in buffer) into each of the wells: A1-12; B1-12; C1-12 (final probe concentration of 25 nM).
9) Place a black lid on the plate and shake the plate at 300 rpm for 20 min at RT using an orbital shaker.

10) Remove the black lid, place the plate in the POLARstar OPTIMA Microplate Fluorometer and start the measurement.

7.3.9 Fluorescence polarisation single shot displacement assay - The single shot displacement assays were performed in duplicate or triplicate. The experimental procedure for setting up the plates for the assay is outlined below:

1) Prepare a protein-probe master mix containing *Sd*KefCTD protein (240 µM final concentration) and the fluorescein probe (33.3 nM final concentration) in buffer.
2) Pipette 30 µL of the master mix into the wells of interest (final *Sd*KefCTD protein concentration of 180 µM and final probe concentration of 25 nM).
3) Pipette 10 µL of the competing ligand (4 mM stock in buffer) into each of the wells of interest (final ligand concentration of 1 mM).
4) Place a black lid on the plate and shake the plate at 300 rpm for 20 min at RT using an orbital shaker.
5) Remove the black lid, place the plate in the POLARstar OPTIMA Microplate Fluorometer and start the measurement.

7.3.10 Fluorescence polarisation equilibrium competition-binding assay - The equilibrium competition-binding assay was performed in duplicate. The experimental procedure for setting up the plates for the assay using **^{t}BuSG** is outlined below:

1) Pipette 10 µL of buffer into wells: (A2-12 & B1-4); (D2-12 & E1-4).
2) Pipette 10 µL of **^{t}BuSG** (51.2 mM stock in buffer) into the wells A1 and D1.
3) Pipette 10 µL of **^{t}BuSG** (51.2 mM stock in buffer) into the wells A2 and D2. Pipette up and down in each of these wells to mix the **^{t}BuSG** with the buffer thoroughly.
4) Transfer 10 µL of diluted **^{t}BuSG** from A2 and D2 to the wells A3 and D3 respectively. Pipette up and down in each of these wells to mix thoroughly.
5) Repeat step 4 for each of the subsequent wells from (A4-12 & B1-4); (D4-12 & E1-4), such that a serial dilution of **^{t}BuSG** of factor 0.5 has been performed across the wells: (A1-12 & B1-4); (D1-12 & E1-4) (final **^{t}BuSG** concentrations 12800-0.39 µM).
6) Prepare a protein-probe master mix containing *Sd*KefCTD protein (240 µM final concentration) and the fluorescein probe (33.3 nM final concentration) in buffer.
7) Pipette 30 µL of the master mix into the wells (A1-12 & B1-4); (D1-12 & E1-4); (final *Sd*KefCTD protein concentration of 180 µM and final probe concentration of 25 nM).
8) Place a black lid on the plate and shake the plate at 300 rpm for 20 min at RT using an orbital shaker.
9) Remove the black lid, place the plate in the POLARstar OPTIMA Microplate Fluorometer and start the measurement.

7.4 Chapter 5 experimental

7.4.1 Constructing a homology model - A homology model of *Sd*KefCTD was constructed using the X-ray crystal structure of *Ec*KefCCTD (PDB code: 3L9X) as a template. Molecular Operating Environment (MOE) 2011.10 molecular modeling software was used for the alignment of sequences and the building of the homology model Hm2*Sd*KefCTD. All molecular modeling experiments were performed on an Intel® Core™2 Duo CPU E6550 @ 2.33 GHz × 2 running Ubuntu 12.04 LTS. The sequence for full-length *Sd*Kef channel was obtained from GenBank (GenBank ID: ABE53663.1) and the sequence of *Ec*KefCCTD was obtained from the 3L9X PDB. MOE-Align was used to align the query sequence of *Sd*KefCTD with the template sequence of *Ec*KefCCTD using a BLOSUM62 comparison matrix with a gapstart penalty of 14, to give a sequence identity of 37.3%. All other parameters were set to their default values. The following alignment and sequence identity matrix (Figure 7.2) were obtained:

```
>3L9X.A
GSHGMRVIIAGFGRFGQITGRLLLSSGVKMVVLDHDPDHIETLRKFGMKVFYGDATRMDL
LESAGAAKAEVLINAIDDPQTNLQLTEMVKEHFPHLQIIARARDVDHYIRLRQAGVEKPE
RETFEGALKTGRLALESLGLGPYEARERADVFRRFNIQMVEEMAMVENDTKARAAVYKRT
SAMLSEIITEDREHLSLIQRHGWQGTEEGKHTGNMADEPETKPSSTSGGLVPRGSSG
>3L9X.B
GSHGMRVIIAGFGRFGQITGRLLLSSGVKMVVLDHDPDHIETLRKFGMKVFYGDATRMDL
LESAGAAKAEVLINAIDDPQTNLQLTEMVKEHFPHLQIIARARDVDHYIRLRQAGVEKPE
RETFEGALKTGRLALESLGLGPYEARERADVFRRFNIQMVEEMAMVENDTKARAAVYKRT
SAMLSEIITEDREHLSLIQRHGWQGTEEGKHTGNMADEPETKPSSTSGGLVPRGSSG
>Sd_KefB_CTD_A
-----MVIIAGYGRFGQVVGRLLSAQGYHLSILDHSPSQIDMLRRFGNKVFYGDAARKDL
LEAAGAKDAQLLVIAIDAPDKALEIVELAHKHYPQLKIVARAIDRRHAYQYLRLGVTSFK
RETFDSAVNLGIEALTLLGNSSTVAERAGDLFSQHDNASLHELAALWGDDHTYGVAVRQR
MEDLKQVLAKDKAAQERLNTCKGGDCLEHHHHHH
>Sd_KefB_CTD_B
-----MVIIAGYGRFGQVVGRLLSAQGYHLSILDHSPSQIDMLRRFGNKVFYGDAARKDL
LEAAGAKDAQLLVIAIDAPDKALEIVELAHKHYPQLKIVARAIDRRHAYQYLRLGVTSFK
RETFDSAVNLGIEALTLLGNSSTVAERAGDLFSQHDNASLHELAALWGDDHTYGVAVRQR
MEDLKQVLAKDKAAQERLNTCKGGDCLEHHHHHH
```

Chains	1	2	3	4
1		100.0	33.1	33.1
2	100.0		100.0	33.1
3	37.3	100.0		100.0
4	37.3	37.3	100.0	

Figure 7.2 Sequence identity matrix of *Sd*KefCTD and *Ec*KefCCTD (3L9X). Performed using MOE-Align: 1) 3L9X.A; 2) 3L9X.B; 3) Sd_KefB_CTD_A; 4) Sd_KefB_CTD_B.

Once the sequences had been aligned, MOE-Homology was used to produce the homology model Hm2*Sd*KefCTD. A maximum number of 25 independent mainchain models were specified to be built; with 10 independent sidechain models built for every distinct mainchain model. The forcefield used was MMFF94x, with R-field solvation. The homology model included **ESG** as an environment for induced fit, so that **ESG** was included in both the clash test when building loops or sidechains and all minimisation stages. All other parameters were set at their default values.

7.4.2 K+ efflux studies - Silvia Ekkermann in the Miller Group, University of Aberdeen, performed all K^+ efflux experiments. The experimental procedure provided by Silvia Ekkermann is below:

K_x media - The minimal medium is represented by the general name K_x, where x indicates the concentration of K^+ in mM. For example, K_{120} medium contains a final concentration of K^+ of approximately 120 mM. The medium is made up from the appropriate buffer to which growth supplements are added (Table 7.4).

K_{120} buffer

	per Litre
$K_2HPO_4 \cdot 3H_2O$	10.5 g
KH_2PO_4	3.12 g
$(NH_4)_2SO_4$	1.05 g

K_x medium containing less than 120 mM K^+

For medium containing less than 120 mM K^+ the concentration of K^+ is adjusted by addition of sterile 1M KCl (or a dilution of this stock solution) to K_0 buffer.

K_0 buffer

	per Litre
$Na_2HPO_4 \cdot 12H_2O$	16.47 g
$NaH_2PO_4 \cdot 2H_2O$	3.12 g
$(NH_4)_2SO_4$	1.05 g

Growth supplements for minimal media

Table 7.4 Stock solutions used as growth supplements for minimal media. All solutions are prepared separately in distilled water except for $(NH_4)_2SO_4{\cdot}FeSO_4{\cdot}6H_2O$ which is made up in 0.1 M HCl*.

Supplement	Stock concentration	Dilution	Final concentration
D-glucose	20% (*w/v*)	100×	0.2% (*w/v*)
Thiamine·HCl	0.01% (*w/v*)	100×	0.0001% (*w/v*)
$MgSO_4{\cdot}\ 7H_2O$	40 mM	100×	0.4 mM
$(NH_4)_2SO_4{\cdot}FeSO_4{\cdot}6H_2O$	6 mM*	1000×	0.6 µM

K^+ efflux was measured, using a Corning 400 Flame Photometer. Freshly transformed *E. coli* MJF335 cells lacking KefB, KefC, and GshA were cultured in K_{115} minimal medium, containing 1 mM GSH, at 37 °C to an OD_{650} of ~0.8−1. The cells were filtered and washed with K_{10} buffer before being suspended in 5 mL of K^+-free K_0 buffer. The cells were added to 30 mL of K_0 medium in a stirred vessel held at 37 °C. After 5 min, compound **141** or the positive control **NEM** were added at final concentrations of 0.5 mM.

The resuspended samples were boiled for 10 min in a sealed rack to release intracellular K^+. Once cooled, the lysed cells were pelleted by centrifugation. The released K^+ in the supernatant was determined by flame photometry. The flame photometer was calibrated at 95% using 500 µM KCl and H_2O was used as a blank to set zero. Equation 7.3 was used to calculate the intracellular K^+ concentration.

$$\frac{(F.P. \div 95) \times K^{+}\ standard}{(OD_{650} \div 2)} = \mu mol\ K^{+}/g\ dry\ cell\ wt$$

Equation 7.3 Equation used to calculate the intracellular K^+ concentration, where F.P. = reading obtained from flame photometer; Potassium standard = μM concentration of KCl used to calibrate flame photometer; OD_{650} = optical density of cells when filtered.

7.4.3 Kirby-Bauer disc diffusion assay – Dr. Anthony Chan in the Conway Group, University of Oxford, performed the Kirby-Bauer disc diffusion assay. The experimental procedure provided by Dr. Anthony Chan is below:

Escherichia coli strains (Table 7.5) or transformants were inoculated onto 2× TY agar plate [16 g/L Tryptone (Oxoid Ltd.), 10 g/L yeast extract (Oxoid Ltd.), 5 g/L NaCl, 1.5% (*w/v*) Oxoid™ Bacto Agar (Thermo Fisher Scientific Inc.)] and incubated for 18-24 h at 37 °C before selection of well-isolated colonies. The full-length *Sd*Kef channel (GenBank ID: ABE53663.1) encoded by the pTrcSdKefH_6 plasmid[1] was transformed into the *E. coli* MJF335 host (Table 1), in which its endogenous *E. coli* Kef systems were knocked out,[1] as a model strain for testing in this study. For each isolate, three to five morphologically similar colonies were selected from the agar plate from the above step, and transferred into a sterile Falcon™ 50 mL conical centrifuge tube (Thermo Fisher Scientific Inc.) containing 10 mL of sterile Mueller-Hinton broth (MHB; Sigma-Aldrich Co.) using an inoculation loop. The bacterial culture was grown aerobically to 0.5 McFarland turbidity standard[9,10] at 37 °C with an agitation speed of 200 rpm During this incubation period, test discs were prepared by impregnating 20 μL of different concentrations of test compounds onto Whatman® antibiotic assay discs of 6 mm diameters (Sigma-Aldrich Co.). After the cell culture reached the cell density of 0.5 McFarland turbidity standard, the cell isolate was inoculated onto Muller-Hinton agar (MHA; Sigma-Aldrich Co.) plates using a sterile cotton swab. Subsequently the compound discs were aseptically placed onto the bacterial plates. After 24 h of incubation at 37 °C, inhibition zone diameters associated with the inhibition effect of the

compounds were recorded by averaging both vertical and horizontal diameters of each zone. The whole experimental procedures were performed according to the standards set by the Clinical and Laboratory Standards Institute.[11]

Table 7.5 *E. coli* strains used in this study.

***E. coli* strain**	**Genotype**	**Reference**
Frag1	F^-, *lacZ82*(Am), λ^-, *rha-4*, *thiE1*, *gal-33*	9
Frag5	Frag1, *Δ(kdpC-kdpA)18*	1
Frag56	Frag5, *gshA::*Tn*10*(Kan)	1
MJF335	Frag5, *lacI*, *kefB*, *kefC::*Tn*10*, *gshA::*Tn*10*(Kan)	12

7.5 Chemistry Experimental

7.5.1 General Experimental

^{1}H NMR spectra were recorded on Bruker AVIIIHD 400 nanobay (400 MHz), Bruker AVII 500 (500 MHz), Bruker AVII 500 with cryoprobe (500 MHz) or AVII 700 with cryoprobe (700 MHz) spectrometers using $CDCl_3$ (unless indicated otherwise) as a reference for the internal deuterium lock. The chemical shift data for each signal are given as δ_H in units of parts per million (ppm). The multiplicity of each signal is indicated by: s (singlet); d (doublet); t (triplet); q (quartet); qn (quintet); sp (septet); dd (doublet of doublets); ddd (doublet of doublet of doublets); or m (multiplet). The number of protons (n) for a given resonance signal is indicated by nH. Coupling constants (*J*) are quoted in Hz and are recorded to the nearest 0.1 Hz. Identical proton coupling constants (*J*) are averaged in each spectrum and reported to the nearest 0.1 Hz. The coupling constants are determined by analysis using Bruker TopSpin software.

^{13}C NMR spectra were recorded on Bruker AVIIIHD 400 nanobay (101 MHz), Bruker AVII 500 with cryoprobe (126 MHz) or AVII 700 with cryoprobe (176 MHz) spectrometers with broadband proton decoupling and internal deuterium lock. The chemical shift data for each signal are given as δ_C in units of parts per million (ppm). Where appropriate, coupling constants (*J*) are quoted in Hz and are recorded to the nearest 0.1 Hz. ^{1}H and ^{13}C spectra were assigned using 2D NMR experiments including COSY, HSQC, HMBC, CLIP-HSQC.

^{19}F NMR spectra were recorded on a Bruker AVIIIHD 400 nanobay (377 MHz) using a broadband proton decoupling pulse sequence and deuterium internal lock. The chemical shift data for each signal are given as δ_F in units of parts per million (ppm).

High-resolution mass spectra were acquired on a Bruker MicroTOF spectrometer or a Thermo Exactive spectrometer from solutions of methanol, acetonitrile or water (ESI), operating in positive or negative mode, as indicated.

Low-resolution mass spectra were acquired on either a Waters LCT Premier spectrometer or Agilent 6120 Quadrupole LC/MS spectrometer (ESI) from solutions of methanol, acetonitrile or water. *m/z* values are reported in Daltons and followed by their percentage abundance in parentheses.

Melting points were determined using a Leica Galen III hot stage microscope. The solvent of crystallisation is shown in parentheses.

Infrared Spectra were obtained from neat samples, either as solids or liquids, using a diamond ATR module. The spectra were recorded on either Bruker Tensor 27 or Shimadzu IRAffinity-1 spectrometers. Absorption maxima are reported in wavenumbers (cm^{-1}) and reported as strong (s), medium (m) or weak (w).

Specific Optical Rotations were measured using a Perkin Elmer Model 241 polarimeter or a Schmidt Haensch Unipol polarimeter, in cells with a path length of 1 dm. The light source was maintained at 589 nm. The concentration (*c*) is expressed in g/100 mL (equivalent to g/0.1 dm^3). Specific rotations are denoted $[\alpha]_D^T$ and are given in implied units of $10^{-1}degcm^2g^{-1}$ (where T = ambient temperature in °C).

Analytical HPLC was carried out on a PerkinElmer Flexar system with a Binary LC Pump and UV/VIS LC Detector. For determination of compound purity, a Dionex Acclaim® 120 column (C18, 5 µm, 120 Å, 4.6 × 150 mm) was employed using a flow rate of 1 mL/min and detection at 254 nm or 220 mm. Samples were injected in DMSO, MeOH, or H_2O. The methods used are described below. All samples were run using method 1, unless stated otherwise.

Method 1:

Step length (min)	Elapsed time (min)	%A	%B
1	1	100	0
10	11	0	100
3	14	0	100
1	15	100	0
5	20	100	0

Moblie Phases: A = 95 : 5 : 0.1, H_2O : MeCN : TFA; B = 5 : 95 : 0.1, MeCN : H_2O : TFA.

Method 2:

Step length (min)	Elapsed time (min)	%A	%B
1	1	100	0
10	11	0	100
8	19	0	100
1	20	100	0
5	25	100	0

Moblie Phases: A = 95 : 5 : 0.1, H_2O : MeCN : TFA; B = 5 : 95 : 0.1, MeCN : H_2O : TFA.

To determine the number of stereoisomers in a product, a ChiralPak® AD-H column (5 μm, 4.6 × 150 mm) was employed using an isocratic method, as stated in the compound characterisation.

Analytical thin layer chromatography (TLC) was carried out on normal phase Merck silica gel 60 F254 aluminium-supported thin layer chromatography sheets, unless otherwise stated. Visualisation was by absorption of UV light (λ_{max} 254 or 365 nm), exposure to iodine vapour or thermal development after dipping in one of: **a** ethanolic solution of: phosphomolybdic acid (PMA), ninhydrin, vanillin and sulfuric acid; **b** aqueous solution of potassium permanganate.

Reverse phase analytical thin layer chromatography (TLC) was carried out on Merck silica gel 60 RP-18 $F_{254}S$ alumininum-supported thin layer chromatography sheets. Visualisation was by absorption of UV light (λ_{max} 254 or 365 nm), or thermal development after dipping in ninhydrin solution or an aqueous solution of potassium permanganate.

Flash column chromatography was performed on a Biotage SP1 System using KP-Sil™ cartridges or carried out manually on Merk Geduran® silica gel 60 (40-63 µm), eluting with solvents as supplied, under a positive pressure of compressed air.

Reverse phase silica gel column chromatography was carried out on C_{18} Fluka silica gel, eluting with the appropriate solvents as supplied, under a positive pressure of compressed nitrogen.

Semi-preparative HPLC was performed using an Agilent ZORBAX 300SB-C18 column (5 µm, 9.4 mm × 250 mm) on a DIONEX P680 HPLC pump using a flow rate of 2.5 mL min^{-1}. The gradients employed were generally, 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA or 95:5 → 5:95; H_2O : MeCN over 15 min, unless stated otherwise. The typical method employed is as follows:

Step	Retention Time / min	Flow (mL/min)	%A	%B
0	0.000	2.500	100.0	0.0
1	3.000	2.500	100.0	0.0
2	18.000	2.500	0.0	100.0
3	20.000	2.500	100.0	0.0
4	25.000	2.500	100.0	0.0

Moblie Phases: A = 95:5:0.1, H_2O : MeCN : TFA; B = 95:5:0.1, MeCN : H_2O : TFA.

Preparative HPLC was performed using a SUPELCOSIL™ LC-Si, column (5 µm, 25 cm × 21.2 mm) on a DIONEX P680 HPLC pump using a flow rate of 2.5 mL min^{-1}. The gradient employed was, 100:0 → 50:50; Hexane : IPA. The method employed is as follows:

Step	Retention Time / min	Flow (mL/min)	%A	%B
0	0.000	10.000	100.0	0.0
1	3.000	10.000	100.0	0.0
2	4.000	10.000	50.0	50.0
3	28.000	10.000	50.0	50.0
4	29.000	10.000	0.0	100.0
5	33.000	10.000	0.0	100.0
6	34.000	10.000	100.0	0.0
7	35.000	10.000	100.0	0.0

Moblie Phases: A = Hexane; B = IPA.

Anhydrous solvents were obtained under the following conditions: Diethyl ether, toluene, CH_2Cl_2, MeCN and THF were dried by passing them through a column of activated basic alumina according to the Grubbs procedure.[13] Anhydrous NEt_3 was distilled from calcium hydride and 3 Å molecular sieves. Anhydrous DMF and MeOH were purchased from Sigma-Aldrich UK in SureSeal^{TM} bottles and used without further purification. All other solvents were used as supplied (analytical or HPLC grade) without purification.

Chemicals were purchased from Acros UK, Apollo Scientific, Enamine, Sigma Aldrich UK, Alfa Aesar UK, Fisher UK, Fluka UK, Fluorochem, Merck, Argo International Limited or TCI-Europe. All solvents and reagents were purified, when necessary, by standard techniques.[3] Where appropriate and if not otherwise stated, all non-aqueous reactions were performed in a flame-dried flask under an inert atmosphere.

In vacuo refers to removal of solvent on a Buchi™ rotary evaporator under reduced pressure in a water bath at 40 °C.

Celite® refers to Celite® 545 filter aid, treated with sodium carbonate, flux-calcined which was purchased from Sigma Aldrich.

Brine refers to a saturated aqueous solution of sodium chloride.

Petroleum Ether refers to the fraction in the boiling point range 40-60 °C.

Light (365 nm, 4 × 15 W) refers to the use of a Philips HB175 Facial Solarium. UVA, 365 nm, P = 4 × 15 W.

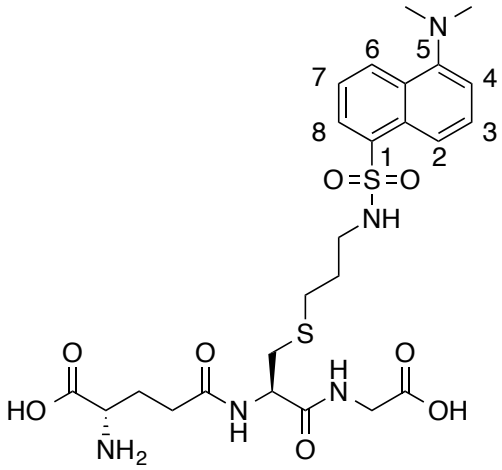
Reactions with microwave irradiation were carried out in a Discover® CEM Corporation microwave synthesiser.

Lyophilization refers to the removal of water from aqueous solutions by freeze drying using a Christ Alpha 2-4-LD lyophilizer.

Microanalyses were obtained from the Elemental Analysis Service, London Metropolitan University, London.

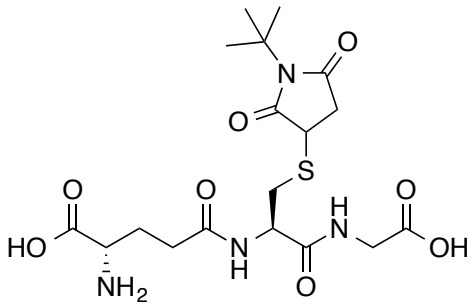
7.5.2 Synthetic procedures

_S_-((5-(Dimethylamino)naphthalen-1-yl)sulfonylaminopropyl) L-glutathione (DNGSH; 35)



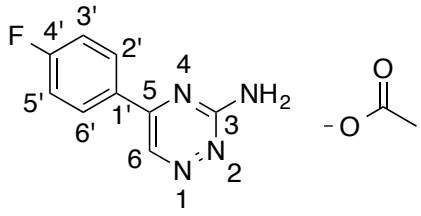
_N_-Allyl-5-(dimethylamino)naphthalene-1-sulfonamide (0.10 g, 0.34 mmol, 1.0 eq.), L-glutathione, **11**, (0.42 g, 1.37 mmol, 4.0 eq.), TCEP·HCl (0.19 g, 0.68 mmol, 2.0 eq.) and 2,2-dimethoxyphenyl acetophenone (17.0 mg, 0.07 mmol, 0.2 eq.) were stirred at RT in THF/H_2O (1:2, 3 mL) in the presence of light (365 nm, 4 × 15 W) for 5 h. After this time, the reaction mixture was washed with CH_2Cl_2 (2 × 50 mL). The aqueous layer was lyophilized and the crude material purified by RP C-18 silica gel chromatography (0:100 → 50:50; MeOH : H_2O), yielding _S_-((5-(dimethylamino)naphthalen-1-yl)sulfonylaminopropyl) L-glutathione (**DNGSH**; **35**) (22.5 mg, 11%) as a hygroscopic yellow solid: R_f 0.4 (MeOH/H_2O 50:50 (RP)); $[\alpha]_D^{25}$ = −10.4 (_c_ 0.25, H_2O) {lit.[3] $[\alpha]_D^{20}$ = −19.2 (_c_ 0.25, H_2O)}; ^{1}H NMR (400 MHz; D_2O): δ 8.55 (1H, d, _J_ 7.8, Ar C^2_H_), 8.44 (1H, d, _J_ 8.7, Ar C^6_H_), 8.32 (1H, d, _J_ 7.4, Ar C^8_H_), 7.83-7.74 (3H, m, Ar C^3_H_ & Ar C^4_H_ & Ar C^7_H_), 4.26 (1H, dd, _J_ 8.8, 5.0, cys-α-C_H_), 3.76-3.68 (3H, m, gly-CH_2, glu-α-C_H_), 3.21 (6H, s, N(CH_3)$_2$), 3.00 (2H, t, _J_ 6.8, linker-CH_2), 2.54 (1H, dd, _J_ 14.2, 5.0, cys-β-CH_2), 2.49-2.38 (3H, m, cys-β-CH_2, glu-β-CH_2), 2.21-2.13 (2H, m, linker-CH_2), 2.12-2.03 (2H, m, glu-ɣ-CH_2), 1.45 (2H, dt, _J_ 13.3, 6.8, linker-CH_2), _m/z_ (ES$^+$) 598 ([M+H]$^+$, 100%). The data are in good agreement with the literature values.[3]

***S*-(*N*-*tert*-Butylsuccinimido) L-glutathione (^tBuSG; 24)**



To a stirred solution of L-glutathione, **11**, (0.2 g, 0.65 mmol, 1.0 eq.) and NaOH (0.03 g, 0.65 mmol, 1.0 eq.) in H_2O (5 mL) was added *N*-*tert*-butyl-maleimide (0.10 g, 0.09 mL, 0.65 mmol, 1.0 eq.). The resulting solution was stirred at RT for 20 min. After this time, the reaction solution was lyophilized, purified by RP C-18 silica gel chromatography, eluting with H_2O to afford *S*-(*N*-*tert*-butylsuccinimido) L-glutathione, (**^tBuSG**; **24**) (mixture of diastereoisomers) as a colourless hygroscopic solid (0.26 g, 86%): $[\alpha]_D^{20}$ = −13.3 (*c* 1.0, H_2O) {lit.[3] $[\alpha]_D^{20}$ = −4.0 (*c* 0.5, H_2O)}; ^{1}H NMR (400 MHz; D_2O): δ 4.68-4.60 (1H, m, cys-α-C*H*), 3.91-3.81 (1H, m, succinimide-C*H*), 3.81-3.67 (3H, m, glu-α-C*H*, gly-CH_2), 3.32 (0.5H, dd, *J* 14.2, 4.9, cys-β-CH_2), 3.26-3.04 (2H, m, cys-β-CH_2, succinimide-CH_2), 2.96 (0.5H, dd, *J* 14.2, 9.2, cys-β-CH_2), 2.60-2.43 (3H, m, glu-ɣ-CH_2, succinimide-CH_2), 2.20-2.07 (2H, m, glu-β-CH_2), 1.51 (4.5H, s, NC(CH_3)$_3$), 1.51 (4.5H, s, NC(CH_3)$_3$); *m/z* (ES^-) 459 ($[M-H]^-$, 100%). The data are in good agreement with the literature values, with the exception of the $[\alpha]_D^{20}$ values, which both have a negative specific rotation, but different magnitudes due to them being run at different concentrations.[14]

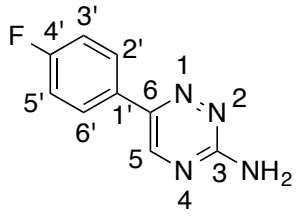
5-(4'-Fluorophenyl)-1,2,4-triazin-3-amine acetate (42)



To a stirred solution of 2,2-dibromo-4'-fluoroacetophenone, **50**, (2.59 g, 8.75 mmol, 1.0 eq.) in anhydrous THF (13.0 mL) under nitrogen was added morpholine (3.20 g, 3.18 mL, 36.76 mmol, 4.2 eq.) at RT. The resulting solution was slowly heated to 50 °C under nitrogen over 1 h and the temperature was maintained for 24 h. After cooling to RT, the suspension was filtered through a fritted funnel and the wet cake was washed with THF (10 mL). The combined filtrate was then concentrated *in vacuo* to give the crude aminal intermediate, **51**, (2.75 g), which was used without further purification.

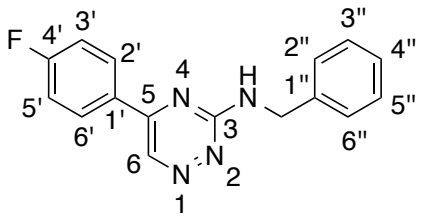
To a stirred solution of the crude aminal intermediate, **51**, (2.75 g) in MeOH (13.7 mL) under nitrogen was added aminoguanidine bicarbonate (1.19 g, 8.75 mmol, 1.0 eq. with respect to dibromide), followed by dropwise addition of glacial acetic acid (1.58 g, 1.50 mL, 26.26 mmol, 3.0 eq. with respect to dibromide) over 10 min at RT. The resulting suspension was stirred at RT for 2 h, and then heated at 80 °C for 14 h. After this time, the suspension was cooled to RT and concentrated to half its volume *in vacuo*, cooled to 0 °C, aged for 1 h, filtered through a fritted funnel and washed with cold MeOH : H_2O (4:1). The collected solid was purified by crystallisation from ethanol : toluene (3:1) to yield 5-(4'-fluorophenyl)-1,2,4-triazin-3-amine acetate, **42**, (0.51 g) as a yellow solid: R_f 0.1 (EtOAc/petrol 40:60); ^{1}H NMR (400 MHz; $CDCl_3$): δ 9.05 (1H, s, Ar C^6H), 8.18-8.09 (2H, m, Ar $C^{2'}H$ & Ar $C^{6'}H$), 7.29-7.20 (2H, m, Ar $C^{3'}H$ & Ar $C^{5'}H$), 5.74 (2H, s, ArNH_2); ^{19}F NMR (376.6 MHz; $CDCl_3$): δ −106.4 (Ar $C^{4'}F$); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 213.0537. $C_9H_7FN_4Na$ requires M^+ 213.0547); *m/z* (ES^+) 191 ($[M+H]^+$, 100%). The data are in good agreement with the available literature values.[15]

6-(4'-Fluorophenyl)-1,2,4-triazin-3-amine (49)



4'-Fluorophenylglyoxal hydrate, **48**, (0.20 g, 1.18 mmol, 1.0 eq.) and aminoguanidine bicarbonate, **47**, (0.16 g, 1.18 mmol, 1.0 eq.) were stirred in aqueous ethanol (2 mL, 90% *v/v*) at 25 °C for 22 h. After this time, the reaction mixture was heated to 50 °C for 1 h before being concentrated *in vacuo*. The resulting residue was partitioned between aqueous 2 M HCl (50 mL) and EtOAc (50 mL) and the aqueous layer was collected. Solid Na_2CO_3 was slowly added portionwise to the aqueous layer at 0 °C until a pH of 7 was attained. EtOAc was then used to extract the aqueous layer and the organic layers were combined, washed with H_2O, brine, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (30:70 → 40:60; EtOAc : petroleum ether) to afford 6-(4'-fluorophenyl)-1,2,4-triazin-3-amine, **49**, (0.08 g, 36%) as a pale yellow solid: R_f 0.2 (EtOAc/petroleum ether 40:60); mp 157-161 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} (neat): 3318 (m), 3196 (m), 1738 (m), 1627 (m), 1602 (m), 1578 (m), 1498 (m), 1472 (s), 1368 (m), 1217 (s), 1159 (m), 1100 (m), 1044 (m); ^{1}H NMR (500 MHz; $CDCl_3$): δ 8.56 (1H, s, Ar C^5H), 7.95-7.90 (2H, m, Ar $C^{2'}H$ & Ar $C^{6'}H$), 7.22-7.16 (2H, m, Ar $C^{3'}H$ & Ar $C^{5'}H$), 5.38 (2H, s, ArNH_2); ^{19}F NMR (376.6 MHz; $CDCl_3$): δ −111.8 (Ar $C^{4'}F$); ^{13}C NMR (126 MHz; $CDCl_3$): δ 163.9 (d, *J* 249.2 Hz, $C^{4'}$), 161.7 (C^3), 149.6 (C^6), 147.7 (C^5), 130.4 (d, *J* 2.9 Hz, $C^{1'}$), 127.7 (d, *J* 8.3 Hz, $C^{2'}$ & $C^{6'}$), 116.3 (d, *J* 21.4 Hz, $C^{3'}$ & $C^{5'}$); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 213.0548. $C_9H_7FN_4Na$ requires M^+ 213.0547); *m/z* (ES^+) 223 ($[M+MeOH+H]^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 97:3 → 3:97; H_2O : MeCN, with 0.25 mM NEt_3 and acetic acid) Ret. Time = 10.198 min, Purity: 96.15%.

_N_-Benzyl-5-(4'-fluorophenyl)-1,2,4-triazin-3-amine (53)

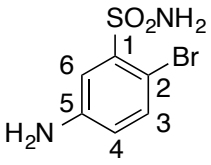


To a mixture of *tert*-butyl nitrite (0.09 g, 0.11 mL, 0.92 mmol, 3.5 eq., 90%) and copper(II) chloride (0.04 g, 0.29 mmol, 1.1 eq.) in anhydrous MeCN (0.3 mL) heated at 65 °C, was added a solution of 5-(4'-fluorophenyl)-1,2,4-triazin-3-amine, **42**, (0.05 g, 0.26 mmol, 1.0 eq.) in anhydrous MeCN (1.0 mL) and CH_2Cl_2 (0.7 mL). The reaction mixture was stirred under argon at 65 °C for 3 h. The solution was then poured into aqueous 6 M hydrochloric acid (10 mL) cooled to 0 °C. The mixture was extracted with diethyl ether (2 × 50 mL) and the combined organic layers were washed with brine (50 mL), then dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo* to afford 3-chloro-5-(4'-fluorophenyl)-1,2,4-triazine, **52**, (0.04 g) as a yellow solid, which was used without further purification.

A solution of 3-chloro-5-(4'-fluorophenyl)-1,2,4-triazine, **52**, (0.04 g, 0.19 mmol, 1.0 eq.), benzyl amine (0.03 g, 0.03 mL, 0.25 mmol, 1.3 eq.) and *N*,*N*-diisopropylethylamine (0.02 g, 0.03 mL, 0.19 mmol, 1.0 eq.) in anhydrous MeCN (0.7 mL) was stirred at 85 °C for 15.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (12:88 → 100:0; EtOAc : petroleum ether) to afford *N*-benzyl-5-(4'-fluorophenyl)-1,2,4-triazin-3-amine, **53**, (0.02 g, 29% over two steps) as a yellow solid: R_f 0.4 (EtOAc/petrol 50:50); mp 128-130 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} (neat) 1601 (m), 1533 (m), 1512 (m), 1476 (m), 1347 (m), 1309 (m), 1287 (m), 1232 (m), 1158 (m), 1119 (m), 1095 (m), 1081 (m), 1045 (m); ^{1}H NMR (500 MHz; $CDCl_3$): δ 9.02 (1H, s, Ar C^6H), 8.12-8.05 (2H, m, Ar $C^{2'}H$ & Ar $C^{6'}H$), 7.43-7.38 (2H, m, Ar $C^{2''}H$ & Ar $C^{6''}H$), 7.38-7.33 (2H, m, Ar $C^{3''}H$ & Ar $C^{5''}H$), 7.32-7.27 (1H, m,

Ar C$^{4''}$*H*), 7.23-7.16 (2H, m, Ar C$^{3'}$*H* & Ar C$^{5'}$*H*), 4.79 (2H, d, *J* 4.7, ArC*H*$_2$NH); ^{19}F NMR (376.6 MHz; $CDCl_3$): δ −107.1 (Ar C$^{4'}$*F*); ^{13}C NMR (126 MHz; $CDCl_3$): δ 165.4 (d, *J* 253.7, *C*$^{4'}$), 161.7 (*C*^{3}), 155.0 (*C*^{5}), 138.5 (*C*$^{1''}$), 137.6 (*C*^{6}), 130.4 (d, *J* 2.9, *C*$^{1'}$), 129.8 (d, *J* 8.6, *C*$^{2'}$ & *C*$^{6'}$), 128.9 (*C*$^{3''}$ & *C*$^{5''}$), 127.8 (*C*$^{2''}$& *C*$^{6''}$), 127.7 (*C*$^{4''}$), 116.5 (d, *J* 21.9, *C*$^{3'}$ & *C*$^{5'}$), 45.4 (Ar*C*H$_2$NH); HRMS *m/z* (ES$^+$) (Found: $[M+H]^+$ 281.1201. $C_{16}H_{14}FN_4$ requires M^+ 281.1197); *m/z* (ES$^+$) 281 ($[M+H]^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 11.801 min, Purity: 95.65%.

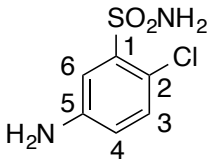
5-Amino-2-bromobenzene-1-sulfonamide (56b)



To a solution of 2-bromo-5-nitrobenzene-1-sulfonamide, **61b**, (0.18 g, 0.63 mmol, 1.0 eq.) in THF/conc. HCl (1:1, 7.6 mL), was added zinc powder (1.03 g, 15.75 mmol, 25.0 eq.) portionwise. The mixture was stirred at RT for 4 h and then the pH was raised to ca. 7 with aqueous saturated $NaHCO_3$. The mixture was extracted with EtOAc (2 × 50 mL) and the combined organic layers were concentrated *in vacuo*. The resulting residue adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (40:60; EtOAc : petroleum ether) to afford 5-amino-2-bromobenzene-1-sulfonamide, **56b**, as a brown solid (0.13 g, 79%): R_f 0.3 (EtOAc/petroleum ether 50:50); mp 150-154 °C (from EtOAc/petroleum ether) [lit.[15] 148-150 °C]; ν_{max}/cm^{-1} (neat) 3455 (m), 3371 (m), 3350 (m), 3248 (m), 2161 (m), 1631 (m), 1591 (m), 1556 (m), 1458 (m), 1417 (m), 1324 (m), 1156 (s), 1141 (m), 1107 (m), 1019 (m); ^{1}H NMR (500 MHz; D_6-acetone): δ 7.43 (1H, d, *J* 2.8, Ar C^{6}*H*), 7.39 (1H, d, *J* 8.5, Ar C^{3}*H*), 6.76 (1H, dd, *J* 8.5, 2.8, Ar C^{4}*H*), 6.49 (2H, s, ArSO_2N*H*$_2$), 5.26 (2H, s, ArN*H*$_2$); ^{13}C NMR (126 MHz; D_6-acetone): δ 149.3 (*C*^{5}), 143.7 (*C*^{1}), 136.0 (*C*^{3}), 119.0 (*C*^{4}), 115.8 (*C*^{6}), 104.1 (*C*^{2}); HRMS *m/z* (ES$^+$) (Found: [M^{79}Br+Na]$^+$ 272.9301. $C_6H_7{}^{79}BrN_2NaO_2S$ requires M^+

272.9304); HRMS *m/z* (ES^+) (Found: $[M^{81}Br+Na]^+$ 274.9274. $C_6H_7{}^{81}BrN_2NaO_2S$ requires M^+ 274.9283); *m/z* (ES^-) 251 ($[M^{81}Br-H]^-$, 100%), 249 ($[M^{79}Br-H]^-$, 67%); Anal. Calcd. (Found C, 28.66; H, 2.71; N, 11.13. $C_6H_7BrN_2O_2S$ requires C, 28.7; H, 2.81; N, 11.16). The data are in good agreement with the available literature values.[16]

5-Amino-2-chlorobenzene-1-sulfonamide (56c)



Method A

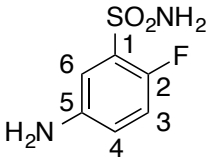
A mixture of 2-chloro-5-nitrobenzene-1-sulfonamide, **61c**, (0.40 g, 1.68 mmol, 1.0 eq.) and palladium on carbon (0.04 g, 0.04 mmol, 0.02 eq., 10% wt) in MeOH (8.4 mL) was stirred under hydrogen at atmospheric pressure for 24 h at RT. After this time, the reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo*. The resulting crude was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (50:50; EtOAc : petroleum ether) to afford 5-amino-2-chlorobenzene-1-sulfonamide, **56c**, as a brown solid (0.03 g, 9%). Data are shown below.

Method B

To a solution of 2-chloro-5-nitrobenzene-1-sulfonamide, **61c**, (0.30 g, 1.27 mmol, 1.0 eq.) in THF/conc. HCl (15.3 mL, 1:1) was added zinc powder (2.07 g, 31.70 mmol, 25.0 eq.) portionwise. The mixture was stirred at RT for 7 h and then the pH was raised to ca. 7 with aqueous saturated $NaHCO_3$. The mixture was extracted with EtOAc (3 × 50 mL). The combined organic components were concentrated *in vacuo*, the resulting residue adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (40:60; EtOAc : petroleum ether) to afford 5-amino-2-chlorobenzene-1-sulfonamide, **56c**, as a brown solid (0.23 g, 89%):

R_f 0.2 (EtOAc/petroleum ether 40:60); mp 162-166 °C (from EtOAc/petroleum ether) [lit.[16] 166-167 °C]; ν_{max}/cm^{-1} (neat) 3456 (m), 3374 (m), 3343 (m), 3244 (m), 1632 (m), 1596 (m), 1561 (m), 1462 (m), 1422 (m), 1326 (m), 1157 (s), 1142 (m), 1118 (m), 1067 (m), 1030 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 7.33 (2H, s, $ArSO_2NH_2$), 7.22 (1H, d, *J* 2.8, Ar C^6H), 7.18 (1H, d, *J* 8.6, Ar C^3H), 6.68 (1H, dd, *J* 8.6, 2.8, Ar C^4H), 5.67 (s, 2H, $ArNH_2$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 148.0 (C^5), 140.9 (C^1), 131.6 (C^3), 117.4 (C^4), 114.8 (C^2), 113.7 (C^6); HRMS *m/z* (ES^-) (Found: $[M-H]^-$ 204.9837. $C_6H_6ClN_2O_2S$ requires M^- 204.9844); *m/z* (ES^-) 205 ($[M-H]^-$, 100%); Anal. Calcd. (Found C, 34.82; H, 3.29; N, 13.48. $C_6H_7ClN_2O_2S$ requires C, 34.87; H, 3.41; N, 13.56). The data are in good agreement with the available literature values.[15]

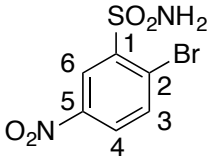
5-Amino-2-fluorobenzene-1-sulfonamide (56d)



A mixture of 2-fluoro-5-nitrobenzene-1-sulfonamide, **61d**, (0.50 g, 2.27 mmol, 1.0 eq.) and palladium on carbon (0.12 g, 0.11 mmol, 0.05 eq., 10% wt) in MeOH (11.4 mL) was stirred under hydrogen at atmospheric pressure at RT for 16 h. After this time, the reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo*. The resulting crude was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (60:40; EtOAc : petroleum ether) to afford 5-amino-2-fluorobenzene-1-sulfonamide, **56d**, as a beige solid (0.39 g, 91%): R_f 0.3 (EtOAc/petroleum ether 60:40); mp 139-141 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} (neat) 3411 (m), 3342 (m), 1492 (m), 1334 (m), 1280 (m), 1213 (m), 1160 (s), 1063 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 7.43 (2H, s, $ArSO_2NH_2$), 7.03 (1H, dd, *J* 10.3, 9.0, Ar C^3H), 6.99 (1H, dd, *J* 6.0, 3.0, Ar C^6H), 6.71 (1H, ddd, *J* 9.0, 4.0, 3.0, Ar C^4H), 5.36 (2H, s, $ArNH_2$); ^{19}F NMR (377 MHz; D_6-DMSO): δ −129.3 (Ar C^2F);

^{13}C NMR (126 MHz; D_6-DMSO): δ 149.4 (d, *J* 239.6, *C*^{2}), 145.2 (d, *J* 1.9, *C*^{5}), 131.3 (d, *J* 15.3, *C*^{1}), 117.5 (d, *J* 85.8, *C*^{4}), 117.4 (d, *J* 101.1, *C*^{3}), 112.1 (*C*^{6}); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 213.0098. $C_6H_7FN_2NaO_2S$ requires M^+ 213.0104); *m/z* (ES^-) 189 ($[M-H]^-$, 100%); Anal. Calcd. (Found C, 37.92; H, 3.72; N, 14.62. $C_6H_7FN_2O_2S$ requires C, 37.89; H, 3.71; N, 14.73). The data are in good agreement with the available literature values.[17]

2-Bromo-5-nitrobenzene-1-sulfonamide (61b)

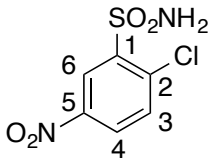


Chlorosulfonic acid (10.52 g, 6.00 mL, 90.27 mmol, 9.1 eq.) was added dropwise to 1-bromo-4-nitrobenzene, **59b**, (2.00 g, 9.90 mmol, 1.0 eq.) over 5 min at 0 °C. The reaction mixture was stirred under argon at 120 °C for 16 h. After this time, the reaction mixture was poured onto crushed ice (100 g) and extracted with $CHCl_3$ (3 × 100 mL), washed with brine (100 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting brown solid (2.64 g), which contained 2-bromo-5-nitrobenzene-1-sulfonyl chloride, **60b**, was used without further purification.

To a stirred solution of the brown solid, which contained 2-bromo-5-nitrobenzene-1-sulfonyl chloride, **60b**, (2.64 g) in EtOAc (70 mL), was added ammonium hydroxide (70 mL, 25% solution in H_2O) dropwise over 5 min at 0 °C, and the reaction mixture was stirred at RT for 4 h. After this time, the organic layer was separated and the aqueous layer was extracted with EtOAc (2 × 50 mL). The organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite®, and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (40:60; EtOAc : petroleum ether), to yield 2-bromo-5-nitrobenzene-1-sulfonamide, **61b**, as a purple solid (0.83 g, 30% over two steps):

R_f 0.4 (EtOAc/petroleum ether 50:50); mp 203-207 °C (from EtOAc/petroleum ether) [lit.[15] 204-205 °C]; ν_{max}/cm^{-1} (neat) 3407 (m), 3285 (m), 1601 (m), 1570 (m), 1536 (m), 1444 (m), 1349 (s), 1246 (m), 1164 (s), 1130 (m), 1092 (m), 1029 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 8.69 (1H, d, *J* 2.7, Ar C^6*H*), 8.30 (1H, dd, *J* 8.7, *J* 2.7, Ar C^4*H*), 8.14 (1H, d, *J* 8.7, Ar C^3*H*), 7.98 (2H, s, $ArSO_2NH_2$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 146.5 (*C*^{5}), 144.0 (*C*^{1}), 136.9 (*C*^{3}), 127.5 (*C*^{4}), 126.4 (*C*^{2}), 123.6 (*C*^{6}); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 302.9054. $C_6H_5BrN_2NaO_4S$ requires M^+ 302.9046); *m/z* (ES^-) 281 ($[M^{81}Br-H]^-$, 100%), 279 ($[M^{79}Br-H]^-$, 89%); Anal. Calcd. (Found C, 25.77; H, 1.69; N, 9.87. $C_6H_5BrN_2O_4S$ requires C, 25.64; H, 1.79; N, 9.97). The data are in good agreement with the available literature values.[16]

2-Chloro-5-nitrobenzene-1-sulfonamide (61c)

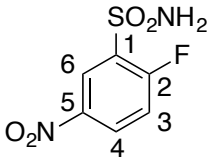


Chlorosulfonic acid (5.08 g, 2.90 mL, 43.63 mmol, 6.9 eq.) was added dropwise to 1-chloro-4-nitrobenzene **59c** (1.00 g, 6.35 mmol, 1.0 eq.) over 5 min at 0 °C. The reaction mixture was stirred under argon at 120 °C for 16 h. After this time, the reaction mixture was poured onto crushed ice (100 g) and extracted with CH_2Cl_2 (3 × 100 mL), washed with brine (100 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting brown solid (1.18 g), which contained 2-chloro-5-nitrobenzene-1-sulfonyl chloride, **60c**, was used without further purification.

To a stirred solution of the brown solid, which contained 2-chloro-5-nitrobenzene-1-sulfonyl chloride, **60c**, (1.04 g) in THF (1.3 mL) was added ammonium hydroxide (1.3 mL, 25% solution in H_2O) dropwise over 5 min at 0 °C and the reaction mixture was stirred at RT for 16 h. After this time, the organic layer was separated and the aqueous layer was extracted with EtOAc

(2 × 50 mL). The organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (20:80 → 40:60; EtOAc : petroleum ether) to yield 2-chloro-5-nitrobenzene-1-sulfonamide, **61c**, as an off-white solid (0.45 g, 34% over two steps): R_f 0.4 (EtOAc/petroleum ether 40:60); mp 184-188 °C (from EtOAc/petroleum ether) [lit.[16] 184-185 °C]; ν_{max}/cm^{-1} (neat) 3409 (m), 3289 (m), 1605 (m), 1574 (m), 1537 (m), 1450 (m), 1386 (m), 1352 (s), 1294 (m), 1245 (m), 1165 (m), 1132 (m), 1105 (m), 1043 (m); ^{1}H NMR (400 MHz; D_6-DMSO): δ 8.67 (1H, d, *J* 2.7, Ar C^6*H*), 8.43 (1H, dd, *J* 8.7, 2.7, Ar C^4*H*), 8.02 (2H, s, $ArSO_2NH_2$), 7.97 (1H, d, *J* 8.7, Ar C^3*H*); ^{13}C NMR (126 MHz; D_6-DMSO): δ 145.9 (*C*^{5}), 142.1 (*C*^{1}), 137.2 (*C*^{2}), 133.3 (*C*^{3}), 127.8 (*C*^{4}), 123.7 (*C*^{6}); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 258.9551. $C_6H_5ClN_2NaO_4S$ requires M^+ 258.9551); *m/z* (ES^-) 235 ($[M-H]^-$, 100%); Anal. Calcd. (Found C, 30.58; H, 2.11; N, 11.92. $C_6H_5ClN_2O_4S$ requires C, 30.45; H, 2.13; N, 11.84). The data are in good agreement with the available literature values.[18]

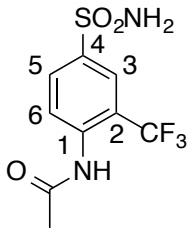
2-Fluoro-5-nitrobenzene-1-sulfonamide (61d)



Chlorosulfonic acid (6.66 g, 3.80 mL, 57.17 mmol, 8.1 eq.) was added dropwise to 1-fluoro-4-nitrobenzene **59d** (1.00 g, 7.09 mmol, 1.0 eq.) over 5 min at 0 °C. The reaction mixture was stirred under argon at 110 °C for 20 h. After this time, the reaction mixture was poured onto crushed ice (100 g) and extracted with diethyl ether (3 × 50 mL). The organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting red-brown oil (1.38 g), which contained 2-fluoro-5-nitrobenzene-1-sulfonyl chloride, **60d**, was used without further purification.

To a solution of the red-brown oil, which contained 2-fluoro-5-nitrobenzene-1-sulfonyl chloride, **60d**, (1.38 g) in EtOAc (20 mL) was added ammonium hydroxide (20 mL, 25% solution in H_2O) dropwise over 5 min at 0 °C and the reaction mixture was stirred at RT for 4 h. After this time, the organic layer was separated and the aqueous layer was extracted with EtOAc (2 × 50 mL). The organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (40:60; EtOAc : petroleum ether) to yield 2-fluoro-5-nitrobenzene-1-sulfonamide, **61d**, as a yellow solid (0.75 g, 48% over two steps): R_f 0.4 (EtOAc/petroleum ether 50:50); mp 133-137 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} (neat) 3381 (m), 3291 (m), 1587 (m), 1537 (m), 1473 (m), 1362 (s), 1342 (s), 1264 (m), 1233 (m), 1173 (m), 1141 (m), 1127 (m), 1064 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 8.57-8.51 (2H, m, Ar C^4*H* & Ar C^6*H*), 8.07 (2H, s, Ar-SO_2NH_2), 7.80-7.73 (1H, m, Ar C^3*H*), ^{19}F NMR (377 MHz; D_6-DMSO): δ −100.9 (Ar C^2*F*); ^{13}C NMR (126 MHz; D_6-DMSO): δ 161.4 (d, *J* 263.1, *C*^{2}), 143.4 (d, *J* 2.9, *C*^{5}), 132.6 (d, *J* 17.2, *C*^{1}), 130.0 (d, *J* 10.9, *C*^{6}), 124.1 (d, *J* 1.9, *C*^{4}), 119.1 (d, *J* 24.7, *C*^{3}); HRMS *m/z* (ES$^+$) (Found: $[M+Na]^+$ 242.9851. $C_6H_5FN_2NaO_4S$ requires M^+ 242.9846); *m/z* (ES$^-$) 219 ($[M-H]^-$, 100%); Anal. Calcd. (Found C, 32.81; H, 2.13; N, 12.73. $C_6H_5FN_2O_4S$ requires C, 32.73; H, 2.29; N, 12.72). The data are in good agreement with the available literature values.[19]

_N_-[4-Sulfamoyl-2-(trifluoromethyl)phenyl]acetamide (67a)

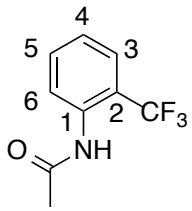


Chlorosulfonic acid (1.00 g, 0.57 mL, 8.56 mmol, 5.6 eq.) was added dropwise to *N*-[2-(trifluoromethyl)phenyl]acetamide, **69a**, (0.31 g, 1.53 mmol, 1.0 eq.) over 5 min at 0 °C. The

reaction mixture was stirred under argon at 60 °C for 21 h. After this time, the reaction mixture was poured onto crushed ice (50 g) and extracted with $CHCl_3$ (2 x 50 mL). The organic layers were combined, washed with brine (50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting off-white solid (0.22 g), which contained 4-acetamido-3-(trifluoromethyl)benzene-1-sulfonyl chloride, **70a**, was used without further purification.

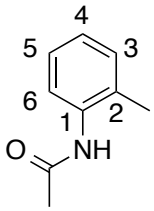
A solution of the off-white solid, which contained 4-acetamido-3-(trifluoromethyl)benzene-1-sulfonyl chloride, **70a**, (0.22 g) in THF (1.53 mL) was added dropwise to ammonium hydroxide (1.53 mL, 25% solution in H_2O) over 5 min at 0 °C. The reaction mixture was stirred at RT for 16 h. After this time, the reaction mixture was concentrated *in vacuo* and the residue was partitioned between H_2O (25 mL) and EtOAc (50 mL). The aqueous layer was extracted with EtOAc (2 × 25 mL), the organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (40:60; EtOAc : petroleum ether) to afford *N*-[4-sulfamoyl-2-(trifluoromethyl)phenyl]acetamide, **67a**, (56.4 mg, 13% over two steps) as a colourless solid: R_f 0.2 (EtOAc/petroleum ether 60:40); mp 195-197 °C (from aqueous ethanol, 95% *v/v*); ν_{max}/cm^{-1} (neat) 3331 (m), 1683 (m), 1586 (m), 1514 (m), 1420 (m), 1376 (m), 1318 (m), 1292 (m), 1275 (m), 1185 (m), 1123 (s), 1056 (s); ^{1}H NMR (400 MHz; CD_3OD): δ 8.21 (d, 1H, *J* 2.0, Ar C^3*H*), 8.12 (dd, 1H, *J* 8.5, 2.0, Ar C^5*H*), 7.83 (d, 1H, *J* 8.5, Ar C^6*H*), 2.20 (s, 3H, ArNHCOC*H*$_3$); ^{19}F NMR (377 MHz; $CDCl_3$): δ −62.3 (ArC*F*$_3$); ^{13}C NMR (126 MHz; CD_3OD): δ 172.7 (ArNH*C*OCH$_3$), 143.4 (*C*^{1}), 139.7 (*C*^{4}), 131.5 (*C*^{5}), 131.1 (*C*^{6}), 126.4 (q, *J* 31.1, *C*^{2}), 125.7 (q, *J* 5.3, *C*^{3}), 124.5 (q, *J* 272.6, Ar*C*F$_3$), 23.1 (ArNHCO*C*H$_3$); HRMS *m/z* (ES$^-$) (Found: [M−H]$^-$ 281.0224. $C_9H_8F_3N_2O_3S$ requires M$^-$ 281.0213); *m/z* (ES$^-$) 281 ([M−H]$^-$, 100%); Anal. Calcd. (Found C, 38.37; H, 3.19; N, 9.83. $C_9H_9F_3N_2O_3S$ requires C, 38.30; H, 3.21; N, 9.93).

***N*-[2-(Trifluoromethyl)phenyl]acetamide (69a)**



To a stirred solution of 2-(trifluoromethyl)aniline, **68a**, (1.00 g, 0.78 mL, 6.21 mmol, 1.0 eq.) and anhydrous NEt_3 (0.63 g, 0.87 mL, 6.21 mmol, 1.0 eq.) at 0 °C in anhydrous $CHCl_3$ (6 mL) was added acetyl chloride (0.54 g, 0.49 mL, 6.83 mmol, 1.1 eq.) dropwise over 5 min. The reaction mixture was stirred under argon at RT for 5.5 h before being concentrated *in vacuo*. The resulting oily residue was then stirred with H_2O (3 mL) until a precipitate appeared. The precipitate was filtered, washed with H_2O (10 mL), dried under high vaccum and purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-[2-(trifluoromethyl)phenyl]acetamide, **69a**, (0.94 g, 74%) as a colourless solid: R_f 0.4 (EtOAc/petroleum ether 50:50); mp 88-92 °C (from aqueous ethanol, 95% *v/v*) (lit.[19-21] 94-95 °C]; ^{1}H NMR (500 MHz; $CDCl_3$): δ 8.12 (1H, d, *J* 7.7, Ar C^3*H*), 7.59 (1H, d, *J* 7.6, Ar C^6*H*), 7.53 (1H, dd, *J* 8.9, 7.7, Ar C^4*H*), 7.47 (1H, s, ArN*H*COCH$_3$), 7.22 (1H, dd, *J* 8.9, 7.6, Ar C^5*H*), 2.20 (3H, s, ArNHCOC*H*$_3$); ^{19}F NMR (470 MHz; $CDCl_3$): δ −60.6 (ArC*F*$_3$); HRMS *m/z* (ES$^+$) (Found: $[M+Na]^+$ 226.0441. $C_9H_8F_3NNaO$ requires $[M+Na]^+$ 226.0450); *m/z* (ES$^+$) 204 ($[M+H]^+$, 100%). The data are in good agreement with literature values.[22]

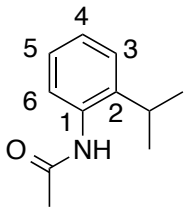
***N*-(2-Methylphenyl)acetamide (69b)**



To a stirred solution of *o*-toluidine, **68b**, (1.00 g, 0.99 mL, 9.33 mmol, 1.0 eq.) and anhydrous NEt_3 (1.42 g, 1.95 mL, 14.00 mmol, 1.5 eq.) at 0 °C in anhydrous CH_2Cl_2 (38 mL) was added acetyl chloride (0.81 g, 0.73 mL, 10.27 mmol, 1.1 eq.) dropwise over 5 min. The reaction mixture was

stirred under nitrogen at RT for 3.5 h. After this time the reaction mixture was diluted with H_2O (30 mL) and CH_2Cl_2 (20 mL). The organic layer was washed with aqueous 1 M HCl (3 × 20 mL) and brine (2 × 20 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-(2-methylphenyl)acetamide, **69b**, (0.85 g, 61%) as a colourless solid: R_f 0.2 (EtOAc/petroleum ether 50:50); mp 99-103 °C (from aqueous ethanol, 95% *v/v*) [lit.[23] 98-100 °C]; ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.72 (1H, d, *J* 8.0, Ar C^3*H*), 7.23-7.04 (4H, m, ArN*H*COCH$_3$ & Ar C^4*H* & Ar C^5*H* & Ar C^6*H*), 2.24 (s, 3H, C*H*$_3$), 2.19 (s, 3H, C*H*$_3$); *m/z* (ES$^+$) 321 ([2M+Na]$^+$, 100%). The data are in good agreement with literature values.[5,[24]]

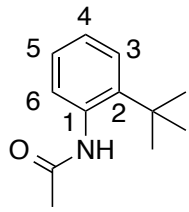
_N_-[2-(Propan-2-yl)phenyl]acetamide (69c)



To a stirred solution of 2-(propan-2-yl)aniline, **68c**, (1.00 g, 1.05 mL, 7.40 mmol, 1.0 eq.) and anhydrous NEt_3 (1.12 g, 1.55 mL, 11.09 mmol, 1.5 eq.) at 0 °C in anhydrous CH_2Cl_2 (30 mL) was added acetyl chloride (0.64 g, 0.58 mL, 8.14 mmol, 1.1 eq.) dropwise over 5 min. The reaction mixture was stirred under argon at RT for 3 h. After this time the reaction mixture was diluted with H_2O (30 mL) and CH_2Cl_2 (20 mL). The organic layer was washed with aqueous 1 M HCl (3 × 20 mL) and brine (2 × 20 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-[2-(propan-2-yl)phenyl]acetamide, **69c**, (1.31 g, 100%) as an off-white solid: R_f 0.3 (EtOAc/petroleum ether 50:50); mp 78-80 °C (from aqueous ethanol, 95% *v/v*) [lit.[24] 71-72 °C]; ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.55-7.62 (1H, m, Ar C*H*), 7.27-7.31 (1H, m, Ar C*H*), 7.16-7.23 (2H, m, Ar C*H*), 7.13 (1H, s, ArN*H*COCH$_3$), 3.04 (1H, sp, *J* 6.9, ArC*H*(CH$_3$)$_2$), 2.19 (3H, s, ArNHCOC*H*$_3$),

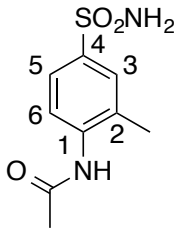
1.24 (6H, d, *J* 6.9, ArCH(C*H*$_3$)$_2$); *m/z* (ES$^+$) 377 ([2M+Na]$^+$, 100%). The data are in good agreement with literature values.[25]

_N_-(2-_tert_-Butylphenyl)acetamide (69d)



To a stirred solution of 2-*tert*-butylaniline, **68d**, (1.00 g, 1.04 mL, 6.70 mmol, 1.0 eq.) and anhydrous NEt_3 (1.02 g, 1.40 mL, 10.05 mmol, 1.5 eq.) at 0 °C in anhydrous CH_2Cl_2 (30 mL) was added acetyl chloride (0.58 g, 0.53 mL, 7.37 mmol, 1.1 eq.) dropwise over 5 min. The reaction mixture was stirred under argon at RT for 3 h. After this time the reaction mixture was diluted with H_2O (30 mL) and CH_2Cl_2 (20 mL). The organic layer was washed with aqueous 1 M HCl (3 × 20 mL) and brine (2 × 20 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-(2-*tert*-butylphenyl)acetamide, **69d**, (1.26 g, 98%) as a colourless solid: R_f 0.3 (EtOAc/petroleum ether 50:50); mp 144-145 °C (from aqueous ethanol, 95% *v/v*) [lit.[25,26] 158-158.5 °C]; ^{1}H NMR (400 MHz; D_6-DMSO): δ 9.22 (1H, s, ArN*H*COCH$_3$), 7.42-7.35 (1H, m, Ar C*H*), 7.24-7.14 (2H, m, Ar C*H*), 7.05-6.99 (1H, m, Ar C*H*), 2.03 (3H, s, ArNHCOC*H*$_3$), 1.32 (9H, s, C(C*H*$_3$)$_3$); *m/z* (ES$^+$) 405 ([2M+Na]$^+$, 100%). The data are in good agreement with literature values.[27]

_N_-(2-Methyl-4-sulfamoylphenyl)acetamide (67b)

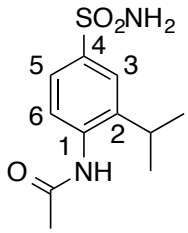


Chlorosulfonic acid (7.86 g, 4.49 mL, 67.50 mmol, 5.3 eq.) was added dropwise to *N*-(2-methylphenyl)acetamide, **69b**, (1.90 g, 12.74 mmol, 1.0 eq.) over 5 min at 0 °C. The reaction mixture was stirred under argon at 60 °C for 0.5 h. After this time, the reaction mixture was poured onto crushed ice (50 g) and extracted with $CHCl_3$ (3 × 100 mL). The organic layers were combined, washed with brine (100 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting pale yellow oil (2.41 g) that contained a mixture of 4-acetamido-3-methylbenzene-1-sulfonyl chloride, **70b**, and 3-acetamido-4-methylbenzene-1-sulfonyl chloride, **71b**, was used without further purification.

A solution of the pale yellow oil containing a mixture of 4-acetamido-3-methylbenzene-1-sulfonyl chloride, **70b**, and 3-acetamido-4-methylbenzene-1-sulfonyl chloride, **71b**, (2.41 g) in THF (30.2 mL) was added dropwise to ammonium hydroxide (30.2 mL, 25% solution in H_2O) over 5 min at 0 °C. The reaction mixture was stirred at RT for 16 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*). This involved two stages of crystallisation, in which *N*-(2-methyl-5-sulfamoylphenyl)acetamide, **72b**, was isolated in the initial crystallisation step and filtered. The mother liquor was concentrated *in vacuo* and the resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-(2-methyl-4-sulfamoylphenyl)acetamide, **67b**, (0.37 g, 13% over two steps) as a colourless solid: R_f 0.2 (EtOAc/petroleum ether 80:20); mp 201-203 °C (from aqueous ethanol, 95% *v/v*); ν_{max}/cm^{-1} (neat) 3352 (m), 3254 (m), 1739 (m), 1662 (m), 1583 (m), 1525 (m), 1439 (m), 1367 (m), 1309 (m), 1280 (m), 1196 (m), 1158 (s), 1125

(m), 1097 (m), 1034 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 9.48 (1H, s, ArN*H*COCH$_3$), 7.70 (1H, d, *J* 8.4, Ar C^{6}*H*), 7.65 (1H, d, *J* 2.0, Ar C^{3}*H*), 7.60 (1H, dd, *J* 8.4, 2.0, Ar C^{5}*H*), 7.25 (2H, s, ArSO$_2$N*H*$_2$), 2.28 (3H, s, ArC*H*$_3$), 2.11 (3H, s, NHCOC*H*$_3$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 168.7 (ArNH*C*OCH$_3$), 139.7 (*C*^{1}), 139.7 (*C*^{4}), 131.1 (*C*^{2}), 127.6 (*C*^{3}), 124.0 (*C*^{6}), 123.6 (*C*^{5}), 23.5 (ArNHCO*C*H$_3$), 18.0 (Ar*C*H$_3$); HRMS *m/z* (ES^{+}) (Found: [M+H]$^{+}$ 229.0631. $C_9H_{13}N_2O_3S$ requires M^{+} 229.0641); *m/z* (ES^{-}) 227 ([M−H]$^{-}$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 7.616 min, Purity: 98.78%. The data are in good agreement with the available literature values.[28]

***N*-[(2-Propan-2-yl)-4-sulfamoylphenyl)acetamide (67c)**

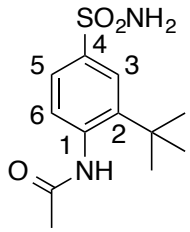


Chlorosulfonic acid (4.18 g, 2.39 mL, 35.88 mmol, 5.3 eq.) was added dropwise to *N*-[2-(propan-2-yl)phenyl]acetamide, **69c**, (1.20 g, 6.77 mmol, 1.0 eq.) over 5 min at 0 °C. The reaction mixture was stirred under argon at 60 °C for 0.5 h. After this time, the reaction mixture was poured onto crushed ice (100 g) and extracted with $CHCl_3$ (3 × 50 mL). The organic layers were combined, washed with brine (100 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting off-white solid (1.75 g), which contained 4-acetamido-3-(propan-2-yl)benzene-1-sulfonyl chloride, **70c**, and 3-acetamido-4-(propan-2-yl)benzene-1-sulfonyl chloride, **71c**, was used without further purification.

A solution of the off-white solid containing 4-acetamido-3-(propan-2-yl)benzene-1-sulfonyl chloride, **70c**, and 3-acetamido-4-(propan-2-yl)benzene-1-sulfonyl chloride, **71c**, (1.75 g) in THF (13.1 mL) was added dropwise to ammonium hydroxide (13.1 mL, 25% solution in H_2O) over

5 min at 0 °C. The reaction mixture was stirred at RT for 16 h. After this time, the reaction mixture was concentrated *in vacuo* and the residue was extracted with EtOAc (3 × 50 mL). The organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*). This involved two stages of crystallisation, in which *N*-[2-(propan-2-yl)-5-sulfamoylphenyl]acetamide, **72c**, was isolated in the initial crystallisation step and filtered. The mother liquor was concentrated *in vacuo* and the resulting residue was purified by crystallisation from MeOH to yield *N*-[2-(propan-2-yl)-5-sulfamoylphenyl]acetamide, **67c**, (0.03 g, 2% over two steps) as a colourless solid: R_f 0.3 (EtOAc/petroleum ether 80:20); mp 217-219 °C (from MeOH); ν_{max}/cm^{-1} (neat) 3317 (m), 1643 (m), 1569 (m), 1513 (m), 1481 (m), 1402 (m), 1368 (m), 1343 (s), 1294 (m), 1275 (m), 1162 (s), 1107 (m); ^{1}H NMR (500 MHz; CD_3OD): δ 7.89 (1H, d, *J* 2.1, Ar C^3*H*), 7.71 (1H, dd, *J* 8.3, 2.1, Ar C^5*H*), 7.50 (1H, d, *J* 8.3, Ar C^6*H*), 3.24 (1H, sp, *J* 6.9, ArC*H*$(CH_3)_2$), 2.18 (3H, s, ArNHCOC*H*$_3$), 1.25 (6H, d, *J* 8.3, ArCH(C*H*$_3)_2$); ^{13}C NMR (126 MHz; CD_3OD): δ 172.7 (ArNH*C*OCH$_3$), 145.6 (*C*^{2}), 143.0 (*C*^{4}), 139.1 (*C*^{1}), 128.3 (*C*^{6}), 125.1 (*C*^{3}), 125.0 (*C*^{5}), 29.1 (Ar*CH*$(CH_3)_2$), 23.4 (ArCH(*C*H$_3)_2$), 23.1 (ArNHCO*C*H$_3$); HRMS *m/z* (ES$^-$) (Found: $[M-H]^-$ 255.0817. $C_{11}H_{15}N_2O_3S$ requires M^- 255.0809); *m/z* (ES$^-$) 255 ($[M-H]^-$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 8.846 min, Purity: 99.26%.

_N_-(2-_tert_-Butyl-4-sulfamoylphenyl)acetamide (67d)



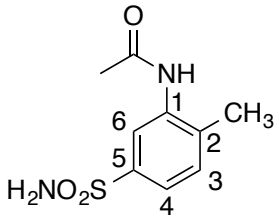
Chlorosulfonic acid (3.87 g, 2.21 mL, 33.25 mmol, 5.3 eq.) was added dropwise over 5 min to *N*-(2-*tert*-butylphenyl)acetamide, **69d**, (1.20 g, 6.27 mmol, 1.0 eq.). The reaction mixture was

stirred under argon at 60 °C for 0.5 h. After this time, the reaction mixture was poured onto crushed ice (50 g) and extracted with $CHCl_3$ (3 × 50 mL). The organic layers were combined, washed with brine (50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting off-white solid (1.80 g), which contained 4-acetamido-3-(*tert*-butyl)benzene-1-sulfonyl chloride, **70d**, and 3-acetamido-4-(*tert*-butyl)benzene-1-sulfonyl chloride, **71d**, was used without further purification.

A solution of the off-white solid containing 4-acetamido-3-(*tert*-butyl)benzene-1-sulfonyl chloride, **70d**, and 3-acetamido-4-(*tert*-butyl)benzene-1-sulfonyl chloride, **71d**, (1.80 g) in THF (12.9 mL) was added dropwise to ammonium hydroxide (12.9 mL, 25% solution in H_2O) over 5 min at 0 °C. The reaction mixture was stirred at RT for 16 h. After this time, the reaction mixture was concentrated *in vacuo* and the residue was extracted with EtOAc (2 × 50 mL). The organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*). This involved two stages of crystallisation, in which *N*-(2-tert-butyl-5-sulfamoylphenyl)acetamide, **72d**, was isolated in the initial crystallisation step and filtered. The resulting mother liquor was concentrated *in vacuo* and the resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-(2-tert-butyl-4-sulfamoylphenyl)acetamide, **67d**, (0.1 g, 6% over two steps) as a colourless solid: R_f 0.2 (EtOAc/petroleum ether 80:20); mp 225-228 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} (neat) 3348 (m), 3296 (m), 3094 (m), 2967 (m), 1667 (m), 1578 (m), 1514 (m), 1408 (m), 1367 (m), 1318 (m), 1267 (m), 1154 (s), 1115 (m), 1069 (m); ^{1}H NMR (500 MHz; CD_3OD): δ 8.03 (1H, d, *J* 2.1, Ar C^3*H*), 7.75 (1H, dd, *J* 8.2, 2.1, Ar C^5*H*), 7.30 (1H, d, *J* 8.2, Ar C^6*H*), 2.18 (3H, s, ArNHCOC*H*$_3$), 1.42 (9H, s, ArC(C*H*$_3$)$_3$); ^{13}C NMR (126 MHz; CD_3OD): δ 173.1 (ArNH*C*OCH$_3$), 149.1 (*C*^{4}), 143.7 (*C*^{2}), 140.2 (*C*^{1}), 133.2 (*C*^{6}), 126.2 (*C*^{3}), 125.5 (*C*^{5}), 36.4 (Ar*C*(CH$_3$)$_3$), 31.0 (ArC(*C*H$_3$)$_3$), 23.2 (ArNHCO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 271.1109. $C_{12}H_{19}N_2O_3S$ requires M$^+$ 271.1111);

m/z (ES^-) 269 ($[M-H]^-$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.128 min, Purity: 95.24%.

_N_-(2-Methyl-5-sulfamoylphenyl)acetamide (72b)

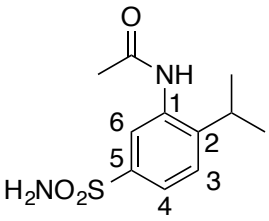


Chlorosulfonic acid (1.24 g, 0.71 mL, 10.66 mmol, 5.3 eq.) was added dropwise to *N*-(2-methylphenyl)acetamide, **69b**, (0.30 g, 2.01 mmol, 1.0 eq.) over 5 min at 0 °C. The reaction mixture was stirred under argon at 60 °C for 0.5 h. After this time, the reaction mixture was poured onto crushed ice (50 g) and extracted with EtOAc (2 x 50 mL). The organic layers were combined, washed with brine (50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting pale yellow oil (0.47 g) that contained a mixture of 4-acetamido-3-methylbenzene-1-sulfonyl chloride, **70b**, and 3-acetamido-4-methylbenzene-1-sulfonyl chloride, **71b**, was used without further purification.

A solution of the pale yellow oil containing a mixture of 4-acetamido-3-methylbenzene-1-sulfonyl chloride, **70b**, and 3-acetamido-4-methylbenzene-1-sulfonyl chloride, **71b**, (0.47 g) in THF (7.8 mL) was added dropwise to ammonium hydroxide (7.8 mL, 25% solution in H_2O) at over 5 min 0 °C. The reaction mixture was stirred at RT for 16 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-(2-methyl-5-sulfamoylphenyl)acetamide, **72b**, (0.15 g, 33% over two steps) as a colourless solid: R_f 0.1 (EtOAc/petroleum ether 80:20); mp 230-234 °C (from aqueous ethanol, 95% *v/v*); ν_{max}/cm^{-1} (neat) 3262 (m), 1626 (m), 1578 (m), 1530 (m), 1403 (m), 1371 (m), 1330 (m), 1302 (m), 1195 (m), 1149 (s), 1132 (m), 1087 (m), 1041 (m);

^{1}H NMR (400 MHz; D_6-DMSO): δ 9.47 (1H, s, N*H*), 7.95 (1H, d, *J* 1.3, Ar C^{6}*H*), 7.50 (1H, dd, *J* 8.0, 1.3, Ar C^{4}*H*), 7.38 (1H, d, *J* 8.0, Ar C^{3}*H*), 7.31 (2H, s, $ArSO_2NH_2$), 2.26 (3H, s, $ArCH_3$), 2.09 (3H, s, $ArNHCOCH_3$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 168.5 ($ArNHCOCH_3$), 141.9 (*C*^{5}), 136.8 (*C*^{1}), 135.3 (*C*^{2}), 130.7 (*C*^{3}), 121.9 (*C*^{4}), 121.8 (*C*^{6}), 23.3 ($ArNHCOCH_3$), 17.9 ($ArCH_3$); HRMS *m/z* (ES^{+}) (Found: $[M+Na]^+$ 251.0461. $C_9H_{12}N_2NaO_3S$ requires M^+ 251.0461); *m/z* (ES^{-}) 227 ($[M-H]^-$, 100%).

_N_-[2-(Propan-2-yl)-5-sulfamoylphenyl]acetamide (72c)

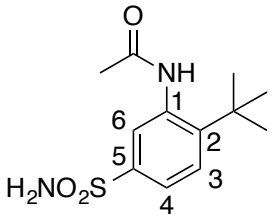


Chlorosulfonic acid (2.84 g, 1.62 mL, 24.4 mmol, 5.3 eq.) was added dropwise to *N*-[2-(propan-2-yl)phenyl]acetamide, **69c**, (0.82 g, 4.60 mmol, 1.0 eq.) over 5 min at 0 °C. The reaction mixture was stirred under argon at 60 °C for 1 h. After this time, the reaction mixture was poured onto crushed ice (50 g) and extracted with $CHCl_3$ (2 x 50 mL). The organic layers were combined, washed with brine (50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting off-white solid (0.96 g), which contained 4-acetamido-3-(propan-2-yl)benzene-1-sulfonyl chloride, **70c**, and 3-acetamido-4-(propan-2-yl)benzene-1-sulfonyl chloride, **71c**, was used without further purification.

A solution of the off-white solid containing 4-acetamido-3-(propan-2-yl)benzene-1-sulfonyl chloride, **70c**, and 3-acetamido-4-(propan-2-yl)benzene-1-sulfonyl chloride, **71c**, (0.96 g) in THF (7.19 mL) was added dropwise to ammonium hydroxide (7.19 mL, 25% solution in H_2O) over 5 min at 0 °C. The reaction mixture was stirred at RT for 2.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the residue was partitioned between H_2O (25 mL) and

EtOAc (50 mL). The aqueous layer was extracted with EtOAc (2 × 25 mL), the organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-[2-(propan-2-yl)-5-sulfamoylphenyl]acetamide, **72c**, (0.32 g, 27% over two steps) as a colourless solid: R_f 0.1 (EtOAc/petroleum ether 60:40); mp 193-194 °C (from aqueous ethanol, 95% *v/v*); ν_{max}/cm^{-1} (neat) 3306 (m), 2965 (m), 1668 (m), 1511 (m), 1410 (m), 1372 (m), 1322 (m), 1279 (m), 1195 (m), 1168 (s), 1113 (m), 1052 (m); ^{1}H NMR (500 MHz; CD_3OD): δ 7.82 (1H, d, *J* 2.0, Ar C^6*H*), 7.76 (1H, dd, *J* 8.3, 2.0, Ar C^4*H*), 7.53 (1H, d, *J* 8.3, Ar C^3*H*), 3.21 (1H, sp, *J* 6.9, ArC*H*$(CH_3)_2$), 2.18 (3H, s, ArNHCOC*H*$_3$), 1.23 (6H, d, *J* 6.9, ArCH(C*H*$_3)_2$); ^{13}C NMR (126 MHz; CD_3OD): δ 172.9 (ArNH*C*OCH_3), 149.8 (*C*^{2}), 142.6 (*C*^{5}), 135.9 (*C*^{1}), 127.8 (*C*^{3}), 126.2 (*C*^{6}), 125.5 (*C*^{4}), 29.2 (Ar*C*H$(CH_3)_2$), 23.4 (ArCH(*C*H$_3)_2$), 22.9 (ArNHCO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: $[M+Na]^+$ 279.0783. $C_{11}H_{16}N_2NaO_3S$ requires M^+ 279.0774); *m/z* (ES$^-$) 255 ($[M-H]^-$, 100%); Anal. Calcd. (Found C, 51.58; H, 6.24; N, 10.82. $C_{11}H_{16}N_2O_3S$ requires C, 51.54; H, 6.29; N, 10.93).

_N_-(2-_tert_-Butyl-5-sulfamoylphenyl)acetamide (72d)

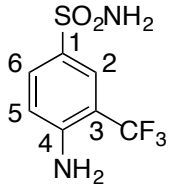


Chlorosulfonic acid (0.48 g, 0.28 mL, 4.15 mmol, 5.3 eq.) was added dropwise over 5 min to *N*-(2-*tert*-butylphenyl)acetamide, **69d**, (0.15 g, 0.78 mmol, 1.0 eq.). The reaction mixture was stirred under argon at 60 °C for 0.5 h. After this time, the reaction mixture was poured onto crushed ice (30 g) and extracted with $CHCl_3$ (2 x 25 mL). The organic layers were combined, washed with brine (25 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting off-white solid (0.16 g), which contained 4-acetamido-3-*tert*-butylbenzene-1-sulfonyl

chloride, **70d**, and 5-acetamido-3-*tert*-butylbenzene-1-sulfonyl chloride, **71d**, was used without further purification.

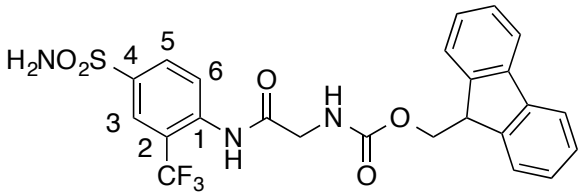
A solution of the off-white solid containing 4-acetamido-3-*tert*-butylbenzene-1-sulfonyl chloride, **70d**, and 5-acetamido-3-*tert*-butylbenzene-1-sulfonyl chloride, **71d**, (0.11 g) in THF (0.75 mL) was added dropwise to ammonium hydroxide (0.75 mL, 25% solution in H_2O) over 5 min at 0 °C. The reaction mixture was stirred at RT for 1 h. After this time, the reaction mixture was concentrated *in vacuo* and the residue was partitioned between H_2O (25 mL) and EtOAc (50 mL). The aqueous layer was extracted with EtOAc (2 × 25 mL), the organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was purified by crystallisation from aqueous ethanol (95% *v/v*) to yield *N*-(2-*tert*-butyl-5-sulfamoylphenyl)acetamide, **72d**, (29.3 mg, 20% over two steps) as a colourless solid: R_f 0.1 (EtOAc/petroleum ether 60:40); mp 237-240 °C (from aqueous ethanol, 95% *v/v*); ν_{max}/cm^{-1} (neat) 3320 (m), 2962 (m), 1667 (m), 1443 (m), 1401 (m), 1370 (m), 1328 (s), 1278 (m), 1226 (m), 1171 (s), 1122 (m), 1070 (m); ^{1}H NMR (500 MHz; CD_3OD): δ 7.75 (1H, dd, *J* 8.5, 2.1, Ar C^4*H*), 7.65 (1H, d, *J* 8.5, Ar C^3*H*), 7.62 (1H, d, *J* 2.1, Ar C^6*H*), 2.18 (3H, s, ArNHCOC*H*$_3$), 1.41 (9H, s, ArC(C*H*$_3$)$_3$); ^{13}C NMR (126 MHz; CD_3OD): 173.3 (ArNH*C*OCH$_3$), 152.8 (*C*^{2}), 143.2 (*C*^{5}), 137.2 (*C*^{1}), 130.4 (*C*^{6}), 129.1 (*C*^{3}), 126.1 (*C*^{4}), 36.4 (Ar*C*(CH$_3$)$_3$), 31.0 (ArC(*C*H$_3$)$_3$), 23.1 (ArNHCO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+Na]$^+$ 293.0937. $C_{12}H_{18}N_2NaO_3S$ requires M$^+$ 293.0930); *m/z* (ES$^-$) 269 ([M−H]$^-$, 100%); Anal. Calcd. (Found C, 53.26; H, 6.69; N, 10.24. $C_{12}H_{18}N_2O_3S$ requires C, 53.31; H, 6.71; N, 10.36).

4-Amino-3-(trifluoromethyl)benzene-1-sulfonamide (74)



A mixture of 4-nitro-3-(trifluoromethyl)benzene-1-sulfonamide, **73**, (0.95 g, 3.52 mmol, 1.0 eq.) and palladium on carbon (0.19 g, 0.18 mmol, 0.05 eq., 10% wt) in MeOH (18 mL) was stirred under hydrogen at atmospheric pressure for 5 h at RT. After this time, the reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo*, to afford 4-amino-3-(trifluoromethyl)benzene-1-sulfonamide, **74**, as a colourless solid (0.84 g, 99%): R_f 0.3 (EtOAc/petroleum ether 50:50); mp 143-145 °C (from MeOH); ν_{max}/cm^{-1} (neat) 3351 (m), 3260 (m), 2160 (m), 1651 (m), 1614 (m), 1574 (m), 1499 (m), 1437 (m), 1311 (m), 1286 (m), 1150 (s), 1098 (s), 1049 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 7.75 (1H, d, *J* 2.1, Ar C^2*H*), 7.65 (1H, dd, *J* 8.8, 2.1, Ar C^6*H*), 7.12 (2H, s, $ArSO_2NH_2$), 6.89 (1H, d, *J* 8.8, Ar C^5*H*), 6.37 (2H, s, $ArNH_2$); ^{19}F NMR (377 MHz; D_6-DMSO): δ −62.3 (ArCF_3); ^{13}C NMR (126 MHz; D_6-DMSO): δ 148.9 (*C*^{1}), 130.6 (*C*^{6}), 129.9 (*C*^{4}), 124.8 (q, *J* 5.7, *C*^{2}), 124.3 (q, *J* 272.3, *C*F_3), 116.4 (*C*^{5}), 109.0 (q, *J* 30.5, *C*^{3}); HRMS *m/z* (ES$^+$) (Found: $[M+Na]^+$ 263.0066. $C_7H_7F_3N_2NaO_2S$ requires M^+ 263.0073); *m/z* (ES$^+$) 263 ($[M+Na]^+$, 100%); Anal. Calcd. (Found C, 34.91; H, 2.86; N, 11.51. $C_7H_7F_3N_2O_2S$ requires C, 35.00; H, 2.94; N, 11.66).

(9*H*-Fluoren-9-yl)methyl (2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethyl)carbamate (75)



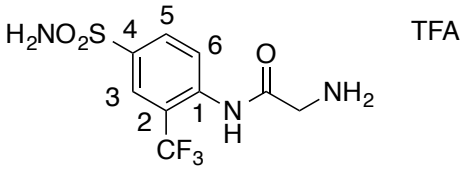
To Fmoc-glycine (0.40 g, 1.35 mmol, 1.0 eq.) suspended in anhydrous CH_2Cl_2 (46 mL) was added DMF (0.08 mL), followed by the dropwise addition of thionyl chloride (1.60 g, 0.98 mL,

13.45 mmol, 10.0 eq.) over 5 min. The reaction mixture was stirred under argon at RT for 3 h. After this time, the reaction mixture was concentrated *in vacuo*, azeotroping with anhydrous toluene three times before the resulting residue containing Fmoc-glycyl chloride was dried under high vacuum.

To a stirred solution of 4-amino-3-(trifluoromethyl)benzene-1-sulfonamide, **74**, (0.10 g, 0.42 mmol) and anhydrous NEt_3 (42.2 mg, 58.1 µL, 0.42 mmol) in anhydrous MeCN (3.0 mL) was added a solution of Fmoc-glycyl chloride (0.39 g) in anhydrous MeCN (7.0 mL) dropwise over 5 min at 0 °C. The reaction mixture was stirred at RT for 15.5 h under argon. After this time, the reaction mixture was quenched with MeOH and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with MeOH and $CHCl_3$ (1:99 → 10:95; MeOH : $CHCl_3$) to afford (9*H*-fluoren-9-yl)methyl (2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethyl)carbamate, **75**, as an off-white solid (71.1 mg, 10% over two steps): R_f 0.2 (MeOH/$CHCl_3$ 5:95); mp 219-221 °C (from MeOH/$CHCl_3$); ν_{max}/cm^{-1} (neat) 3256 (m), 1739 (m), 1663 (m), 1538 (m), 1434 (m), 1330 (m), 1312 (m), 1283 (m), 1254 (m), 1166 (s), 1126 (s), 1059 (s); ^{1}H NMR (500 MHz; D_6-DMSO): δ 9.70 (1H, s, ArN*H*CO), 8.12 (1H, d, *J* 2.0, Ar C^3*H*), 8.09 (1H, dd, *J* 8.5, 2.0, Ar C^5*H*), 7.90 (2H, d, *J* 7.5, Fmoc Ar C*H*), 7.86 (1H, d, *J* 8.5, Ar C^6*H*), 7.76 (1H, t, *J* 5.9, N*H*CH_2), 7.72 (2H, d, *J* 7.4, Fmoc Ar C*H*), 7.56 (2H, s, $ArSO_2NH_2$), 7.42 (2H, dd, *J* 8.4, 7.5, Fmoc Ar C*H*), 7.33 (2H, dd, *J* 8.4, 7.4, Fmoc Ar C*H*), 4.33 (2H, d, *J* 6.7, Fmoc CHC*H*$_2$O), 4.24 (1H, t, *J* 6.7, Fmoc C*H*CH_2O), 3.88 (2H, d, *J* 5.9, NHC*H*$_2$); ^{19}F NMR (470 MHz; D_6-DMSO): δ −59.7 (ArC*F*$_3$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 169.3 (O*C*ONH), 156.7 (CH_2*C*ONH), 143.8 (Ar *C*), 141.4 (Ar *C*), 138.1 (Ar *C*), 130.3 (*C*^{5}), 129.2 (*C*^{6}), 127.7 (Fmoc Ar *C*H), 127.1 (Fmoc Ar *C*H), 125.2 (Fmoc Ar *C*H), 124.1 (q, *J* 4.8, *C*^{3}), 123.3 (q, *J* 30.0, *C*^{2}), 122.9 (q, *J* 274.0, *C*F_3), 120.2 (Fmoc Ar *C*H), 119.6 (Ar *C*), 65.8 (Fmoc CH*C*H_2O), 46.6 (Fmoc *C*HCH_2O), 43.9 (NH*C*H_2); HRMS *m/z* (ES$^+$) (Found: $[M+Na]^+$ 542.0971. $C_{24}H_{20}F_3N_3NaO_5S$

requires M^+, 542.0968); *m/z* (ES^-) 518 ($[M-H]^-$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 12.41 min, Purity: 96.73%.

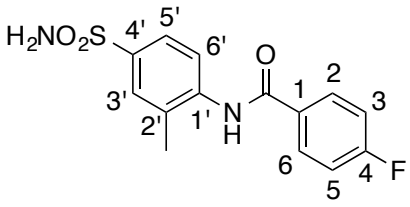
2-Oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethan-1-aminium 2,2,2-trifluoroacetate (76)



To a stirred solution of (9*H*-fluoren-9-yl)methyl (2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethyl)carbamate, **75**, (0.07 g, 0.13 mmol, 1.0 eq.) in anhydrous DMF (1.4 mL) was added piperidine (0.28 mL) dropwise over 5 min. The reaction mixture was stirred under argon at RT for 18 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with MeOH and $CHCl_3$ (1:99 → 10:90; MeOH : $CHCl_3$) to afford a slightly impure colourless residue containing 2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethan-1-aminium, **76**, (0.03 g). 0.01 g of the slightly impure residue containing 2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethan-1-aminium, **76**, was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA), to afford 2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethan-1-aminium 2,2,2-trifluoroacetate, **76**, as a colourless solid (5.81 mg, 10%): R_f 0.03 (MeOH/$CHCl_3$ 5:95); ν_{max}/cm^{-1} (neat): 3256 (m), 1739 (m), 1663 (m), 1538 (m), 1434 (m), 1330 (m), 1312 (m), 1283 (m), 1254 (m), 1166 (s), 1126 (s), 1059 (m); ^{1}H NMR (500 MHz; CD_3OD): δ 8.24 (1H, d, *J* 1.9, Ar C^3*H*), 8.16 (1H, dd, *J* 8.5, 1.9, Ar C^5*H*), 7.93 (1H, d, 8.5, Ar C^6*H*), 3.98 (2H, s, C*H*$_2$); ^{19}F NMR (470 MHz; CD_3OD): δ −62.6 (s, ArC*F*$_3$), −76.9 (s, OCOC*F*$_3$); ^{13}C NMR (126 MHz; CD_3OD): δ 167.3 (ArNH*C*O), 163.0 (q, *J* 34.3, O*C*OCF$_3$), 143.8 (*C*^{1}), 138.5 (*C*^{4}), 131.7 (*C*^{5}), 130.6 (*C*^{6}), 125.9 (q, *J* 5.5, *C*^{3}), 124.4 (q, *J* 273.0, Ar*C*F$_3$) 42.2 (*C*H$_2$), (*C*^{2} and OCO*C*F$_3$ are

not observed); HRMS *m/z* (ES^+) (Found: $[M+H]^+$ 298.0465. $C_9H_{11}F_3N_3O_3S$ requires M^+ 298.0468); *m/z* (ES^+) 298 ($[M+H]^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA; 10 min) Ret. Time = 7.00 min, Purity: 96.52%.

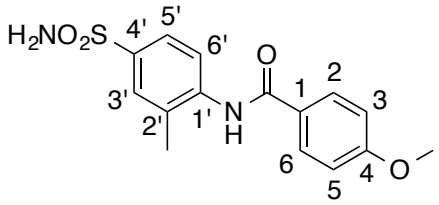
4-Fluoro-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide (78a)



To a stirred solution of 4'-amino-3'-(methyl)benzene-1'-sulfonamide, **77**, (0.10 g, 0.54 mmol, 1.0 eq.) and anhydrous NEt_3 (0.08 g, 0.11 mL, 0.81 mmol, 1.5 eq.) at 0 °C in anhydrous THF (2.0 mL) was added 4-fluorobenzoyl chloride (0.09 g, 0.07 mL, 0.59 mmol, 1.1 eq.) dropwise over 5 min. The reaction mixture was stirred under nitrogen at RT for 22 h. After this time the reaction mixture was concentrated *in vacuo* and the resulting residue was partitioned between H_2O (25 mL) and EtOAc (25 mL). The organic layer was collected and washed with brine (25 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (12:88 → 100:0; EtOAc : petroleum ether) to afford 4-fluoro-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide, **78a**, (0.07 g, 45%) as a colourless solid: R_f 0.2 (EtOAc/petroleum ether 50:50); mp 234-237 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} (neat): 3301 (w), 2160 (w), 1659 (m), 1603 (w), 1578 (w), 1530 (m), 1505 (m), 1317 (m), 1283 (w), 1241 (m), 1194 (w), 1159 (s), 1136 (m), 1111 (m); ^{1}H NMR (500 MHz; CD_3OD): δ 8.07-8.02 (2H, m, Ar C^2H & Ar C^6H), 7.84 (1H, d, *J* 2.0, $C^{3'}H$), 7.77 (1H, dd, *J* 8.3, 2.0, $C^{5'}H$), 7.59 (1H, d, *J* 8.3, $C^{6'}H$), 7.30-7.24 (2H, m, Ar C^3H & Ar C^5H), 2.38 (s, 3H, ArCH_3); ^{19}F NMR (376.6 MHz; CD_3OD): δ −109.7 (Ar C^4F); ^{13}C NMR (126 MHz; CD_3OD): δ 167.8 (NH*C*O), 166.6 (d, *J* 251.1, C^4), 142.7 ($C^{4'}$), 140.8 ($C^{1'}$), 135.9 ($C^{2'}$), 131.8 (d, *J* 3.4, C^1), 131.5 (d, *J* 9.6, C^2 & C^6), 129.5 ($C^{3'}$), 127.6 ($C^{6'}$), 125.4 ($C^{5'}$), 116.6 (d, *J* 22.6, C^3 & C^5), 18.3 (Ar*C*H_3); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 331.0514.

$C_{14}H_{13}FN_2NaO_3S$ requires M^+ 331.0523); *m/z* (ES^-) 307 ($[M-H]^-$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 10.550 min, Purity: 96.80%.

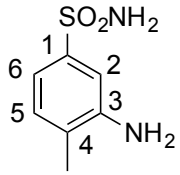
4-Methoxy-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide (78b)



To a stirred solution of 4'-amino-3'-(methyl)benzene-1'-sulfonamide, **77**, (0.06 g, 0.33 mmol, 1.0 eq.) and anhydrous NEt_3 (0.05 g, 0.07 mL, 0.50 mmol, 1.5 eq.) at 0 °C in anhydrous THF (1.5 mL) was added a solution of 4-methoxybenzoyl chloride (0.06 g, 0.37 mmol, 1.1 eq.) in anhydrous THF (1.0 mL) dropwise over 5 min. The reaction mixture was stirred under nitrogen at RT for 22 h. After this time the reaction mixture was concentrated *in vacuo* and the resulting residue was partitioned between H_2O (25 mL) and EtOAc (25 mL). The organic layer was collected and washed with brine (25 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (12:88 → 100:0; EtOAc : petroleum ether) to afford 4-methoxy-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide, **78b**, (0.05 g, 48%) as a colourless solid: R_f 0.2 (EtOAc/petroleum ether 50:50); mp 243-247 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} (neat): 3253 (w), 1644 (m), 1607 (m), 1507 (m), 1398 (w), 1309 (m), 1259 (m), 1156 (s), 1137 (m), 1024 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 9.87 (1H, s, N*H*CO), 8.00-7.96 (2H, m, Ar C^2H & Ar C^6H), 7.72 (1H, d, *J* 1.9, Ar $C^{3'}H$), 7.66 (1H, dd, *J* 8.3, 1.9, Ar $C^{5'}H$), 7.58 (1H, d, *J* 8.3, Ar $C^{6'}H$), 7.31 (2H, s, $ArSO_2NH_2$), 7.10-7.05 (2H, m, Ar C^3H & Ar C^5H), 3.84 (3H, s, $ArCH_3$), 2.31 (s, 3H, $ArOCH_3$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 164.8 (NH*C*O), 162.1 (C^4), 140.8 ($C^{4'}$), 139.8 ($C^{1'}$), 133.7 ($C^{2'}$), 129.7 (C^2 & C^6), 127.6 ($C^{3'}$), 126.2 (C^1), 126.2 ($C^{6'}$), 123.6 ($C^{5'}$),

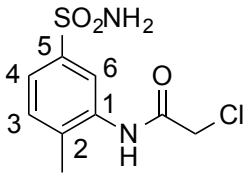
113.7 (C^3 & C^5), 55.5 (ArCH$_3$), 18.0 (ArOCH$_3$); HRMS *m/z* (ES$^-$) (Found: [M−H]$^-$ 319.0756. $C_{15}H_{15}N_2O_4S$ requires M^+ 319.0758); *m/z* (ES$^-$) 319 ([M−H]$^-$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 10.392 min, Purity: 95.09%.

3-Amino-4-methylbenzene-1-sulfonamide (80)



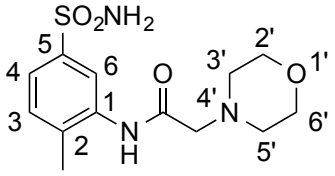
A mixture of 4-methyl-3-nitrobenzene-1-sulfonamide, **79**, (0.50 g, 2.31 mmol, 1.0 eq.) and palladium on carbon (0.12 g, 0.12 mmol, 0.05 eq., 10% wt) in MeOH (12 mL) was stirred under hydrogen at atmospheric pressure for 63.5 h at RT. After this time, the reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo*, to afford 3-amino-4-methylbenzene-1-sulfonamide, **80**, as a colourless solid (0.43 g, 99%): R_f 0.2 (EtOAc/petroleum ether 50:50); mp 170-174 °C (from EtOAc) [lit.[28] 177-179 °C]; ν_{max}/cm^{-1} (neat) 3291 (w), 2970 (w), 1739 (s), 1630 (w), 1453 (m), 1368 (m), 1313 (m), 1229 (m), 1216 (m), 1151 (m), 1105 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 7.08-7.04 (4H, m, Ar C^5H, Ar C^2H, ArSO$_2$NH_2), 6.89 (1H, dd, *J* 7.8, 1.8, Ar C^6H), 5.26 (2H, s, ArNH_2), 2.08 (3H, s, ArCH_3); ^{13}C NMR (126 MHz; D_6-DMSO): δ 146.9 (C^3), 142.3 (C^1), 130.0 (C^5), 124.7 (C^4), 112.8 (C^6), 110.4 (C^2), 17.4 (ArCH$_3$); HRMS *m/z* (ES$^+$) (Found: [M+Na]$^+$ 209.0356. $C_7H_{10}N_2NaO_2S$ requires M^+ 209.0355); *m/z* (ES$^+$) 187 ([M+H]$^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 97:3 → 3:97; H_2O : MeCN, with 0.25 mM NEt_3 and acetic acid) Ret. Time = 8.317 min, Purity: 98.78%. The data are in good agreement with the available literature values.[29]

2-Chloro-*N*-(2-methyl-5-sulfamoylphenyl)acetamide (81)



To a stirred solution of 3-amino-4-methylbenzene-1-sulfonamide, **80**, (0.04 g, 0.23 mmol, 1.0 eq.) in aqueous 2 M NaOH (0.47 mL) was added chloroacetyl chloride (0.03 g, 0.02 mL, 0.30 mmol, 1.3 eq.) as a solution in CH_2Cl_2 (0.47 mL). The reaction mixture was stirred at RT for 0.5 h and after this time the layers were separated and the aqueous phase extracted with EtOAc (2 × 50 mL). The combined organic layers were washed with aqueous 1 M HCl (50 mL), aqueous saturated $NaHCO_3$ (50 mL), dried with anhydrous Na_2SO_4, and concentrated *in vacuo* to yield 2-chloro-*N*-(2-methyl-5-sulfamoylphenyl)acetamide, **81**, (21.5 mg, 36%) as a colourless solid: R_f 0.3 (MeOH/$CHCl_3$ 5:95); mp 175-176 °C (from EtOAc); ν_{max}/cm^{-1} (neat) 3259 (m), 1660 (m), 1536 (m), 1402 (m), 1337 (m), 1251 (m), 1149 (s), 1127 (m), 1085 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 9.85 (1H, s, ArN*H*$COCH_2Cl$), 7.92 (1H, d, *J* 1.8, Ar C^6H), 7.56 (1H, dd, *J* 8.0, 1.8, Ar C^4H), 7.43 (1H, d, *J* 8.0, Ar C^3H), 7.35 (2H, s, $ArSO_2NH_2$), 4.34 (2H, s, $ArNHCOCH_2Cl$), 2.27 (3H, s, $ArCH_3$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 165.3 ($ArNHCOCH_2Cl$), 142.1 (C^5), 136.0 (C^1), 135.9 (C^2), 131.0 (C^3), 122.8 (C^4), 122.0 (C^6), 43.1 ($ArNHCOCH_2Cl$), 17.8 ($ArCH_3$); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 285.0078. $C_9H_{11}ClN_2NaO_3S$ requires M^+ 285.0071); *m/z* (ES^-) 261 ($[M-H]^-$, 100%).

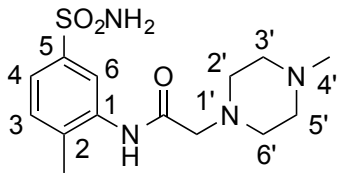
***N*-(2-Methyl-5-sulfamoylphenyl)-2-(morpholin-4'-yl)acetamide (82)**



To a stirred suspension of 2-chloro-*N*-(2-methyl-5-sulfamoylphenyl)acetamide, **81**, (0.17 g, 0.64 mmol, 1.0 eq.) in MeOH (44.5 mL) at 80 °C was added

morpholine (2.57 g, 2.55 mL, 29.5 mmol, 46.1 eq.) in ethanol (14.8 mL) dropwise over 5 min. The reaction mixture was heated at 80 °C for 6 h, cooled to RT, and diluted with H_2O (25 mL). The aqueous layer was extracted with EtOAc (2 × 50 mL), and the organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with MeOH and $CHCl_3$ (1:99 → 5:95; MeOH : $CHCl_3$) to afford *N*-(2-methyl-5-sulfamoylphenyl)-2-(morpholin-4'-yl)acetamide, **82**, (0.09 g, 47%) as a colourless solid: R_f 0.3 (MeOH/$CHCl_3$ 5:95); mp 201-204 °C (from MeOH/$CHCl_3$); ν_{max}/cm^{-1} (neat) 2970 (m), 1739 (s), 1669 (m), 1576 (m), 1523 (m), 1447 (m), 1366 (m), 1302 (m), 1217 (m), 1158 (m), 1107 (m), 1013 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 9.57 (1H, s, ArN*H*CO), 8.26 (1H, d, *J* 1.7, Ar C^6*H*), 7.50 (1H, dd, *J* 8.0, 1.7, Ar C^4*H*), 7.41 (1H, d, *J* 8.0, Ar C^3*H*), 7.32 (2H, s, $ArSO_2NH_2$), 3.66 (4H, t, *J* 4.3, C$^{2'}$*H*$_2$ & C$^{6'}$*H*$_2$), 3.17 (2H, s, NHCOC*H*$_2$N), 2.57 (4H, t, *J* 4.3, C$^{3'}$*H*$_2$ & C$^{5'}$*H*$_2$), 2.30 (3H, s, ArC*H*$_3$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 168.2 (ArNH*C*O), 142.2 (*C*^{5}), 136.3 (*C*^{1}), 133.4 (*C*^{2}), 130.6 (*C*^{3}), 121.6 (*C*^{4}), 119.4 (*C*^{6}), 66.3 (C$^{2'}$ & C$^{6'}$), 61.7 (NHCO*C*H$_2$N), 53.2 (C$^{3'}$ & C$^{5'}$), 17.6 (Ar*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 314.1171. $C_{13}H_{20}N_3O_4S$ requires M$^+$ 314.1169); *m/z* (ES$^-$) 312 ([M−H]$^-$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 7.445 min, Purity: 99.14%.

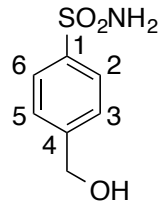
_N_-(2-Methyl-5-sulfamoylphenyl)-2-(4'-methylpiperazin-1'-yl)acetamide (83)



To a stirred suspension of 2-chloro-*N*-(2-methyl-5-sulfamoylphenyl)acetamide, **81**, (35.6 mg, 0.14 mmol, 1.0 eq.) in MeOH (4.2 mL) at 80 °C was added 1-methylpiperazine (0.64 g, 0.71 mL, 6.39 mmol, 47.2 eq.) in ethanol (1.4 mL) dropwise over 5 min. The mixture was heated at 80 °C for 6 h, cooled to RT, and diluted with H_2O (25 mL). The aqueous layer was extracted with

EtOAc (2 × 50 mL), and the organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with MeOH and $CHCl_3$ (2:98 → 20:80; MeOH : $CHCl_3$) to afford *N*-(2-methyl-5-sulfamoylphenyl)-2-(4'-methylpiperazin-1'-yl)acetamide, **83**, (22.4 mg, 51%) as a colourless solid: R_f 0.1 (MeOH/$CHCl_3$ 10:90); mp 273-276 °C (from MeOH/$CHCl_3$); ν_{max}/cm^{-1} (neat) 2802 (m), 1673 (m), 1578 (m), 1521 (m), 1448 (m), 1410 (m), 1371 (m), 1327 (s), 1302 (m), 1133 (m), 1082 (m), 1009 (m); ^1H NMR (500 MHz; D_6-DMSO): δ 9.55 (s, 1H, ArN*H*COCH$_2$N), 8.34 (1H, d, *J* 1.8, Ar C^6*H*), 7.49 (1H, dd, *J* 8.0, 1.8, Ar C^4*H*), 7.41 (1H, d, *J* 8.0, Ar C^3*H*), 7.32 (2H, s, ArSO$_2$N*H$_2$*), 3.16 (2H, s, ArNHCOC*H$_2$*N), 2.58 (4H, s, C$^{2'}$*H$_2$* & C$^{6'}$*H$_2$*), 2.40 (4H, s, C$^{3'}$*H$_2$* & C$^{5'}$*H$_2$*), 2.30 (3H, s, ArC*H$_3$*), 2.18 (3H, s, NC*H$_3$*); ^{13}C NMR (126 MHz; D_6-DMSO): δ 168.3 (ArNH*C*OCH$_2$N), 142.3 (*C*^{5}), 136.3 (*C*^{1}), 132.6 (*C*^{2}), 130.6 (*C*^{3}), 121.4 (*C*^{4}), 118.7 (*C*^{6}), 61.4 (ArNHCO*C*H$_2$N), 54.9 (C$^{3'}$ & C$^{5'}$), 52.8 (C$^{2'}$ & C$^{6'}$), 45.7 (N*C*H$_3$), 17.5 (Ar*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: $[M+H]^+$ 327.1494. $C_{14}H_{23}N_4O_3S$ requires M^+ 327.1485); *m/z* (ES$^+$) 327 ($[M+H]^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 7.393 min, Purity: 95.99%.

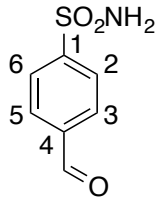
4-(Hydroxymethyl)benzene-1-sulfonamide (88)



Borane•THF (8.93 g, 9.94 mL, 103.88 mmol, 41.8 eq., 1 M in THF) was added dropwise over 0.5 h to a stirred solution of 4-sulfamoylbenzoic acid, **87**, (0.50 g, 2.49 mmol, 1.0 eq.) in anhydrous THF (40 mL) under nitrogen at 0 °C. The reaction mixture was stirred for 10 min at 0 °C before being stirred under nitrogen at RT for 3.5 h. After this time, the reaction mixture was cooled to 0 °C and quenched by the dropwise addition of MeOH (20 mL) over 5 min. After 15 min, aqueous 2 M HCl (25 mL) was added and the mixture was stirred at RT for 0.5 h. The mixture was then

concentrated *in vacuo* and the resulting residue was partitioned between brine (25 mL) and EtOAc (50 mL). The aqueous layer was extracted with EtOAc (4 × 25 mL), and the organic layers were combined, dried over anhydrous $MgSO_4$, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (30:70; EtOAc : petroleum ether) to afford 4-(hydroxymethyl)benzene-1-sulfonamide, **88**, (0.30 g, 64%) as a colourless solid: R_f 0.1 (EtOAc/petroleum ether 50:50); mp 123-125 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} (neat) 3464 (m), 3330 (m), 3233 (m), 1601 (w), 1545 (m), 1456 (m), 1399 (m), 1309 (s), 1201 (m), 1150 (s), 1091 (m), 1040 (s); ^{1}H NMR (400 MHz; D_6-DMSO): δ 7.77 (2H, d, *J* 8.3, Ar C^2*H* & Ar C^6*H*), 7.48 (2H, d, *J* 8.3, Ar C^3*H* & Ar C^5*H*), 7.31 (2H, s, $ArSO_2NH_2$), 5.39 (1H, t, *J* 5.7, $ArCH_2OH$), 4.57 (2H, d, *J* 5.7, $ArCH_2OH$); ^{13}C NMR (101 MHz; D_6-DMSO): δ 146.7 (*C*^{1}), 142.5 (*C*^{4}), 126.5 (*C*^{3} & *C*^{5}), 125.5 (*C*^{2} & *C*^{6}), 62.3 ($ArCH_2OH$); HRMS *m/z* (ES^-) (Found: $[M-H]^-$ 186.0227. $C_7H_8NO_3S$ requires M^- 186.0230); *m/z* (ES^-) 186 ([M−H], 100%); Anal. Calcd. (Found C, 44.88; H, 4.92; N, 7.36. $C_7H_9NO_3S$ requires C, 44.91; H, 4.85; N, 7.48). The data are in good agreement with the available literature values.[30]

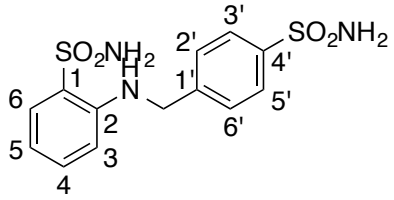
4-Formylbenzene-1-sulfonamide (89)



To a stirred solution of 4-(hydroxymethyl)benzene-1-sulfonamide, **88**, (0.24 g, 1.28 mmol, 1.0 eq.) in anhydrous MeCN (5 mL) was added manganese dioxide (0.51 g, 5.87 mmol, 4.6 eq.). The reaction mixture was stirred under argon at RT for 48 h. After this time the solvent was removed *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (30:70; EtOAc : petroleum ether) to afford 4-formylbenzene-1-sulfonamide, **89**, (0.16 g, 67%) as a colourless solid: R_f 0.4 (EtOAc/petroleum ether 50:50); mp 117-120 °C (from EtOAc/petroleum ether)

[lit.[30,31] 118-120 °C H_2O]; ν_{max}/cm^{-1} (neat) 3333 (m), 3251 (m), 1711 (m), 1599 (w), 1550 (m), 1378 (m), 1326 (s), 1298 (m), 1203 (s), 1151 (s), 1091 (m); ^{1}H NMR (400 MHz; D_6-DMSO): δ 10.09 (1H, s, *H*CO), 8.10 (2H, d, *J* 8.5, Ar C^3H & Ar C^5H), 8.02 (2H, d, *J* 8.5, Ar C^2H & Ar C^6H), 7.60 (2H, s, $ArSO_2NH_2$); ^{13}C NMR (126 MHz; D_6-DMSO): δ 192.6 (Ar*C*HO), 148.8 (C^1), 138.1 (C^4), 130.1 (C^3 & C^5), 126.4 (C^2 & C^6); HRMS *m/z* (ES^-) (Found: $[M-H]^-$ 184.0069. $C_7H_6NO_3S$ requires M^- 184.0074); *m/z* (ES^-) 184 ($[M-H]^-$, 100%). The data are in good agreement with the available literature values.[32]

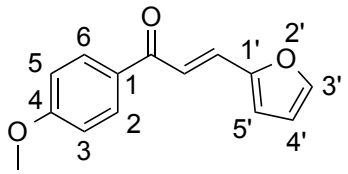
2-((4'-Sulfamoylbenzyl)amino)benzene-1-sulfonamide (85)



To a stirred solution of 4-formylbenzene-1-sulfonamide, **89**, (0.13 g, 0.68 mmol, 1.4 eq.) in ethanol (7 mL) was added 2-aminobenzene-1-sulfonamide (0.10 g, 0.58 mmol, 1.2 eq.). Acetic acid was added dropwise to the solution until a pH of 4 was attained, and the reaction mixture was stirred at RT for 15 min. After this time, $NaCNBH_3$ (0.03 g, 0.48 mmol, 1.0 eq.) was added, and the reaction mixture was stirred at 25 °C for 6 h. The solvent was removed *in vacuo* and the resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (50:50; EtOAc : petroleum ether) to afford 2-((4'-sulfamoylbenzyl)amino)benzene-1-sulfonamide, **85**, (0.09 g, 55%) as a colourless solid: R_f 0.1 (EtOAc/petroleum ether 50:50); mp 198-200 °C (from aqueous ethanol, 95% *v/v*); ν_{max}/cm^{-1} (MeOH) 3265 (m), 2360 (w), 1598 (m), 1487 (m), 1411 (w), 1328 (s), 1158 (s), 1096 (w); ^{1}H NMR (500 MHz; CD_3OD): δ 7.86 (2H, d, *J* 8.5, Ar $C^{3'}H$ & Ar $C^{5'}H$), 7.75 (1H, dd, *J* 7.7, 1.6, Ar C^6H), 7.56 (2H, d, *J* 8.5, Ar $C^{2'}H$ & Ar $C^{6'}H$), 7.25 (1H, ddd, *J* 8.9, 7.9, 1.6, Ar C^4H), 6.69 (1H, ddd, *J* 8.9, 7.7, 1.0, Ar C^5H), 6.62 (1H, dd, *J* 7.9, 1.0, Ar C^3H), 4.60 (2H, s, ArCH_2NH); ^{13}C NMR (126 MHz; CD_3OD): δ 146.2 ($C^{4'}$), 145.4 (C^2), 143.7 (C^1), 134.7 (C^4), 129.8 (C^6), 128.4 ($C^{2'}$ & $C^{6'}$), 127.5 ($C^{3'}$ & $C^{5'}$),

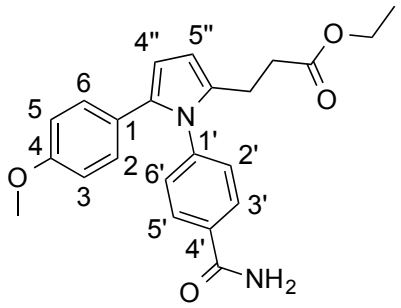
126.4 ($C^{1'}$), 116.9 (C^{5}), 113.7 (C^{3}), 47.3 (ArCH_2NH); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 364.0399. $C_{13}H_{15}N_3NaO_4S_2$ requires M^+ 364.0396); *m/z* (ES^-) 340 ($[M-H]^-$, 100%); Anal. Calcd. (Found C, 45.64; H, 4.50; N, 12.24. $C_{13}H_{15}N_3O_4S_2$ requires C, 45.74; H, 4.43; N, 12.31).

(2*E*)-3'-(furan-2'-yl)-1'-(4-methoxyphenyl)prop-2-en-1-one (91)



To a stirred solution of 4-methoxyacetophenone, **90**, (4.2 g, 27.97 mmol, 1.0 eq.) in MeOH (60 mL) was added solution of 2-furaldehyde (2.92 g, 2.52 mL, 30.42 mmol, 1.09 eq.) and sodium methoxide (1.53 g, 28.32 mmol, 1.01 eq.). The reaction mixture was stirred under nitrogen at RT for 23 h, followed by removal of the solvent *in vacuo*. The resultant mixture was diluted with H_2O (100 mL) and extracted with EtOAc (3 × 100 mL). The combined organic layers were dried over anhydrous Na_2SO_4 and the solvent was removed *in vacuo* to obtain the product (*E*)-3'-furan-2'-yl-1'-(4-methoxyphenyl)propenone, **91**, as an orange solid (6.19 g, 100%): R_f 0.5 (EtOAc/petroleum ether 20:80); mp 72-74 °C (from $CHCl_3$) [lit.[32] 76-78 °C]; ^{1}H NMR (400 MHz; $CDCl_3$): δ 8.02 (2H, dd, *J* 8.9, Ar C^2H & Ar C^6H), 7.57 (1H, d, *J* 15.5, C*H*), 7.49 (1H, d, *J* 1.6, Ar $C^{5'}H$), 7.45 (1H, d, *J* 15.5, C*H*), 6.94 (2H, d, *J* 8.9, Ar C^3H & Ar C^5H), 6.67 (1H, d, *J* 3.2, $C^{3'}H$), 6.47 (1H, dd, *J* 3.2, 1.6, $C^{4'}H$), 3.84 (3H, s, ArOCH_3); *m/z* (ES^+) 479 ($[2M+Na]^+$, 100%). The data are in good agreement with the literature values[33]

Ethyl 3''-[1-(4'-carbamoylphenyl-5''-(4-methoxyphenyl)-1*H*-pyrrol-2''-yl]propanoate (93)

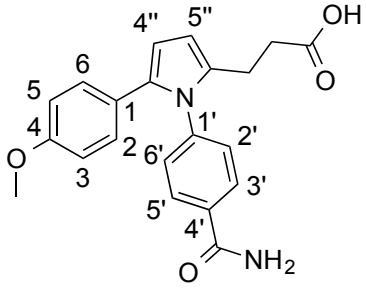


To a stirred solution of (*E*)-3-furan-2-yl-1-(4-methoxy-phenyl)-propenone, **91**, (5.94 g, 26.02 mmol, 1.0 eq.) in ethanol (112 mL) was added conc. HCl (27.8 ml) and the resulting reaction mixture was heated at 80 °C for 16 h. After this time, the reaction mixture was concentrated *in vacuo*, and diluted with CH_2Cl_2 (100 mL). The organic layer was washed with H_2O (2 × 50 mL) before being dried over anhydrous Na_2SO_4 and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (0:100 → 40:60; EtOAc : petroleum ether), to yield ethyl 7-(4-methoxyphenyl)-4,7-dioxoheptanoate, **92**, as an orange oil (3.60 g), which was used without further purification.

To a stirred solution of the orange oil containing ethyl 7-(4-methoxyphenyl)-4,7-dioxoheptanoate, **92**, (0.60 g) in ethanol (10 mL), was added 4-aminobenzamide (0.28 g, 2.06 mmol), followed by the addition of 4-toluenesulfonic acid monohydrate (3.9 mg, 0.02 mmol). The reaction mixture was heated at 80 °C for 16 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with MeOH and $CHCl_3$ (0:100 → 5:95; MeOH : $CHCl_3$), to yield ethyl 3''-[1-(4'-carbamoylphenyl-5''-(4-methoxyphenyl)-1*H*-pyrrol-2''-yl]propanoate, **93**, as an off-white solid (0.15 g, 9% over two steps): R_f 0.6 (MeOH/$CHCl_3$ 5:95); mp 100-102 °C (from MeOH/$CHCl_3$); ν_{max}/cm^{-1} (neat) 3176 (w), 1732 (m), 1651 (m), 1610 (m), 1573 (w), 1523 (m), 1482 (m), 1415 (m), 1395 (m), 1285 (m), 1247 (s), 1179 (s), 1113 (w),

1029 (m); ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.81 (2H, d, *J* 8.5, Ar C$^{3'}$*H* & Ar C$^{5'}$*H*), 7.24 (2H, d, *J* 8.5, Ar C$^{2'}$*H* & Ar C$^{6'}$*H*), 6.94 (2H, d, *J* 8.8, Ar C^{2}*H* & Ar C^{6}*H*), 6.68 (2H, d, *J* 8.8, Ar C^{3}*H* & Ar C^{5}*H*), 6.28 (1H, d, *J* 3.6, C$^{4''}$*H*), 6.10 (1H, d, *J* 3.6, C$^{35''}$*H*), 4.10 (2H, q, *J* 7.2, COO*C*$_2$H$_3$), 3.73 (3H, s, ArOC*H*$_3$), 2.79 (2H, dd, *J* 9.0, 7.0, C*H*$_2$CH$_2$CO), 2.52 (2H, dd, *J* 9.0, 7.0, CH$_2$C*H*$_2$CO), 1.23 (3H, t, *J* 7.2, COOCH$_2$C*H*$_3$); ^{13}C NMR (101 MHz; $CDCl_3$): δ 172.8 (*C*OOCH$_2$CH$_3$), 168.6 (*C*ONH$_2$), 158.2 (*C*^{4}), 142.5 (*C*$^{1'}$), 134.7 (*C*^{1}), 133.6 (*C*$^{1''}$), 132.5 (*C*$^{4'}$), 129.5 (*C*^{2} & *C*^{6}), 128.9 (*C*$^{2'}$ & *C*$^{6'}$), 128.5 (*C*$^{3'}$ & *C*$^{5'}$), 125.7 (*C*$^{3''}$), 113.7 (*C*^{3} & *C*^{5}), 108.6 (*C*$^{4''}$), 107.1 (*C*$^{5''}$), 60.7 (COO*C*H$_2$CH$_3$), 55.3 (ArO*C*H$_3$), 33.5 (CH$_2$*C*H$_2$CO), 22.7 (*C*H$_2$CH$_2$CO), 14.3 (COOCH$_2$*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+Na]$^+$ 415.1638. $C_{23}H_{24}N_2NaO_4$ requires M$^+$ 415.1628); *m/z* (ES$^+$) 415 ([M+Na]$^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 12.513 min, Purity: 98.99%.

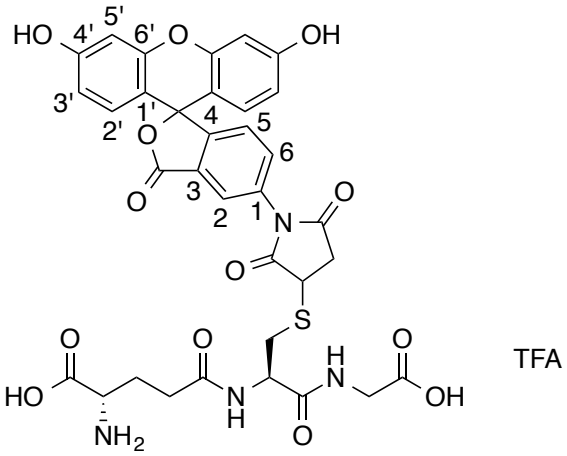
3''-[1-(4'-Carbamoylphenyl)-5''-(4-methoxyphenyl)-1*H*-pyrrol-2''-yl]propanoic acid (94)



3-[1-(4-Carbamoyl-2-methyl-phenyl)-5-(4-methoxy-phenyl)-1*H*-pyrrol-2-yl] propanoic acid ethyl ester, **93**, (0.138 g, 0.35 mmol, 1.0 eq.) was dissolved in ethanol (6.4 mL). To the ethanolic solution was added H_2O (0.8 mL) followed by the addition of aqueous 1 M NaOH (0.82 mL, 0.82 mmol, 2.3 eq.). The reaction mixture was stirred at room temperature for 1 h and then at 45 °C for 1 h. After this time, reaction mixture was concentrated *in vacuo* and the resulting residue was diluted with H_2O (100 mL) and extracted with EtOAc (3 × 50 mL). The pH of the aqueous layer was adjusted to pH 2 with aqueous 1 M HCl and then extracted with EtOAc (3 × 50 mL). The combined organic layers were dried over anhydrous Na_2SO_4 and the

concentrated *in vacuo* to obtain 3''-[1-(4'-carbamoylphenyl)-5''-(4-methoxyphenyl)-1*H*-pyrrol-2''-yl]propanoic acid, **94**, as a colourless solid (0.11 g, 87%) R_f 0.0 (EtOAc); mp 210-214 °C (from EtOAc); ν_{max}/cm^{-1} (neat) 1700 (m), 1655 (m), 1610 (m), 1525 (m), 1397 (m), 1283 (m), 1251 (s), 1220 (m), 1181 (m), 1036 (m); ^{1}H NMR (400 MHz; D_6-DMSO): δ 12.18 (s, 1H, COO*H*), 8.06 (s, 1H, CON*H*H), 7.91 (d, 2H, *J* 8.5, Ar C$^{3'}$*H* & Ar C$^{5'}$*H*), 7.47 (s, 1H, CONH*H*), 7.27 (d, 2H, *J* 8.5, Ar C$^{2'}$*H* & Ar C$^{6'}$*H*), 6.93 (d, 2H, *J* 8.8, Ar C^2*H* & Ar C^6*H*), 6.74 (d, 2H, *J* 8.8, Ar C^3*H* & Ar C^5*H*), 6.23 (d, 1H, *J* 3.5, C$^{4''}$*H*), 6.05 (d, 1H, *J* 3.5, C$^{5''}$*H*), 3.67 (s, 3H, ArOC*H*$_3$), 2.60 (dd, 2H, *J* 9.0, 7.0, C*H*$_2$CH$_2$COOH), 2.42 (dd, 2H, *J* 9.0, 7.0, CH$_2$C*H*$_2$COOH); ^{13}C NMR (101 MHz; D_6-DMSO): δ 173.6 (*C*OOH), 167.1 (*C*ONH$_2$), 157.6 (*C*^{4}), 141.2 (*C*$^{1'}$), 133.8 (*C*$^{1''}$), 133.4 (*C*$^{3''}$), 133.3 (*C*$^{4'}$), 128.8 (*C*^{2} & *C*^{6}), 128.4 (*C*$^{3'}$ & *C*$^{5'}$), 128.3 (*C*$^{2'}$ & *C*$^{6'}$), 125.4 (*C*^{1}), 113.6 (*C*^{3} & *C*^{5}), 108.1 (*C*$^{4''}$), 106.5 (*C*$^{5''}$), 55.0 (ArO*C*H$_3$), 32.7 (CH$_2$*C*H$_2$CO), 22.2 (*C*H$_2$CH$_2$CO); HRMS *m/z* (ES$^+$) (Found: [M+Na]$^+$ 387.1326. $C_{21}H_{20}N_2NaO_4$ requires M$^+$ 387.1315); *m/z* (ES$^+$) 387 ([M+Na]$^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 10.809 min, Purity: 99.51%.

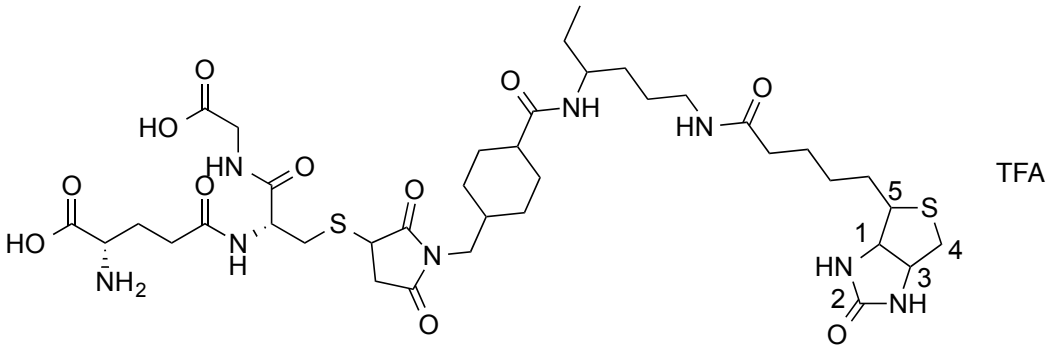
_S_-(_N_-(5-Fluoresceinyl)succinimido) L-glutathione 2,2,2-trifluoroacetate (95)



To a stirred solution of L-glutathione, **11**, (10.0 mg, 0.03 mmol, 1.1 eq.) in DMF (0.5 mL) was added *N*-(5-fluoresceinyl)maleimide (12.6 mg, 0.03 mmol, 1.0 eq.) and the resulting solution was stirred at RT for 144 h. After this time, the reaction solution was lyophilized and the resulting

solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *S*-(*N*-(5-fluoresceinyl)succinimido) L-glutathione 2,2,2-trifluoroacetate, **95**, (mixture of diastereoisomers) as a fluorescent orange solid (13.9 mg, 56%): $[\alpha]_D^{20}$ = −18.0 (*c* 0.4, MeOH); ν_{max}/cm^{-1} (neat) 2980 (w), 1697 (m), 1639 (m), 1342 (m), 1261 (m), 1165 (s), 1117 (m), 1036 (m); ^{1}H NMR (500 MHz; CD_3OD): δ 8.03 (0.5H, dd, *J* 1.7, 0.4, Ar C^2*H*), 8.01 (0.5H, dd, *J* 1.7, 0.4, Ar C^2*H*), 7.75 (0.5H, dd, *J* 8.2, 1.7, Ar C^6*H*), 7.75 (0.5H, dd, *J* 8.2, 1.7, Ar C^6*H*), 7.34 (1H, dd, *J* 8.2, 0.4, Ar C^5*H*), 6.70 (2H, d, *J* 2.4, Ar C$^{5'}$*H* & Ar C$^{9'}$*H*), 6.65 (2H, d, *J* 8.6, Ar C$^{2'}$*H* & Ar C$^{12'}$*H*), 6.56 (2H, dd, *J* 8.6, 2.4, Ar C$^{3'}$*H* & Ar C$^{11'}$*H*), 4.81-4.76 (1H, m, cys-α-C*H*), 4.28-4.21 (1H, m, succinimide-C*H*), 4.04-3.99 (1H, m, glu-α-C*H*), 3.98-3.93 (2H, m, gly-CH_2), 3.55-3.40 (1.5H, m, succinimide-CH_2, cys-β-CH_2), 3.34-3.25 (1H, m, cys-β-CH_2), 3.06-2.98 (0.5H, m, cys-β-CH_2), 2.82-2.71 (1H, m, succinimide-CH_2), 2.61 (2H, dd, *J* 7.9, 7.0, glu-ɣ-CH_2), 2.30-2.11 (2H, m, glu-β-CH_2); ^{19}F NMR (376.5 MHz; CD_3OD): δ −76.9 (OCOCF_3); ^{13}C NMR (126 MHz; CD_3OD): δ 177.9 (succinimide-*C*O), 177.8 (succinimide-*C*O), 175.6 (succinimide-*C*O), 174.5 (glu-*C*O), 172.7 (cys-*C*O), 172.7 (cys-*C*O), 172.7 (gly-*C*OOH), 171.8 (glu-*C*OOH), 171.7 (glu-*C*OOH), 170.4 (Ar*C*OOC), 170.3 (Ar*C*OOC), 162.8 (q, *J* 35.8, O*C*OCF$_3$), 161.5 (*C*$^{4'}$ & *C*$^{10'}$), 154.1 (*C*$^{6'}$ & *C*$^{8'}$), 153.7 (*C*^{4}), 135.2 (*C*^{1}), 135.0 (*C*^{6}), 134.9 (*C*^{6}), 130.3 (*C*$^{2'}$ & *C*$^{12'}$), 129.2 (*C*^{3}), 126.2 (*C*^{5}), 124.5 (*C*^{2}), 124.3 (*C*^{2}), 118.2 (q, *J* 291.1, OCO*C*F$_3$), 113.8 (*C*$^{3'}$ & *C*$^{11'}$), 110.9 (*C*$^{1'}$ & *C*$^{13'}$), 110.9 (*C*$^{1'}$ & *C*$^{13'}$), 103.6 (*C*$^{5'}$ & *C*$^{9'}$), 54.3 (cys-α-*C*H), 53.9 (cys-α-*C*H), 53.7 (glu-α-*C*H), 53.7 (glu-α-*C*H), 41.9 (succinimide-*C*H), 41.9 (gly-*C*H$_2$), 41.9 (gly-*C*H$_2$), 41.2 (succinimide-*C*H), 37.3 (succinimide-*C*H$_2$), 37.1 (succinimide-*C*H$_2$), 34.8 (cys-β-*C*H$_2$), 34.7 (cys-β-*C*H$_2$), 32.5 (glu-ɣ-*C*H$_2$), 27.1 (glu-β-*C*H$_2$), 27.0 (glu-β-*C*H$_2$), (O*C*(Ar)$_3$ observed in HMBC at 87.9 but not ^{13}C); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 735.15871. $C_{34}H_{31}N_4O_{13}S$ requires M$^+$ 735.16028); *m/z* (ES$^+$) 169 ([M+H]$^+$, 100%), 735 ([M+H]$^+$, 3%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Diastereomer 1: Ret. Time = 8.941 min, Purity: 31.46%; Diastereomer 2: Ret. Time = 9.032 min, Purity: 66.11%.

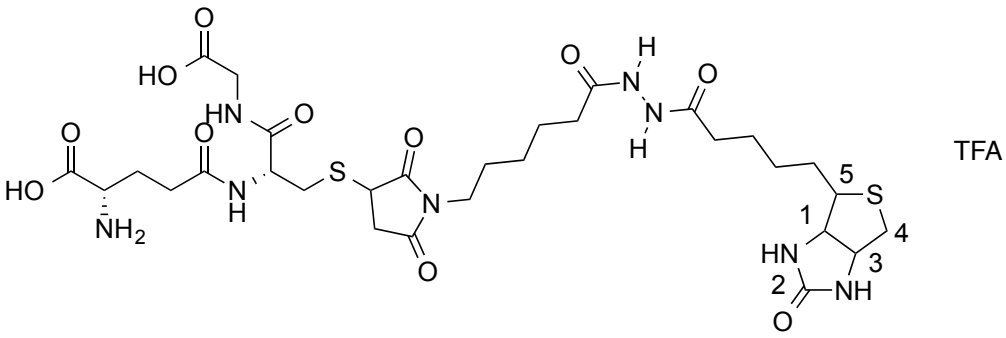
_S_-((1-Biotinamido)-4-[4'-(succinimidomethyl)cyclohexanecarboxamido]hexane) L-glutathione 2,2,2-trifluoroacetate (96)



To a stirred solution of L-glutathione, **11**, (9.8 mg, 0.03 mmol, 1.2 eq.) in THF/H_2O (1:1, 1 mL) was added (1-biotinamido)-4-[4'-(maleimidomethyl)cyclohexanecarboxamido]hexane (15.0 mg, 0.03 mmol, 1.0 eq.). The resulting solution was stirred at RT for 49 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 H_2O : MeCN : TFA → 79:21:0.1 H_2O : MeCN : TFA) to afford *S*-((1-Biotinamido)-4-[4'-(succinimidomethyl)cyclohexanecarboxamido]hexane) L-glutathione 2,2,2-trifluoroacetate, **96**, (mixture of diastereoisomers) as a colourless solid (11.6 mg, 44%): R_f 0.5 (MeOH/H_2O 30:70 (RP)); $[\alpha]_D^{20}$ = +6.0 (*c* 0.25, DMSO); ν_{max}/cm^{-1} (neat) 2933 (w), 1696 (s), 1543 (m), 1403 (w), 1177 (s); ^{1}H NMR (700 MHz; CD_3OD): δ 4.74 (0.5H, dd, *J* 8.4, 5.5, cys-α-C*H*), 4.70 (0.5H, dd, *J* 8.4, 5.5, cys-α-C*H*), 4.49 (1H, dd, *J* 7.9, 4.8, C^3*H*), 4.30 (1H, dd, *J* 7.9, 4.5, C^1*H*), 4.06-3.99 (2H, m, succinimide-C*H*, glu-α-C*H*), 3.98-3.91 (2H, m, gly-CH_2), 3.45 (0.5H, dd, *J* 13.9, 5.5, cys-β-CH_2), 3.35 (2H, d, *J* 7.1, linker-CH_2), 3.25-3.18 (3H, m, C^5*H,* succinimide-CH_2, cys-β-CH_2), 3.18-3.12 (4H, m, linker-CH_2 × 2), 2.95-2.90 (1.5H, m, C^4*H*, cys-β-CH_2), 2.71 (1H, d, *J* 12.8, C$^{4'}$*H*), 2.59 (2H, t, *J* 7.0, glu-ɣ-CH_2), 2.53 (0.5H, dd, *J* 18.5, 3.7, succinimide-CH_2), 2.48 (0.5H, dd, *J* 18.5, 3.7, succinimide-CH_2), 2.29-2.21 (1H, m, glu-β-CH_2), 2.21-2.15 (3H, m, glu-β-CH_2, linker-CH_2), 2.15-2.08 (1H, m, C*H*), 1.83-1.55 (9H, m, C*H*, linker-CH_2× 4), 1.55-1.47 (4H, m, linker-CH_2 × 2), 1.47-1.38 (4H, m, linker-CH_2 × 2), 1.38-1.28 (4H, m, C*H*, CH_3), 1.07-0.97 (2H, m, linker-CH_2); ^{19}F NMR (376.6 MHz; CD_3OD): δ −77.0 (OCOCF_3); ^{13}C NMR (176 MHz; CD_3OD): δ 179.2 (succinimide-*C*O), 179.2 (succinimide-*C*O), 178.8 (succinimide-*C*O), 177.1 (succinimide-*C*O),

177.0 (linker-*C*O), 176.0 (linker-*C*O), 174.5 (glu-*C*O), 174.4 (glu-*C*O), 172.7 (gly-*C*OOH), 172.6 (cys-*C*O), 172.6 (cys-*C*O), 171.6 (glu-*C*OOH), 171.6 (glu-*C*OOH), 166.1 (C^2), 63.4 (C^1), 61.6 (C^3), 57.0 (C^5), 54.3 (cys-α-*C*H), 53.7 (cys-α-*C*H), 53.6 (glu-α-*C*H), 53.6 (glu-α-*C*H), 46.2 (*C*H), 46.2 (*C*H), 45.7 (linker-CH_2), 45.7 (linker-CH_2), 41.9 (gly-CH_2), 41.9 (gly-CH_2), 41.3 (succinimide-*C*H), 41.1 (C^4 & $C^{4'}$), 40.4 (succinimide-*C*H), 40.2 (linker-CH_2), 40.0 (linker-CH_2), 37.2 (*C*H), 37.1 (*C*H), 37.0 (succinimide-CH_2), 36.8 (linker-CH_2), 36.7 (succinimide-CH_2), 34.7 (cys-β-CH_2), 34.6 (cys-β-CH_2), 32.4 (glu-ɣ-CH_2), 32.4 (glu-ɣ-CH_2), 31.0 (linker-CH_2), 30.9 (linker-CH_2), 30.3 (linker-CH_2), 30.1 (linker-CH_2), 30.1 (linker-CH_2), 29.8 (linker-CH_2), 29.5 (linker-CH_2), 27.5 (CH_3), 27.5 (CH_3), 27.1 (glu-β-CH_2), 27.0 (glu-β-CH_2), 27.0 (linker-CH_2), 24.2 (*C*H); HRMS *m/z* (ES^-) (Found: $[M-H]^-$ 867.3783. $C_{38}H_{59}N_8O_{11}S_2$ requires M^- 867.3750); *m/z* (ES^-) 867 ($[M-H]^-$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Diastereomer 1: Ret. Time = 8.522 min, Purity: 9.31%; Diastereomer 2: Ret. Time = 8.625 min, Purity: 86.56%.

(1*S*)-1-Carboxy-4-(((2*R*)-1-((carboxymethyl)amino)-3-((2,5-dioxo-1-(6-oxo-6-(2-(5-(2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanoyl)hydrazinyl)hexyl)pyrrolidin-3-yl)thio)-1-oxopropan-2-yl)amino)-4-oxobutan-1-aminium 2,2,2-trifluoroacetate (97)

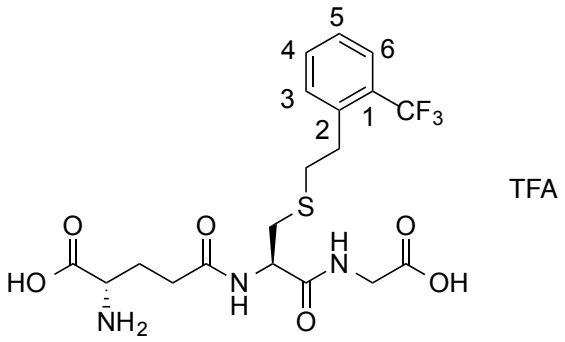


To a stirred solution of L-glutathione, **11**, (12.0 mg, 0.04 mmol, 1.2 eq.) in DMF (1 mL) was added biotin-maleimide (15.0 mg, 0.03 mmol, 1.0 eq.), and the resulting mixture was stirred at RT for 55 h. After this time, H_2O (2 mL) was added and the reaction mixture was lyophilized. The resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column;

95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford (1*S*)-1-carboxy-4-(((2*R*)-1-((carboxymethyl)amino)-3-((2,5-dioxo-1-(6-oxo-6-(2-(5-(2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanoyl)hydrazinyl)hexyl)pyrrolidin-3-yl)thio)-1-oxopropan-2-yl)amino)-4-oxobutan-1-aminium 2,2,2-trifluoroacetate, **97**, (mixture of diastereoisomers) as a colourless solid (8.9 mg, 31%): R_f 0.9 (H_2O (RP)); $[\alpha]_D^{20}$ = +13.6 (*c* 0.4, DMSO); ν_{max}/cm^{-1} (neat) 2940 (w), 1738 (s), 1691 (m), 1369 (m), 1204 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 13.93 (1H, s, COO*H*), 12.63 (1H, s, COO*H*), 9.64 (2H, s, CON*H*N*H*CO), 8.40 (1H, t, *J* 5.7, gly-N*H*), 8.35 (1H, dd, *J* 8.3, 5.9, cys-N*H*), 8.24 (3H, s, glu-NH_3^+), 6.41 (1H, s, Biotin-N*H*), 6.36 (1H, s, Biotin-N*H*), 4.60-4.50 (1H, m, cys-α-C*H*), 4.31 (1H, dd, *J* 7.8, 5.1, C^3*H*), 4.16-4.10 (1H, m, C^1*H*), 4.06-3.98 (1H, m, succinimide-C*H*), 3.95 (1H, s, glu-α-C*H*), 3.77 (2H, d, *J* 5.7, gly-CH_2), 3.50-3.29 (2H, m, linker-CH_2), 3.25-3.13 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 3.12-3.03 (1.5H, m, C^5*H*, cys-β-CH_2), 2.94 (0.5H, dd, *J* 13.0, 8.9, cys-β-CH_2), 2.82 (1H, dd, *J* 12.4, 5.1, C^4*H*), 2.77 (0.5H, dd, *J* 13.0, 9.7, cys-β-CH_2), 2.58 (1H, d, *J* 12.4, C$^{4'}$*H*), 2.55-2.44 (1H, m, succinimide-CH_2), 2.44-2.27 (2H, m, glu-ɣ-CH_2), 2.14-1.90 (6H, m, glu-β-CH_2, linker-CH_2 × 2), 1.67-1.57 (1H, m, linker-CH_2 × 0.5), 1.57-1.41 (7H, m, linker-CH_2 × 3.5), 1.41-1.27 (2H, m, linker-CH_2), 1.27-1.13 (2H, m, linker-CH_2); ^{19}F NMR (376.6 MHz; D_6-DMSO): δ −73.6 (OCOCF_3); ^{13}C NMR (126 MHz; D_6-DMSO): δ 176.8 (succinimide-*C*O), 176.7 (succinimide-*C*O), 175.1 (succinimide-*C*O), 175.1 (succinimide-*C*O), 171.1 (glu-*C*OOH), 171.0 (O*C*NHNHCO), 171.0 (OCNHNH*C*O), 171.0 (glu-*C*O), 170.9 (cys-*C*O), 170.9 (cys-*C*O), 170.3 (gly-*C*OOH), 170.2 (gly-*C*OOH), 162.7 (NH*C*ONH), 157.7 (q, *J* 30.8, O*C*OCF$_3$), 117.3 (q, *J* 299.1, OCO*C*F$_3$), 61.0 (*C*^{1}), 59.2 (*C*^{3}), 55.4 (*C*^{5}), 52.0 (cys-α-*C*H), 51.7 (cys-α-*C*H), 51.6 (glu-α-*C*H), 40.8 (gly-*C*H$_2$), 39.6 (*C*^{4} & *C*$^{4'}$), 39.5 (succinimide-*C*H), 39.0 (succinimide-*C*H), 38.1 (linker-*C*H$_2$), 38.0 (linker-*C*H$_2$), 35.9 (succinimide-*C*H$_2$), 35.8 (succinimide-*C*H$_2$), 33.3 (cys-β-*C*H$_2$), 33.2 (cys-β-*C*H$_2$), 33.0 (linker-*C*H$_2$), 32.9 (linker-*C*H$_2$), 30.6 (glu-ɣ-*C*H$_2$), 30.5 (glu-ɣ-*C*H$_2$), 28.1 (linker-*C*H$_2$), 28.0 (linker-*C*H$_2$), 26.8 (linker-*C*H$_2$), 25.9 (glu-β-*C*H$_2$), 25.9 (glu-β-*C*H$_2$), 25.7 (linker-*C*H$_2$), 25.1 (linker-*C*H$_2$), 24.6 (linker-*C*H$_2$); HRMS *m/z* (ES$^-$) (Found: [M−H]$^-$ 757.2677. $C_{30}H_{45}N_8O_{11}S_2$ requires M$^-$ 757.2655); *m/z* (ES$^-$) 757 ([M−H]$^-$, 100%);

Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 7.938 min, Purity: 99.03%.

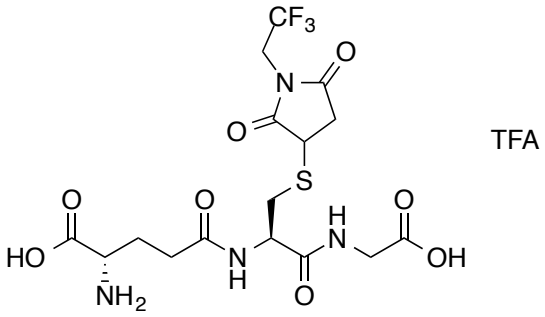
_S_-2-(2-(Trifluoromethyl)phenyl)ethyl L-glutathione 2,2,2-trifluoroacetate (98)



2-Trifluoromethyl styrene (58.8 mg, 50 μL, 0.34 mmol, 1.0 eq.), L-glutathione, **11**, (420 mg, 1.37 mmol, 4.0 eq.) and 2,2-dimethoxyphenyl acetophenone (17.5 mg, 0.07 mmol, 0.2 eq.) were stirred at RT in THF/H_2O (1:2, 3 mL) in the presence of light (365 nm, 4 × 15 W) for 5 h. After this time the reaction was filtered, and the solid washed with ethanol and H_2O. The resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *S*-2-(2-(trifluoromethyl)phenyl)ethyl L-glutathione 2,2,2-trifluoroacetate, **98**, as a colourless solid (41.1 mg, 20%): $[\alpha]_D^{25}$ = −9.9 (*c* 1.0, DMSO); ν_{max}/cm^{-1} (neat) 3368 (w), 3341 (w), 2970 (w), 1738 (m), 1672 (m), 1643 (m), 1584 (w), 1512 (m), 1454 (m), 1433 (m), 1352 (m), 1312 (s), 1258 (m), 1231 (m), 1217 (m), 1163 (m), 1136 (m), 1109 (s), 1082 (m), 1061 (m), 1040 (m); ^{1}H NMR (500 MHz; D_6-DMSO): δ 8.55 (1H, dd, *J* 6.8, 5.9, gly-N*H*), 8.33 (1H, d, *J* 8.5, cys-N*H*), 7.68 (1H, d, *J* 7.7, Ar C^6*H*), 7.63 (1H, dd, *J* 8.1, 7.7, Ar C^4*H*), 7.56 (1H, d, *J* 7.7, Ar C^3*H*), 7.44 (1H, dd, *J* 8.1, 7.7, Ar C^5*H*), 4.56-4.48 (1H, m, cys-α-C*H*), 3.80-3.68 (3H, m, gly-CH_2, glu-α-C*H*), 3.02-2.92 (3H, m, cys-β-CH_2, linker-CH_2), 2.79-2.74 (2H, m, linker-CH_2), 2.68 (1H, dd, *J* 13.5, 9.4, cys-β-CH_2), 2.40-2.28 (2H, m, glu-ɣ-CH_2), 2.03-1.90 (2H, m, glu-β-CH_2); ^{19}F NMR (377 MHz; D_6-DMSO): δ −58.2 (ArCF_3), −73.4 (OCOCF_3); ^{13}C NMR (126 MHz; D_6-DMSO): δ 171.3 (gly-*C*OOH), 170.9 (glu-*C*OOH), 170.7 (glu-*C*O), 170.6 (cys-*C*O), 157.7 (q, *J* 30.3, O*C*OCF$_3$), 138.7 (*C*^{2}), 132.6 (*C*^{4}), 131.6 (*C*^{3}), 127.0 (q, *J* 28.9, *C*^{5}), 126.9 (*C*^{6}), 125.7 (q, *J* 5.8, *C*^{1}),

124.7 (q, *J* 274.0, Ar*C*F_3), 117.5 (q, *J* 299.8, OCO*C*F_3), 52.2 (glu-α-*C*H), 52.1 (cys-α-*C*H), 40.8 (gly-*C*H_2), 33.7 (cys-β-*C*H_2), 32.3 (linker-*C*H_2 × 2), 31.0 (glu-ɣ-*C*H_2), 26.4 (glu-β-*C*H_2); HRMS *m/z* (ES^+) (Found: $[M+H]^+$ 480.14037. $C_{19}H_{25}N_3O_6SF_3$ requires M^+ 480.14107); *m/z* (ES^+) 480 ($[M+H]^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.613 min, Purity: 95.48%.

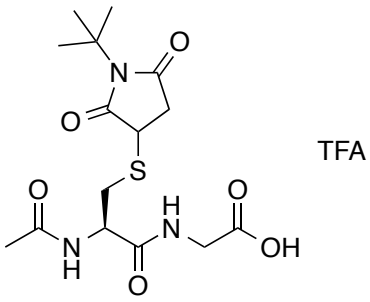
_S_-(_N_-2,2,2-Trifluoroethylsuccinimido) L-glutathione 2,2,2-trifluoroacetate (99)



To a stirred solution of L-glutathione, **11**, (30.5 mg, 0.10 mmol, 1.0 eq.) and NaOH (4.0 mg, 0.10 mmol, 1.0 eq.) in H_2O (0.8 mL) was added 1-(2,2,2-trifluoroethyl)-2,5-dihydro-1*H*-pyrrole-2,5-dione (17.5 mg, 0.10 mmol, 1.0 eq.). The resulting solution was stirred at RT for 9.5 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *S*-(*N*-2,2,2-trifluoroethylsuccinimido) L-glutathione 2,2,2-trifluoroacetate, **99**, (mixture of diastereoisomers) as a colourless solid (50.3 mg, 84%): R_f 0.9 (H_2O (RP)); $[\alpha]_D^{20}$ = −9.2 (*c* 1.0, H_2O); ν_{max}/cm^{-1} (neat) 2980 (w), 1711 (m), 1655 (m), 1524 (m), 1396 (m), 1344 (m), 1260 (m), 1163 (m), 1125 (s), 1028 (m); ^1H NMR (500 MHz; D_2O): δ 4.72-4.67 (1H, m, cys-α-C*H*), 4.34-4.26 (2H, m, C*H*$_2$CF_3), 4.20-4.13 (1H, m, succinimide-C*H*), 4.06-3.97 (3H, m, gly-C*H*$_2$, glu-α-C*H*), 3.45-3.32 (1.5H, m, cys-β-C*H*$_2$, succinimide-C*H*$_2$), 3.26 (0.5H, dd, *J* 14.1, 5.4, cys-β-C*H*$_2$), 3.18 (0.5H, dd, *J* 14.1, 8.0, cys-β-C*H*$_2$), 3.04 (0.5H, dd, *J* 14.1, 8.7, cys-β-C*H*$_2$), 2.83-2.73 (1H, m, succinimide-C*H*$_2$), 2.65-2.51 (2H, m, glu-γ-C*H*$_2$), 2.29-2.15 (2H, m, glu-β-C*H*$_2$); ^{19}F NMR (377 MHz; D_2O): δ −70.08 (ArC*F*$_3$), −70.09 (ArC*F*$_3$), −75.7 (OCOC*F*$_3$);

^{13}C NMR (126 MHz; D_2O): δ 178.1 (succinimide-*C*O), 178.1 (succinimide-*C*O), 176.6 (succinimide-*C*O), 176.6 (succinimide-*C*O), 174.4 (glu-*C*O), 174.4 (glu-*C*O), 172.9 (gly-*C*OOH), 172.9 (gly-*C*OOH), 172.3 (cys-*C*O), 172.2 (cys-*C*O), 172.1 (glu-*C*OOH), 163.0 (q, *J* 35.4, O*C*OCF$_3$), 123.1 (q, *J* 278.9, CH$_2$*C*F$_3$), 116.3 (q, *J* 291.4, OCO*C*F$_3$), 53.1 (cys-α-*C*H), 52.6 (glu-α-*C*H), 52.6 (glu-α-*C*H), 41.1 (gly-*C*H$_2$), 41.1 (gly-*C*H$_2$), 40.6 (succinimide-C*H*), 39.9 (succinimide-C*H*), 39.7 (q, *J* 35.8, *C*H$_2$CF$_3$), 35.9 (succinimide-*C*H$_2$), 35.6 (succinimide-*C*H$_2$), 32.9 (cys-β-*C*H$_2$), 32.7 (cys-β-*C*H$_2$), 30.9 (glu-ɣ-*C*H$_2$), 30.9 (glu-ɣ-*C*H$_2$), 25.6 (glu-β-*C*H$_2$), 25.6 (glu-β-*C*H$_2$); HRMS *m/z* (ES$^+$) (Found: [M+Na]$^+$ 509.09155. $C_{16}H_{21}N_4O_8SNaF_3$ requires M$^+$ 509.09244); *m/z* (ES$^+$) 487 ([M+H]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Diastereomer 1: Ret. Time = 8.039 min, Purity: 50.60%; Diastereomer 2: Ret. Time = 8.126 min, Purity: 48.87%.

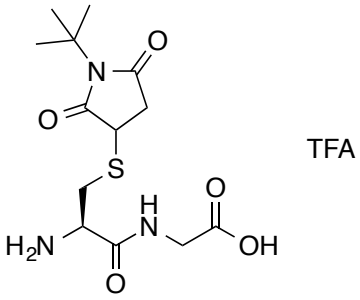
_N_-Acetyl-_S_-(_N-tert_-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (101)



To a stirred solution of *S*-(*N-tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate, **102**, (15.0 mg, 0.03 mmol, 1.0 eq.) and anhydrous NEt$_3$ (6.8 mg, 9.4 µL, 0.07 mmol, 2.0 eq.) in anhydrous THF (0.3 mL) was added acetyl chloride (2.9 mg, 2.6 µL, 0.04 mmol, 1.1 eq.). The reaction mixture was stirred under argon at RT for 6.5 h. After this time, the reaction mixture was diluted with H_2O (0.5 mL), lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *N*-acetyl-*S*-(*N-tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate, **101**, (mixture of diastereoisomers) as a colourless solid (9.0 mg, 55%):

R_f 1.0 (H_2O (RP)); $[\alpha]_D^{20}$ −14.8 (*c* 0.46, H_2O); ν_{max}/cm^{-1} (neat) 2971 (w), 1738 (m), 1694 (s), 1533 (m), 1346 (m), 1264 (m), 1202 (m), 1164 (s); ^1H NMR (500 MHz; D_2O): δ 4.69-4.61 (1H, m, cys-α-C*H*), 4.01 (2H, s, gly-CH_2), 3.92-3.84 (1H, m, succinimide-C*H*), 3.31 (0.5H, dd, *J* 14.0, 5.1, cys-β-CH_2) 3.24-3.10 (2H, m, cys-β-CH_2, succinimide-CH_2), 3.00 (0.5H, dd, *J* 14.0, 8.7, cys-β-CH_2), 2.57 (0.5H, dd, *J* 4.3, 3.6, succinimide-CH_2), 2.54 (0.5H, dd, *J* 4.3, 3.6, succinimide-CH_2), 2.08 (1.5H, s, COCH_3), 2.07 (1.5H, s, COCH_3), 1.54 (4.5H, s, NC(CH_3)$_3$), 1.53 (4.5H, s, NC(CH_3)$_3$); ^{19}F NMR (377 MHz; D_2O): δ −75.6 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.2 (succinimide-*C*O), 180.2 (succinimide-*C*O), 179.1 (succinimide-*C*O), 179.1 (succinimide-*C*O), 174.4 (*C*OCH$_3$), 174.3 (*C*OCH$_3$), 172.9 (*C*OOH), 172.4 (cys-*C*O), 172.4 (cys-*C*O), 163.0 (q, *J* 35.4, O*C*OCF$_3$), 116.4 (q, *J* 290.5, OCO*C*F$_3$), 59.2 (N*C*(CH$_3$)$_3$), 59.1 (N*C*(CH$_3$)$_3$), 53.1 (cys-α-*C*H), 52.7 (cys-α-*C*H), 41.1 (gly-*C*H$_2$), 40.8 (succinimide-*C*H), 40.0 (succinimide-*C*H), 36.2 (succinimide-*C*H$_2$), 35.9 (succinimide-*C*H$_2$), 32.7 (cys-β-*C*H$_2$), 32.3 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$), 21.7 (CO*C*H$_3$); HRMS *m/z* (ES^-) (Found: $[M-H]^-$ 372.12347. $C_{15}H_{22}O_6N_3S$ requires M^- 372.12348); *m/z* (ES^+) 374 ($[M+H]^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.519 min, Purity: 98.82%. Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Diastereomer 1: Ret. Time = 21.810 min, 32.99%; Diastereomer 2: Ret. Time = 23.640 min, 66.99%.

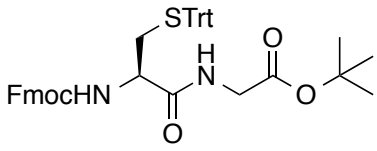
_S_-(_N-tert_-Butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (102)



To a stirred solution of L-cys-gly, **103**, (15.0 mg, 0.08 mmol, 1.0 eq.) and NaOH (3.4 mg, 0.08 mmol, 1.0 eq.) in H_2O (0.7 mL) was added *N-tert*-butyl-maleimide (12.9 mg, 12.2 µL, 0.08 mmol, 1.0 eq.). The resulting solution was stirred at RT for 1.5 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 81:19:0.1; H_2O : MeCN : TFA) to afford *S*-(*N-tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate, **102**, (mixture of diastereoisomers) as a colourless solid (21.9 mg, 58%): R_f 0.3 (H_2O (RP)); $[\alpha]_D^{20}$ = +25.8 (*c* 0.65, H_2O); ν_{max}/cm^{-1} (neat); 2971 (w), 1676 (s), 1349 (m), 1265 (m), 1200 (s); ^{1}H NMR (500 MHz; D_2O): δ 4.39 (0.5H, dd, *J* 7.5, 5.8, cys-α-C*H*), 4.35 (0.5H, dd, *J* 7.5, 5.8, cys-α-C*H*), 4.08-3.96 (2H, m, gly-CH_2), 3.95-3.89 (1H, m, succinimide-C*H*), 3.49 (0.5H, dd, *J* 14.7, 5.8, cys-β-CH_2), 3.41-3.30 (1H, m, cys-β-CH_2), 3.22-3.12 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 2.58 (0.5H, dd, *J* 5.6, 5.2, succinimide-CH_2), 2.55 (0.5H, dd, *J* 5.6, 5.2, succinimide-CH_2), 1.54 (9H, s, NC(CH_3)$_3$); ^{19}F NMR (377 MHz; D_2O): δ −75.6 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.5 (succinimide-*C*O), 180.2 (succinimide-*C*O), 178.8 (succinimide-*C*O), 178.8 (succinimide-*C*O), 173.6 (*C*OOH), 173.6 (*C*OOH), 168.4 (cys-*C*O), 168.3 (cys-*C*O), 163.0 (q, *J* 35.6, O*C*OCF_3), 116.4 (q, *J* 290.9, OCO*C*F_3), 59.2 ((N*C*(CH_3)$_3$), 52.6 (cys-α-*C*H), 52.3 (cys-α-*C*H), 41.8 (gly-*C*H_2), 41.8 (gly-*C*H_2), 41.0 (succinimide-*C*H), 40.1 (succinimide-*C*H), 35.9 (succinimide-*C*H_2), 35.6 (succinimide-*C*H_2), 32.8 (cys-β-*C*H_2), 32.2 (cys-β-*C*H_2), 27.3 (NC(*C*H_3)$_3$); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 354.1098. $C_{13}H_{21}N_3NaO_5S$ requires M^+ 354.1094); *m/z* (ES^+) 332 ($[M+H]^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA)

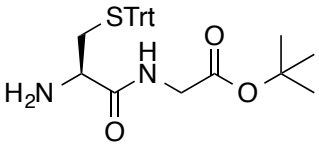
Diastereomer 1: Ret. Time = 8.517 min, Purity: 40.55%; Diastereomer 2: Ret. Time = 8.605 min, Purity: 57.21%.

_N_-Fmoc-_S_-trityl-L-cysteinylglycine _tert_-butyl ester (105)



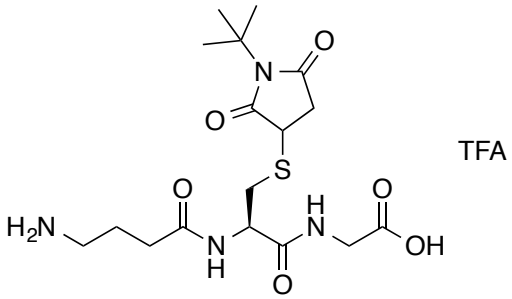
_N_-Fmoc-L-Cys(Trt)-OH, **104**, (0.75 g, 1.28 mmol, 1.05 eq.), glycine _tert_-butyl ester hydrochloride (0.20 g, 1.22 mmol, 1.0 eq.) and HBTU (0.49 g, 1.28 mmol, 1.05 eq.) were stirred in anhydrous DMF (3 mL) for 5 min. After this time, DIPEA (0.32 g, 0.42 mL, 2.44 mmol, 2.0 eq.) was added and the reaction mixture was stirred under argon at RT for 24 h. The reaction mixture was partitioned between EtOAc (50 mL) and aqueous 0.5 M LiCl (50 mL) and the organic layer was collected, dried over anhydrous Na_2SO_4, filtered and concentrated _in vacuo_. The resulting crude material was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether and NEt_3 (10:89:1 → 80:19:1; EtOAc : petroleum ether : NEt_3) to yield _N_-Fmoc-_S_-trityl-L-cysteinylglycine _tert_-butyl ester, **105**, (0.85 g, 99%) as a colourless solid: R_f 0.8 (EtOAc/petroleum ether 40:60); mp 79-83 °C (from EtOAc/petroleum ether); $[\alpha]_D^{20}$ = +5.8 (_c_ 1.0, $CHCl_3$) {lit.[33] $[\alpha]_D^{25}$ = +2.4 (_c_ 0.25, $CHCl_3$)}; ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.73-7.08 (23H, m, Ar C_H_), 6.25 (1H, s, gly-N_H_), 4.92 (1H, d, _J_ 7.5, cys-N_H_), 4.36-4.26 (2H, m, Fmoc-CH_2), 4.13 (1H, t, _J_ 7.5, 6.8, Fmoc-C_H_), 3.85-3.67 (3H, m, cys-α-C_H_, gly-CH_2), 2.67-2.52 (2H, m, cys-β-CH_2), 1.37 (9H, s, OC(CH_3)$_3$); _m/z_ (ES$^+$) 721 ([M+Na]$^+$, 100%). The data are in good agreement with literature values.[33]

_S_-Trityl-L-cysteinylglycine _tert_-butyl ester (106)



_N_-Fmoc-_S_-trityl-L-cysteinylglycine _tert_-butyl ester, **105**, (0.75 g, 1.07 mmol, 1.0 eq.) was stirred under argon at RT in 20% piperidine in DMF (6 mL) for 5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether and NEt_3 (12:87:1 → 99:0:1; EtOAc : petroleum ether : NEt_3) to yield _S_-trityl-L-cysteinylglycine _tert_-butyl ester, **106**, (0.38 g, 74%) as a yellow oil: R_f 0.2 (EtOAc/petroleum ether 50:50); $[\alpha]_D^{20}$ = +13.2 (*c* 1.0, $CHCl_3$) {lit.[33] $[\alpha]_D^{25}$ = +9.2 (*c* 0.5, $CHCl_3$)}; ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.45 (1H, t, *J* 5.4, gly-N*H*), 7.40-7.02 (15H, m, Ar C*H*), 3.77 (1H, dd, *J* 18.3, 5.4, gly-CH_AH$_B$), 3.69 (1H, dd, *J* 18.3, 5.4, gly-CH$_A$$H_B$), 2.91 (1H, dd, *J* 8.8, 3.8, cys-α-C*H*), 2.65 (1H, dd, *J* 12.8, 3.8, cys-β-CH_2), 2.45 (1H, dd, *J* 12.8, 8.8, cys-β-CH_2), 1.33 (9H, s, OC(CH_3)$_3$); *m/z* (ES$^+$) 477 ([M+H]$^+$, 100%). The data are in good agreement with literature values.[34]

ɣ-Aminobutyric acid-_S_-(_N-tert_-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (109)



_S_-Trityl-L-cysteinylglycine _tert_-butyl ester, **106**, (89.3 mg, 0.19 mmol, 1.0 eq.), _N_-Boc-ɣ-aminobutyric acid (38.1 mg, 0.19 mmol, 1.0 eq.), EDC•HCl (41.3 mg, 0.22 mmol, 1.15 eq.), HBTU (7.1 mg, 0.02 mmol, 0.1 eq.) and DIPEA (48.4 mg, 65.3 µL, 0.37 mmol, 2.0 eq.) were stirred under argon at RT in anhydrous CH_2Cl_2 (9 mL) for 24 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was adsorbed onto Celite® and

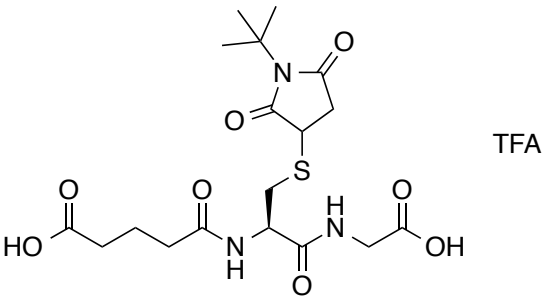
purified by silica gel chromatography, eluting with EtOAc, petroleum ether and NEt_3 (12:87:1 → 99:0:1; EtOAc : petroleum ether : NEt_3) to yield *N*-Boc-ɣ-aminobutyric acid-*S*-trityl-L-cysteinylglycine *tert*-butyl ester, **107**, (0.10 g) as a colourless oil, which was used without further purification.

N-Boc-ɣ-aminobutyric acid-*S*-trityl-L-cysteinylglycine *tert*-butyl ester, **107**, (95.0 mg) and triethylsilane (50.1 mg, 68.8 µL, 0.43 mmol) were stirred in 20% TFA/CH_2Cl_2 (5 mL) at RT for 6.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was purified by RP C-18 silica gel chromatography, eluting with H_2O to afford ɣ-aminobutyric acid-L-cysteinylglycine trifluoroacetic acid salt, **108**, (33.3 mg) as a colourless residue, which was used without further purification.

To a stirred solution of ɣ-aminobutyric acid-L-cysteinylglycine trifluoroacetic acid salt, **108**, (25.0 mg) in H_2O (0.5 mL) was added *N*-*tert*-butyl-maleimide (10.2 mg, 9.6 µL, 0.07 mmol). The resulting solution was stirred at RT for 19.5 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford ɣ-aminobutyric acid-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate, **109**, (mixture of diastereoisomers) as a colourless solid (25.9 mg, 37% over three steps): $[\alpha]_D^{20}$ = −14.5 (*c* 1.0, H_2O); ν_{max}/cm^{-1} (neat) 2940 (w), 1690 (m), 1653 (s), 1528 (m), 1460 (m), 1418 (m), 1346 (m), 1263 (m), 1231 (m), 1167 (s), 1130 (s); ^{1}H NMR (400 MHz; D_2O): δ 4.66-4.57 (1H, m, cys-α-C*H*), 3.97 (2H, s, gly-CH_2), 3.90-3.79 (1H, m, succinimide-C*H*), 3.28 (0.5H, dd, *J* 14.2, 5.1, cys-β-CH_2), 3.22-3.04 (2H, m, cys-β-CH_2, succinimide-CH_2), 3.04-2.89 (2.5H, m, CH_2, cys-β-CH_2), 2.54 (0.5H, dd, *J* 8.1, 4.3, succinimide-CH_2), 2.49 (0.5H, dd, *J* 8.1, 4.3, succinimide-CH_2), 2.46-2.38 (2H, m, CH_2), 1.99-1.87 (2H, m, CH_2), 1.49 (4.5H, s, NC(CH_3)$_3$), 1.49 (4.5H, s, NC(CH_3)$_3$); ^{19}F NMR (377 MHz; D_2O): δ −75.6 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.5 (succinimide-*C*O),

180.5 (succinimide-*C*O), 179.4 (succinimide-*C*O), 179.4 (succinimide-*C*O), 175.4 (CH_2CH_2*C*ONH), 175.3 (CH_2CH_2*C*ONH), 173.2 (gly-*C*OOH), 173.2 (gly-*C*OOH), 172.8 (cys-*C*O), 172.7 (cys-*C*O), 163.3 (q, *J* 35.2, O*C*OCF_3), 116.7 (q, *J* 292.0, OCO*C*F_3), 59.6 (N*C*$(CH_3)_3$), 59.6 (N*C*$(CH_3)_3$), 53.4 (cys-α-*C*H), 53.0 (cys-α-*C*H), 41.5 (gly-*C*H_2), 41.2 (succinimide-*C*H), 40.4 (succinimide-*C*H), 39.2 (*C*H_2), 36.6 (succinimide-*C*H_2), 36.3 (succinimide-*C*H_2), 33.0 (cys-β-*C*H_2), 32.8 (cys-β-*C*H_2), 32.5 (*C*H_2), 27.7 (NC(*C*$H_3)_3$), 23.1 (*C*H_2); HRMS *m/z* (ES^+) (Found: $[M+H]^+$ 417.17911 $C_{17}H_{29}N_4O_6S$ requires M^+ 417.18023); *m/z* (ES^+) 833 ($[2M+H]^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Diastereomer 1: Ret. Time = 8.501 min, Purity: 28.42%; Diastereomer 2: Ret. Time = 8.553 min, Purity: 70.16%.

Glutaric acid-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (112)



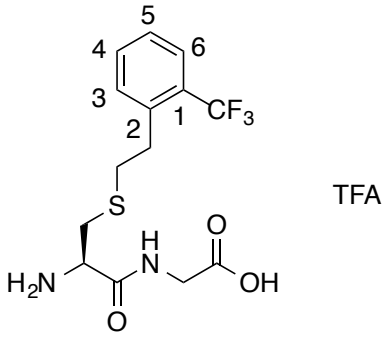
S-Trityl-L-cysteinylglycine *tert-butyl* ester, **106**, (0.21 g, 0.44 mmol, 1.0 eq.) and anhydrous NEt_3 (0.06 g, 0.08 mL, 0.57 mmol, 1.3 eq.) were stirred under argon at RT in anhydrous $CHCl_3$ (2.0 mL) for 5 min. After this time, glutaric anhydride (0.06 g, 0.48 mmol, 1.1 eq.) was added and the reaction mixture was stirred at RT for 15.5 h. The reaction solution was concentrated *in vacuo* and the resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether and formic acid (60:40:1; EtOAc : petroleum ether : formic acid) to yield slightly impure glutaric acid-*S*-trityl-L-cysteinylglycine *tert*-butyl ester, **110**, (21.9 mg) as a colourless oil, which was used without further purification.

Glutaric acid-*S*-trityl-L-cysteinylglycine *tert*-butyl ester, **110**, (15.9 mg) and triethylsilane (94.1 mg, 0.13 mL, 0.81 mmol) were stirred at RT under argon in 20% TFA/CH_2Cl_2 (9.4 mL) at RT for 6.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was purified by RP C-18 silica gel chromatography, eluting with H_2O to afford slightly impure glutaric acid-L-cysteinylglycine trifluoroacetic acid salt, **111**, (23.7 mg) as a colourless solid, which was used without further purification.

To a stirred solution of glutaric acid-L-cysteinylglycine trifluoroacetic acid salt, **111**, (22.4 mg) in H_2O (0.4 mL) was added *N*-*tert*-butyl-maleimide (8.5 mg, 8.0 µL, 0.06 mmol). The resulting solution was stirred at RT for 14 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford glutaric acid-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate, **112**, (mixture of diastereoisomers) as a colourless solid (20.2 mg, 12% over three steps): $[\alpha]_D^{20}$ = −20.0 (*c* 1.0, H_2O); ν_{max}/cm^{-1} (neat) 2970 (w), 2941 (w), 1738 (m), 1690 (s), 1530 (m), 1456 (m), 1366 (m), 1346 (m), 1263 (m), 1229 (m), 1217 (s), 1202 (s), 1159 (s), 1042 (m); ^{1}H NMR (500 MHz; D_2O): δ 4.70-4.63 (1H, m, cys-α-C*H*), 4.01 (2H, s, gly-CH_2), 3.91-3.83 (1H, m, succinimide-C*H*), 3.33 (0.5H, dd, *J* 14.2, 5.1, cys-β-CH_2), 3.23 (0.5H, dd, *J* 14.2, 5.1, cys-β-CH_2), 3.19-3.09 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 2.98 (0.5H, dd, *J* 14.2, 9.2, cys-β-CH_2), 2.55 (0.5H, dd, *J* 5.2, 4.6, succinimide-CH_2), 2.51 (0.5H, dd, *J* 5.2, 4.6, succinimide-CH_2), 2.49-2.33 (4H, m, CH_2), 1.96-1.86 (2H, m, CH_2), 1.53 (4.5H, s, NC(CH_3)$_3$), 1.52 (4.5H, s, NC(CH_3)$_3$); ^{19}F NMR (376.6 MHz; D_2O): δ −75.7 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.2 (succinimide-*C*O), 180.1 (succinimide-*C*O), 179.0 (succinimide-*C*O), 179.0 (succinimide-*C*O), 177.7 (CH_2CH_2*C*OOH), 176.1 (CH_2CH_2*C*ONH), 176.0 (CH_2CH_2*C*ONH), 172.9 (gly-*C*OOH), 172.4 (cys-*C*O), 172.4 (cys-*C*O), 162.9 (q, *J* 35.5, O*C*OCF_3), 116.3 (q, *J* 292.0, OCO*C*F_3), 59.1 (N*C*(CH_3)$_3$), 59.1 (N*C*(CH_3)$_3$), 53.1 (cys-α-*C*H), 52.5 (cys-α-*C*H), 41.1 (gly-*C*H_2), 41.1 (gly-*C*H_2), 40.8 (succinimide-*C*H), 39.9 (succinimide-*C*H),

36.1 (succinimide-*C*H_2), 35.9 (succinimide-*C*H_2), 34.4 (*C*H_2), 32.8 (cys-β-*C*H_2), 32.7 (*C*H_2), 32.5 (cys-β-*C*H_2), 27.3 (NC(*C*H_3)$_3$), 20.4 (*C*H_2); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 468.13940. $C_{18}H_{27}N_3NaO_8S$ requires M^+ 468.14111); *m/z* (ES^+) 468 ($[M+Na]^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.116 min, Purity: 97.61%.

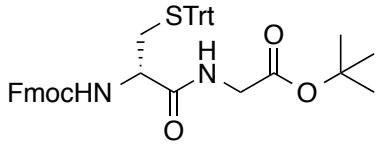
_S_-2-(2-(Trifluoromethyl)phenyl)ethyl L-cys-gly 2,2,2-trifluoroacetate (113)



2-Trifluoromethyl styrene (24.2 mg, 20.6 µL, 0.14 mmol, 1.0 eq.), L-cys-gly, **103**, (0.1 g, 0.56 mmol, 4.0 eq.) and 2,2-dimethoxyphenyl acetophenone (7.2 mg, 0.03 mmol, 0.2 eq.) were stirred at RT in THF/H_2O (1:2, 1.3 mL) in the presence of light (365 nm, 4 × 15 W) for 5 h. After this time, the reaction solution was lyophilized and the crude material purified by RP C-18 silica gel chromatography, eluting with MeCN and H_2O (0:100 → 50:50; MeCN : H_2O) to yield a colourless solid which was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 60:40:0.1; H_2O : MeCN : TFA) to afford *S*-2-(2-(trifluoromethyl)phenyl)ethyl L-cys-gly 2,2,2-trifluoroacetate, **113**, as a colourless solid (11.7 mg, 19%): R_f 0.5 (MeCN/H_2O 50:50 (RP)); $[\alpha]_D^{25}$ = +1.0 (*c* 0.8, DMSO); ν_{max}/cm^{-1} (neat) 2980 (w), 1663 (m), 1639 (m), 1607 (m), 1584 (m), 1493 (m), 1452 (m), 1398 (m), 1312 (m), 1233 (m), 1198 (m), 1175 (m), 1138 (m), 1103 (s), 1059 (m), 1038 (m), ^{1}H NMR (500 MHz; D_2O): δ 7.74 (1H, d, *J* 7.8, Ar C^6*H*), 7.61 (1H, dd, *J* 8.3, 7.5, Ar C^4*H*), 7.50 (1H, d, *J* 7.5, Ar C^3*H*), 7.44 (1H, dd, *J* 8.3, 7.8, Ar C^5*H*), 4.23 (1H, dd, *J* 7.3, 5.9, cys-α-C*H*), 4.01 (1H, d, *J* 17.8, gly-C*H*$_2$), 3.90 (1H, d, *J* 17.8, gly-C*H*$_2$), 3.19-3.01 (4H, m, cys-β-C*H*$_2$, linker-C*H*$_2$), 2.90 (2H, dd, *J* 8.8, 7.8, linker-C*H*$_2$); ^{19}F NMR (377 MHz; D_2O): δ

−59.3 (ArC*F*$_3$), −75.6 (OCOC*F*$_3$); ^{13}C NMR (126 MHz; D_2O): δ 173.9 (*C*OOH), 168.5 (cys-*C*O), 163.0 (q, *J* 35.4, O*C*OCF$_3$), 138.2 (*C*^{2}), 138.2 (*C*^{2}), 132.3 (*C*^{4}), 131.4 (*C*^{3}), 127.8 (q, *J* 29.5, *C*^{1}), 126.9 (*C*^{5}), 126.1 (q, *J* 5.7, *C*^{6}), 124.6 (q, *J* 272.5, Ar*C*F$_3$), 116.3 (q, *J* 290.9, OCO*C*F$_3$), 52.3 (cys-α-*C*H), 41.9 (gly-C*H*$_2$), 32.7 (linker-*C*H$_2$), 32.1 (linker-*C*H$_2$), 31.9 (cys-β-*C*H$_2$), 31.9 (cys-β-*C*H$_2$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 351.09789. $C_{14}H_{18}N_2O_3SF_3$ requires M$^+$ 351.09847); *m/z* (ES$^+$) 351 ([M+H]$^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.625 min, Purity: 98.70%.

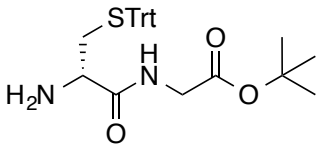
_N_-Fmoc-_S_-trityl-D-cysteinylglycine _tert_-butyl ester (115)



N-Fmoc-D-Cys(Trt)-OH, **114**, (0.50 g, 0.85 mmol, 1.05 eq.), glycine *tert*-butyl ester hydrochloride (0.14 g, 0.81 mmol, 1.0 eq.) and HBTU (0.32 g, 0.85 mmol, 1.05 eq.) were stirred in anhydrous DMF (2 mL) for 5 min. After this time, DIPEA (0.21 g, 0.28 mL, 1.62 mmol, 2.0 eq.) was added and the reaction mixture was stirred under nitrogen at RT for 120 h. The reaction mixture was partitioned over EtOAc (50 mL) and aqueous 0.5 M LiCl (50 mL) and the organic layer was collected, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting crude material was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether and NEt_3 (10:89:1 → 80:19:1; EtOAc : petroleum ether : NEt_3) to yield *N*-Fmoc-*S*-trityl-D-cysteinylglycine *tert*-butyl ester, **115**, (0.50 g, 88%) as a colourless solid: R_f 0.4 (EtOAc/petroleum ether 30:70); mp 85-88 °C (from EtOAc/petroleum ether); $[\alpha]_D^{20}$ = −5.4 (*c* 1.0, $CHCl_3$); ν_{max}/cm^{-1} (neat); 3296 (w), 3057 (w), 1728 (m), 1663 (m), 1595 (m), 1520 (m), 1489 (m), 1445 (m), 1368 (m), 1319 (m), 1217 (m), 1152 (m), 1082 (m), 1034 (m); ^{1}H NMR (500 MHz; $CDCl_3$): δ 7.82-7.18 (23H, m, Ar C*H*), 6.34 (1H, s, gly-N*H*), 5.03 (1H, d, *J* 7.3, cys-N*H*), 4.44-4.36 (2H, m, Fmoc-C*H*$_2$), 4.22 (1H, t, *J* 7.3, 6.8, Fmoc-C*H*),

3.93-3.77 (3H, m, cys-α-C*H*, gly-CH_2), 2.75-2.61 (2H, m, cys-β-CH_2), 1.46 (9H, s, OC(CH_3)$_3$); ^{13}C NMR (126 MHz; $CDCl_3$): δ 170.1 (cys-*C*O), 168.4 (*C*OOC(CH_3)$_3$), 156.1 (Fmoc-*C*ONH), 144.5 (Ar *C*), 143.8 (Ar *C*), 141.4 (Ar *C*), 129.7 (Ar *C*H), 128.2 (Ar *C*H), 127.9 (Ar *C*H), 127.2 (Ar *C*H), 127.1 (Ar *C*H), 125.2 (Ar *C*H), 120.1 (Ar *C*H), 82.5 (O*C*(CH_3)$_3$), 67.5 (Fmoc-*C*H_2), 67.2 (S*C*Ph_3), 54.0 (cys-α-*C*H), 47.2 (Fmoc-*C*H), 42.2 (gly-*C*H_2), 33.9 (cys-β-*C*H_2), 28.1 (OC(*C*H_3)$_3$); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 721.26843. $C_{43}H_{42}N_2NaO_5S$ requires M^+ 721.27066); *m/z* (ES^+) 117 (100%), 721 ($[M+Na]^+$, 3%); Analytical HPLC @ 254 nm Method 2 (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 16.693 min, Purity: 97.97%.

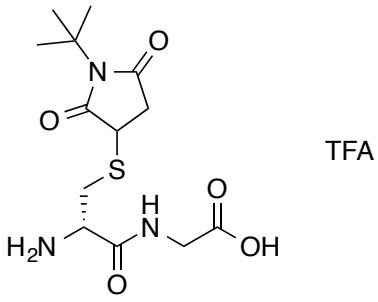
_S_-Trityl-D-cysteinylglycine _tert_-butyl ester (116)



N-Fmoc-*S*-trityl-D-cysteinylglycine *tert*-butyl ester, **115**, (0.4 g, 0.57 mmol, 1.0 eq.) was stirred under argon at RT in 20% piperidine in DMF (3.2 mL) for 4 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether and NEt_3 (12:87:1 → 99:0:1; EtOAc : petroleum ether : NEt_3) to yield *S*-trityl-D-cysteinylglycine *tert*-butyl ester, **116**, (0.13 g, 47%) as a yellow oil: R_f 0.3 (EtOAc/petroleum ether 50:50); $[\alpha]_D^{20}$ = −12.0 (*c* 0.59, $CHCl_3$); ν_{max}/cm^{-1} ($CHCl_3$); 3374 (w), 2980 (m), 2930 (w), 1740 (m), 1672 (m), 1595 (m), 1514 (m), 1491 (m), 1445 (m), 1393 (m), 1368 (m), 1244 (m), 1155 (s), 1082 (m), 1034 (m); ^{1}H NMR (500 MHz; $CDCl_3$): δ 7.51 (1H, t, *J* 5.3, gly-N*H*), 7.48-7.15 (15H, m, Ar C*H*), 3.90 (1H, dd, *J* 18.4, 5.3, gly-CH_AH$_B$), 3.81 (1H, dd, *J* 18.4, 5.3, gly-CH$_A$$H_B$), 3.06-2.96 (1H, m, cys-α-C*H*), 2.82-2.73 (1H, m, cys-β-CH_2), 2.55 (1H, dd, *J* 12.8, 8.8, cys-β-CH_2), 1.45 (9H, s, OC(CH_3)$_3$); ^{13}C NMR (126 MHz; $CDCl_3$): δ 173.2 (cys-*C*O), 169.0 (*C*OOC(CH_3)$_3$), 144.7 (Ar *C*), 129.8 (Ar *C*H),

128.1 (Ar *C*H), 127.0 (Ar *C*H), 82.3 (O*C*(CH_3)$_3$), 67.2 (S*C*Ph$_3$), 54.1 (cys-α-*C*H), 41.9 (gly-*C*H_2), 37.4 (cys-β-*C*H_2), 28.2 (OC(*C*H_3)$_3$); HRMS *m/z* (ES^+) (Found: $[M+H]^+$ 477.22004. $C_{28}H_{33}N_2O_3S$ requires M^+ 477.22064); *m/z* (ES^+) 477 ($[M+H]^+$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 11.734 min, Purity: 95.46%.

_S_-(_N-tert_-Butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate (118)

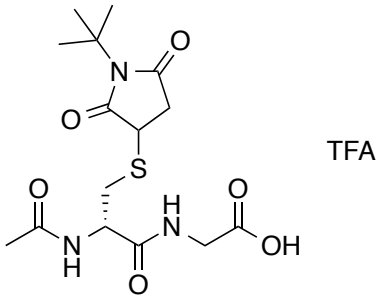


S-Trityl-D-cysteinyl *tert*-butyl ester, **116**, (0.11 g, 0.23 mmol, 1.0 eq.) and triethylsilane (0.06 g, 0.08 mL, 0.5 mmol, 2.2 eq.) were stirred under argon at RT in 20% TFA in CH_2Cl_2 (5.8 mL) for 7 h. After this time the reaction solution was concentrated *in vacuo* and resulting residue was purified by RP C-18 silica gel chromatography, eluting with H_2O to afford D-cysteinylglycine trifluoroacetic acid salt, **117**, as an off-white sticky foam (55.2 mg), which was used without further purification.

To a stirred solution of D-cysteinylglycine trifluoroacetic acid salt, **117**, (35.5 mg) in H_2O (1.0 mL), was added *N-tert*-butyl-maleimide (21.0 mg, 19.8 µL, 0.14 mmol). The resulting solution was stirred at RT for 16 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *S*-(*N-tert*-butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate, **118**, (mixture of diastereoisomers) as a colourless solid (47.1 mg, 71% over two steps): $[\alpha]_D^{20}$ = −26.4 (*c* 1.0, H_2O); ν_{max}/cm^{-1} (neat); 2970 (w), 2941 (w), 1738 (m), 1682 (s), 1460 (m),

1418 (m), 1348 (m), 1265 (m), 1229 (m), 1198 (s), 1167 (s), 1134 (s), 1024 (m); ^{1}H NMR (500 MHz; D_2O): δ 4.39 (0.5H, dd, *J* 7.3, 5.8, cys-α-C*H*), 4.35 (0.5H, dd, *J* 7.3, 5.7, cys-α-C*H*), 4.09 (1H, s, gly-CH_2), 4.08 (1H, s, gly-CH_2), 3.94-3.90 (1H, m, succinimide-C*H*), 3.48 (0.5H, dd, *J* 14.8, 5.9, cys-β-CH_2), 3.41-3.30 (1H, m, cys-β-CH_2), 3.22-3.11 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 2.57 (0.5H, dd, *J* 5.4, 5.0, succinimide-CH_2), 2.54 (0.5H, dd, *J* 5.4, 5.0, succinimide-CH_2), 1.53 (9H, s, NC(CH_3)$_3$); ^{19}F NMR (377 MHz; D_2O): δ −75.6 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.4 (succinimide-*C*O), 180.1 (succinimide-*C*O), 178.8 (succinimide-*C*O), 178.8 (succinimide-*C*O), 172.7 (*C*OOH), 172.7 (*C*OOH), 168.5 (cys-*C*O), 168.5 (cys-*C*O), 162.9 (q, *J* 35.4, O*C*OCF$_3$), 116.3 (q, *J* 291.5, OCO*C*F$_3$), 59.2 (N*C*(CH$_3$)$_3$), 52.5 (cys-α-*C*H), 52.3 (cys-α-*C*H), 41.2 (gly-*C*H$_2$), 41.2 (gly-*C*H$_2$), 41.0 (succinimide-*C*H), 40.1 (succinimide-*C*H), 35.9 (succinimide-*C*H$_2$), 35.6 (succinimide-*C*H$_2$), 32.8 (cys-β-*C*H$_2$), 32.2 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 332.12676. $C_{13}H_{22}N_3O_5S$ requires M$^+$ 332.12747); *m/z* (ES$^+$) 332 ([M+H]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Diastereomer 1: Ret. Time = 8.448 min, Purity: 35.99%; Diastereomer 2: Ret. Time = 8.541 min, Purity: 61.26%.

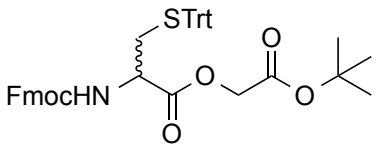
_N_-Acetyl-_S_-(_N-tert_-butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate (119)



To a stirred solution of *S*-(*N-tert*-butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate, **118**, (27.5 mg, 0.06 mmol, 1.0 eq.) and anhydrous NEt$_3$ (12.5 mg, 17.2 µL, 0.12 mmol, 2.0 eq.) in anhydrous THF (0.6 mL) was added acetyl chloride (5.3 mg, 4.8 µL, 0.07 mmol, 1.1 eq.). The reaction mixture was stirred under argon at RT for 4 h. After this time the reaction mixture was

diluted with H_2O (0.5 mL), lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *N*-acetyl-*S*-(*N*-*tert*-butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate, **119**, (mixture of diastereoisomers) as a colourless solid (13.9 mg, 46%): $[\alpha]_D^{20}$ = +15.0 (*c* 1.0, H_2O); ν_{max}/cm^{-1} (neat); 3304 (w), 2980 (w), 1769 (w), 1694 (s), 1659 (m), 1531 (m), 1460 (m), 1402 (m), 1369 (m), 1342 (m), 1263 (m), 1233 (m), 1198 (m), 1163 (m), 1042 (m); ^{1}H NMR (500 MHz; D_2O): δ 4.68-4.61 (1H, m, cys-α-C*H*), 4.01 (2H, s, gly-CH_2), 3.92-3.84 (1H, m, succinimide-C*H*), 3.30 (0.5H, dd, *J* 14.1, 5.1, cys-β-CH_2), 3.24-3.09 (2H, m, cys-β-CH_2, succinimide-CH_2), 2.99 (0.5H, dd, *J* 14.1, 8.7, cys-β-CH_2), 2.57 (0.5H, dd, *J* 4.4, 3.9, succinimide-CH_2), 2.53 (0.5H, dd, *J* 4.4, 3.9, succinimide-CH_2), 2.07 (1.5H, s, COCH_3), 2.06 (1.5H, s, COCH_3), 1.53 (s, 9H, NC(CH_3)$_3$); ^{19}F NMR (377 MHz; D_2O): δ −75.7 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.2 (succinimide-*C*O), 180.2 (succinimide-*C*O), 179.1 (succinimide-*C*O), 179.0 (succinimide-*C*O), 174.4 (*C*OCH$_3$), 174.3 (*C*OCH$_3$), 172.9 (*C*OOH), 172.9 (*C*OOH), 172.4 (cys-*C*O), 172.4 (cys-*C*O), 163.0 (q, *J* 35.1, O*C*OCF$_3$), 116.3 (q, *J* 290.9, OCO*C*F$_3$), 59.2 (N*C*(CH$_3$)$_3$), 59.1 (N*C*(CH$_3$)$_3$), 53.1 (cys-α-*C*H), 52.7 (cys-α-*C*H), 41.1 (gly-*C*H$_2$), 40.8 (succinimide-*C*H), 40.0 (succinimide-*C*H), 36.2 (succinimide-*C*H$_2$), 35.9 (succinimide-*C*H$_2$), 32.7 (cys-β-*C*H$_2$), 32.3 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$), 21.7 (CO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 374.13761. $C_{15}H_{24}N_3O_6S$ requires M$^+$ 374.13803); *m/z* (ES$^+$) 374 ([M+H]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.237 min, Purity: 97.26%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Diastereomer 1: Ret. Time = 40.368 min, 53.55%; Diastereomer 2: Ret. Time = 51.065 min, 44.86%.

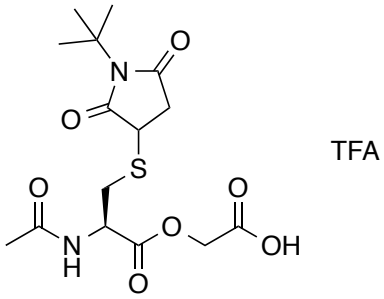
2-(*tert*-Butoxy)-2-oxoethyl *N*-Fmoc-*S*-trityl-L-cysteinate (120)



To a stirred mixture of *N*-Fmoc-L-Cys(Trt)-OH, **104**, (0.10 g, 0.17 mmol, 1.0 eq.) in anhydrous THF (3.4 mL) was added PyBOP (0.12 g, 0.22 mmol, 1.3 eq.) and anhydrous NEt_3 (0.05 g, 0.07 mL, 0.50 mmol, 2.95 eq.). After 1 h *tert*-butyl-2-hydroxyacetate (28.0 mg, 0.21 mmol, 1.2 eq.) was added dropwise over 5 min in anhydrous THF (1 mL). The reaction mixture was stirred under argon at RT for 20.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting crude material was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (5:95 → 40:60; EtOAc : petroleum ether) to yield 2-(*tert*-butoxy)-2-oxoethyl *N*-Fmoc-*S*-trityl-L-cysteinate, **120**, (racemic mixture) as a colourless solid (97.2 mg, 81%): R_f 0.4 (EtOAc/petroleum ether 20:80); $[\alpha]_D^{20}$ = −0.6 (*c* 1.0, $CHCl_3$); ν_{max}/cm^{-1} 2980 (w), 1746 (m), 1721 (m), 1595 (w), 1504 (m), 1491 (m), 1447 (m), 1422 (m), 1393 (m), 1383 (m), 1369 (m), 1315 (m), 1248 (m), 1211 (m), 1182 (m), 1148 (s), 1105 (m), 1057 (m), 1034 (m), 1001 (m); ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.86-7.77 (2H, m, Ar C*H*), 7.70-7.62 (2H, m, Ar C*H*), 7.51-7.41 (8H, m, Ar C*H*), 7.39-7.22 (11H, m, Ar C*H*), 5.26 (1H, d, *J* 8.0, N*H*), 4.55 (2H, s, OCH_2CO), 4.46-4.38 (3H, m, Fmoc CHCH_2O, cys-α-C*H*), 4.32-4.25 (1H, m, Fmoc C*H*CH_2O), 2.89-2.71 (2H, m, cys-β-CH_2), 1.50 (9H, s, $CO_2C(CH_3)_3$); ^{13}C NMR (101 MHz; $CDCl_3$): δ 170.1 (cys-*C*O), 166.1 ($CO_2C(CH_3)_3$), 155.7 (Fmoc *C*O), 144.4 (Ar *C*), 144.0 (Ar *C*), 143.8 (Ar *C*), 141.4 (Ar *C*), 129.7 (Ar *C*H), 128.1 (Ar C*H*), 127.8 (Ar C*H*), 127.2 (Ar C*H*), 127.0 (Ar C*H*), 125.3 (Ar C*H*), 120.1 (Ar C*H*), 82.8 ($CO_2C(CH_3)_3$), 67.3 (S*C*Ph_3), 67.2 (Fmoc CHCH_2O), 62.0 (OCH_2CO), 53.0 (cys-α-*C*H), 47.2 (Fmoc *C*HCH_2O), 33.9 (cys-β-*C*H_2), 28.1 ($CO_2C(CH_3)_3$); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 722.25295. $C_{43}H_{41}NNaO_6S$ requires M^+ 722.25468); *m/z* (ES^+) 359 (½$[M+H+NH_4]^+$, 100%); Analytical HPLC @ 254 nm Method 2 (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 17.672 min, Purity: 97.64%;

Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Enantiomer 1: Ret. Time = 23.945 min, 45.18%; Enantiomer 2: Ret. Time = 30.247 min, 54.62%.

2-((*N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinyl)oxy)acetic acid 2,2,2-trifluoroacetate (125)



To a stirred solution of Boc-L-Cys(Trt)-OH, **121**, (0.10 g, 0.22 mmol, 1.0 eq.) and *tert*-butyl-2-hydroxyacetate (55.6 mg, 0.42 mmol, 1.95 eq.) in anhydrous $CHCl_3$ / anhydrous MeCN (1:2.8, 1.9 mL) at 0 °C was added *N,N'*-dicyclohexylcarbodiimide (86.8 mg, 0.42 mmol, 1.95 eq.). The resulting solution was stirred under argon at 0 °C for a further 2 h before being stirred at RT for 15.5 h. After this time, the reaction mixture was filtered under reduced pressure, the filtrate was concentrated *in vacuo* and the resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (5:95 → 40:60; EtOAc : petroleum ether) to yield 2-(*tert*-butoxy)-2-oxoethyl *N*-Boc-*S*-trityl-L-cysteinate, **122**, as an impure oil (0.11 g), which was used without further purification.

2-(*tert*-Butoxy)-2-oxoethyl *N*-Boc-*S*-trityl-L-cysteinate, **122**, (0.11 g) and triethylsilane (0.08 g, 0.11 mL, 0.69 mmol) were stirred under argon at RT in 20% TFA in CH_2Cl_2 (8.0 mL) for 2 h. After this time the reaction solution was concentrated *in vacuo* and the resulting residue was purified by RP C-18 silica gel chromatography, eluting with H_2O to afford 2-((L-cysteinyl)oxy)acetic acid

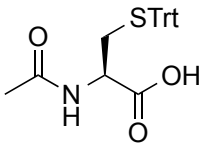
2,2,2-trifluoroacetate, **123**, as an impure colourless solid (32.2 mg), which was used without further purification.

To a stirred solution of 2-((L-cysteinyl)oxy)acetic acid 2,2,2-trifluoroacetate, **123**, (32.2 mg) in H_2O (3.6 mL) was added *N-tert*-butyl-maleimide (27.5 mg, 26.0 µL, 0.18 mmol). The resulting solution was stirred at RT for 7 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford 2-((*S*-(*N-tert*-butylsuccinimido) L-cysteinyl)oxy)acetic acid 2,2,2-trifluoroacetate, **124**, (mixture of diastereoisomers) as a colourless solid (25.5 mg), which was used without further purification.

To a stirred solution of 2-((*S*-(*N-tert*-butylsuccinimido) L-cysteinyl)oxy)acetic acid 2,2,2-trifluoroacetate, **124**, (25.5 mg) and anhydrous NEt_3 (11.5 mg, 15.9 µL, 0.11 mmol) in anhydrous THF (0.5 mL) was added acetyl chloride (5.0 mg, 4.5 µL, 0.06 mmol). The reaction mixture was stirred under argon at RT for 3.5 h. After this time, the reaction mixture was diluted with H_2O, lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford 2-((*N*-acetyl-*S*-(*N-tert*-butylsuccinimido) L-cysteinyl)oxy)acetic acid 2,2,2-trifluoroacetate, **125**, (mixture of diastereoisomers) as a colourless solid (11.6 mg, 11% over four steps): $[\alpha]_D^{20}$ = −24.0 (*c* 0.75, H_2O); ν_{max}/cm^{-1} (neat); 2980 (m), 1746 (m), 1694 (s), 1541 (m), 1460 (w), 1420 (m), 1371 (m), 1344 (m), 1263 (m), 1163 (s), 1053 (m); ^{1}H NMR (500 MHz; D_2O): δ 4.86-4.73 (3H, m, cys-α-C*H*, OC$H_2$$CO_2$H), 3.94-3.85 (1H, m, succinimide-C*H*), 3.42 (0.5H, dd, *J* 14.2, 4.9, cys-β-CH_2), 3.31 (0.5H, dd, *J* 14.2, 4.9, cys-β-CH_2), 3.25-3.10 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 3.06 (0.5H, dd, *J* 14.2, 8.9, cys-β-CH_2), 2.58 (0.5H, dd, *J* 4.4, 3.2, succinimide-CH_2) 2.54 (0.5H, dd, *J* 4.4, 3.2, succinimide-CH_2), 2.07 (1.5H, s, COCH_3), 2.06 (1.5H, s, COCH_3), 1.53 (4.5H, s, NC(CH_3)$_3$), 1.53 (4.5H, s, NC(CH_3)$_3$); ^{19}F NMR (377 MHz; D_2O): δ

−75.7 (OCOC*F*$_3$); ^{13}C NMR (126 MHz; D_2O): δ 180.1 (succinimide-*C*O), 180.0 (succinimide-*C*O), 179.0 (succinimide-*C*O), 179.0 (succinimide-*C*O), 174.3 (*C*OCH$_3$), 174.2 (*C*OCH$_3$), 171.2 (cys-*C*O), 171.2 (cys-*C*O), 171.1 (*C*O$_2$H), 162.9 (q, *J* 35.3, O*C*OCF$_3$), 116.3 (q, *J* 290.7, OCO*C*F$_3$), 61.9 (O*C*H$_2$CO$_2$H), 59.2 (N*C*(CH$_3$)$_3$), 59.2 (N*C*(CH$_3$)$_3$), 52.5 (cys-α-*C*H), 52.0 (cys-α-*C*H), 40.9 (succinimide-*C*H), 39.9 (succinimide-*C*H), 36.2 (succinimide-*C*H$_2$), 35.8 (succinimide-*C*H$_2$), 32.2 (cys-β-*C*H$_2$), 31.7 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$), 21.6 (CO*C*H$_3$), 21.5 (CO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 375.12147. $C_{15}H_{23}O_7N_2S$ requires M$^+$ 375.12205); *m/z* (ES$^+$) 375 ([M+H]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.776 min, Purity: 95.47%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Diastereomer 1: Ret. Time = 20.625 min, 50.48%; Diastereomer 2: Ret. Time = 29.359 min, 49.32%.

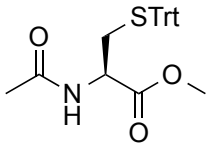
_N_-Acetyl-_S_-trityl-L-cysteine (127)



To a stirred solution of *N*-acetyl-L-cysteine, **126**, (1.00 g, 6.13 mmol, 1.0 eq.) in anhydrous DMF (16 mL) was added trityl chloride (2.56 g, 9.19 mmol, 1.5 eq.). The resulting reaction mixture was stirred under argon at RT for 20 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was partitioned between diethyl ether (50 mL) and aqueous 0.5 M LiCl (50 mL). The aqueous layer was extracted with diethyl ether (2 × 50 mL) and the combined organic extracts were dried over anhydrous $MgSO_4$, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, MeOH and H_2O (100:0:0 → 66.7:16.7:16.7; EtOAc : MeOH : H_2O) to yield *N*-acetyl-*S*-trityl-L-cysteine, **127**, (1.21 g, 49%) as a colourless solid:

R_f 0.6 (EtOAc/MeOH/H_2O 4:1:1); $[\alpha]_D^{20}$ = +33.9 (*c* 1.0, $CHCl_3$) {lit.[34] $[\alpha]_D^{20}$ = +40.7 (*c* 1.0, $CHCl_3$)}; mp 76-80 °C (from EtOAc/MeOH/H_2O); ν_{max}/cm^{-1} 3055 (w), 1724 (m), 1624 (m), 1522 (m), 1489 (m), 1443 (m), 1373 (m), 1302 (m), 1213 (m), 1182 (m), 1121 (m), 1082 (m), 1034 (m), 1001 (m); ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.37-7.28 (6H, m, Ar C*H*), 7.25-7.17 (6H, m, Ar C*H*), 7.17-7.09 (3H, m, Ar C*H*), 5.87 (1H, d, *J* 7.4, N*H*), 4.43-4.31 (1H, m, cys-α-C*H*), 2.69 (1H, dd, *J* 12.7, 6.2, cys-β-CH_2), 2.60 (1H, dd, *J* 12.7, 4.7, cys-β-CH_2), 1.87 (3H, s, CH_3); ^{13}C NMR (126 MHz; $CDCl_3$): δ 173.3 (*C*OOH), 171.3 (*C*OCH_3), 144.3 (Ar *C*), 129.6 (Ar *C*H), 128.2 (Ar *C*H), 127.1 (Ar *C*H), 67.3 (S*C*Ph_3), 51.5 (cys-α-*C*H), 33.2 (cys-β-*C*H_2), 23.0 (*C*H_3); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 428.12922. $C_{24}H_{23}NNaO_3S$ requires M^+ 428.12909); *m/z* (ES^-) 404 ($[M-H]^-$, 100%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 12.678 min, Purity: 98.90%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Ret. Time = 14.947 min, 98.96%. The data are in good agreement with the literature values.[35]

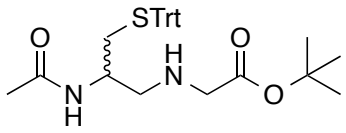
Methyl *N*-acetyl-*S*-trityl-L-cysteinate (128)



To a stirred solution of *N*-acetyl-*S*-trityl-L-cysteine, **127**, (0.50 g, 1.23 mmol, 1.0 eq.) in MeOH (9.9 mL) at 0 °C was added thionyl chloride (0.18 g, 0.11 mL, 1.48 mmol, 1.2 eq.) dropwise over 5 min and the reaction mixture was stirred under argon at RT for 3.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was partitioned between brine (50 mL) and CH_2Cl_2 (50 mL). The aqueous layer was extracted with CH_2Cl_2 (2 × 50 mL), and the organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting crude material was adsorbed onto Celite® and purified by silica gel

chromatography, eluting with EtOAc and petroleum ether (12:88 → 80:20; EtOAc : petroleum ether) to yield *N*-acetyl-*S*-trityl-L-cysteine methyl ester, **128**, as a colourless solid (0.34 g, 66%): R_f 0.3 (EtOAc/petroleum ether 50:50); $[\alpha]_D^{20}$ = +52.9 (*c* 1.0, $CHCl_3$); mp 147-150 °C (from EtOAc/petroleum ether); ν_{max}/cm^{-1} 3057 (w), 3017 (w), 2970 (w), 2949 (w), 1746 (s), 1670 (m), 1636 (m), 1591 (m), 1557 (m), 1487 (m), 1443 (m), 1371 (m), 1306 (m), 1271 (m), 1229 (m), 1217 (m), 1167 (m), 1155 (m), 1126 (m), 1082 (m), 1032 (m), 1011 (m), 1001 (m); $[\alpha]_D^{20}$ = +52.9 (*c* 1.0, $CHCl_3$); ^{1}H NMR (500 MHz; $CDCl_3$): δ 7.34-7.29 (6H, m, Ar C*H*), 7.24-7.18 (6H, m, Ar C*H*), 7.18-7.12 (3H, m, Ar C*H*), 5.82 (1H, d, *J* 7.7, N*H*), 4.57-4.50 (1H, m, cys-α-C*H*), 3.64 (3H, s, COOCH_3), 2.62 (1H, dd, *J* 12.3, *J* 5.6, cys-β-CH_2), 2.56 (1H, dd, *J* 12.3, *J* 4.7, cys-β-CH_2), 1.88 (3H, s, COCH_3); ^{13}C NMR (126 MHz; $CDCl_3$): δ 171.1 (*C*OOCH$_3$), 169.7 (*C*OCH$_3$), 144.4 (Ar *C*), 129.6 (Ar *C*H), 128.1 (Ar *C*H), 127.1 (Ar *C*H), 67.1 (S*C*Ph$_3$), 52.8 (COO*C*H$_3$), 51.2 (cys-α-*C*H), 33.9 (cys-β-*C*H$_2$), 23.2 (CO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+Na]$^+$ 442.14351. $C_{25}H_{25}NNaO_3S$ requires M$^+$ 442.14474); *m/z* (ES$^+$) 243 (100%), 442 ([M+Na]$^+$, 60%); Analytical HPLC @ 254 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 13.719 min, Purity: 96.42%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Ret. Time = 17.780 min, 99.10%.

_N_-Acetyl-_S_-trityl-_tert_-butyl (2-amino-3-mercaptopropyl)glycinate (130)



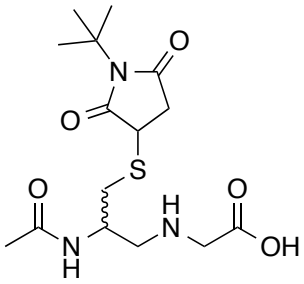
To a solution of *N*-acetyl-*S*-trityl-L-cysteine methyl ester, **128**, (0.38 g, 0.90 mmol, 1.0 eq.) in anhydrous toluene (9.0 mL) was added a 1 M solution of diisobutyl aluminum hydride in toluene (0.97 mL, 0.97 mmol, 1.1 eq.) and the resulting reaction mixture was stirred under argon at −60 °C for 2 h. After this time, methanol (10 mL) followed by aqueous 0.8 M sodium potassium tartrate (10 mL) was added dropwise over 5 min at −60 °C and the reaction mixture was warmed

to RT. Diethyl ether (50 mL) was added, the organic layer was separated and the aqueous layer was extracted with diethyl ether (2 × 50 mL). The combined organic layers were washed with brine (50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite®, and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (12:88 → 100:0; EtOAc : petroleum ether), to yield slightly impure *N*-(1-oxo-3-(tritylthio)propan-2-yl)acetamide, **129**, (94.2 mg), which was used without further purification.

To a stirred solution of *N*-(1-oxo-3-(tritylthio)propan-2-yl)acetamide, **129**, (94.2 mg) in anhydrous methanol (1.4 mL) was added glycine *tert*-butyl ester hydrochloride (0.12 g, 0.73 mmol), and $NaBH_3CN$ (15.2 mg, 0.24 mmol) in methanol (1 mL) at 0 °C. The reaction mixture was stirred under argon at RT for 17 h. After this time the reaction mixture was concentrated *in vacuo* and the residue was partitioned between H_2O (25 mL) and CH_2Cl_2 (50 mL). The aqueous layer was extracted with CH_2Cl_2 (4 × 25 mL), and the organic layers were combined, washed with brine (50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with MeOH and $CHCl_3$ (0:100 → 5:95; MeOH : $CHCl_3$) to afford *N*-acetyl-*S*-trityl-*tert*-butyl (2-amino-3-mercaptopropyl)glycinate, **130**, (36.2 mg, 8% over two steps) as a colourless oil: R_f 0.4 (MeOH/$CHCl_3$ 5:95); $[\alpha]_D^{20}$ = +0.7 (*c* 1.0, $CHCl_3$); ν_{max}/cm^{-1} ($CHCl_3$) 3289 (w), 3055 (w), 2980 (w), 2930 (w), 1732 (m), 1651 (m), 1595 (w), 1547 (m), 1489 (m), 1445 (m), 1393 (m), 1368 (m), 1288 (m), 1246 (m), 1153 (s), 1082 (m), 1034 (m), 1001 (w); ^{1}H NMR (500 MHz; $CDCl_3$): δ 7.40-7.08 (15H, m, Ar C*H*), 5.92 (1H, d, *J* 7.8, N*H*COCH$_3$), 3.91-3.81 (1H, m, SCH$_2$C*H*CH$_2$), 3.17-3.04 (2H, m, gly-C*H*$_2$), 2.66 (1H, dd, *J* 12.4, 5.7, SCH$_2$CHC*H*$_2$), 2.49 (1H, dd, *J* 12.4, 5.3, SC*H*$_2$CHCH$_2$), 2.38 (1H, dd, *J* 12.4, 4.9, SCH$_2$CHC*H*$_2$), 2.24 (1H, dd, *J* 12.4, 4.9, SC*H*$_2$CHCH$_2$), 1.83 (3H, s, NHCOC*H*$_3$), 1.37 (9H, s, OC(C*H*$_3$)$_3$); ^{13}C NMR (126 MHz; $CDCl_3$): δ 171.8 (*C*OOC(CH$_3$)$_3$), 169.9 (NH*C*OCH$_3$), 144.7 (Ar *C*), 129.7 (Ar *C*H), 128.1 (Ar *C*H), 126.9 (Ar *C*H), 81.5 (O*C*(CH$_3$)$_3$),

66.9 (S*C*Ph$_3$), 51.7 (gly-*C*H$_2$), 51.1 (SCH$_2$CH*C*H$_2$), 48.4 (SCH$_2$*C*HCH$_2$), 34.0 (S*C*H$_2$CHCH$_2$), 28.2 (OC(*C*H$_3$)$_3$), 23.5 (NHCO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 505.25070 $C_{30}H_{37}N_2O_3S$ requires M$^+$ 505.25194); *m/z* (ES$^+$) 505 ([M+H]$^+$, 100%); Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Ret. Time = 16.576 min, 88.18%.

_S_-(_N-tert_-Butylsuccinimido) (2-acetamido-3-mercaptopropyl)glycine 2,2,2-trifluoroacetate (132)

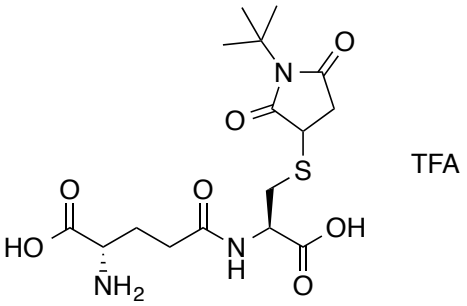


N-Acetyl-*S*-trityl-*tert*-butyl (2-amino-3-mercaptopropyl)glycinate, **130**, (34.6 mg, 0.07 mmol, 1.0 eq.) and triethylsilane (25.3 mg, 34.7 µL, 0.22 mmol, 3.2 eq.) were stirred under argon at RT in 20% TFA in CH_2Cl_2 (2.5 mL) for 2.5 h. After this time the reaction solution was concentrated *in vacuo* and resulting residue was purified by RP C-18 silica gel chromatography, eluting with H_2O to afford (2-acetamido-3-mercaptopropyl)glycine 2,2,2-trifluoroacetate, **131**, as a colourless solid (13.4 mg), which was used without further purification.

To a stirred solution of (2-acetamido-3-mercaptopropyl)glycine 2,2,2-trifluoroacetate, **131**, (13.4 mg) in H_2O (0.48 mL) was added *N*-*tert*-butylmaleimide (10.0 mg, 9.4 µL, 0.06 mmol). The resulting solution was stirred at RT for 8 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *S*-(*N*-*tert*-butylsuccinimido) (2-acetamido-3-mercaptopropyl)glycine 2,2,2-trifluoroacetate, **132**, (mixture of

diastereoisomers) as a colourless solid (5.4 mg, 17% over two steps): $[\alpha]_D^{20}$ = +0.5 (*c* 0.59, H_2O); ν_{max}/cm^{-1} (neat) 3278 (w), 2980 (w), 1771 (w), 1694 (s), 1546 (m), 1456 (m), 1419 (m), 1370 (m), 1344 (m), 1264 (m), 1197 (s), 1166 (s), 1131 (s), 1043 (m); ^{1}H NMR (500 MHz; D_2O): δ 4.50-4.36 (1H, m, SCH_2CHCH_2), 3.94-3.82 (3H, m, gly-CH_2, succinimide-C*H*), 3.46 (0.5H, dd, *J* 3.8, 3.0, SCH_2CHCH_2), 3.43 (0.5H, dd, *J* 3.8, 3.0, SCH_2CHCH_2), 3.28-3.06 (2.5, m, SCH_2CHCH_2, SCH_2CHCH_2, succinimide-CH_2), 2.98 (1H, d, *J* 6.8, SCH_2CHCH_2), 2.83 (0.5H, dd, *J* 13.5, 8.7, SCH_2CHCH_2), 2.59-2.48 (1H, m, succinimide-CH_2), 2.06 (1.5H, s, NHCOCH_3), 2.05 (1.5H, s, NHCOCH_3), 1.53 (4.5H, s, $NC(CH_3)_3$), 1.53 (4.5H, s, $NC(CH_3)_3$); ^{19}F NMR (376.6 MHz; D_2O): δ −75.6 ($OCOCF_3$); ^{13}C NMR (126 MHz; D_2O): δ 180.2 (succinimide-*C*O), 180.1 (succinimide-*C*O), 179.0 (succinimide-*C*O), 179.0 (succinimide-*C*O), 175.1 ($NHCOCH_3$), 175.0 ($NHCOCH_3$), 169.3 (*C*OOH), 163.0 (q, *J* 35.7, $OCOCF_3$), 116.3 (q, *J* 291.6, $OCOCF_3$), 59.2 ($NC(CH_3)_3$), 59.1 ($NC(CH_3)_3$), 50.0 (SCH_2CHCH_2), 50.0 (SCH_2CHCH_2), 48.0 (gly-CH_2), 46.6 (SCH_2CHCH_2), 46.0 (SCH_2CHCH_2), 40.6 (succinimide-*C*H), 39.9 (succinimide-*C*H), 36.0 (succinimide-*C*H_2), 35.9 (succinimide-*C*H_2), 33.2 (SCH_2CHCH_2), 33.1 (SCH_2CHCH_2), 27.3 ($NC(CH_3)_3$), 22.0 ($NHCOCH_3$); HRMS *m/z* (ES^+) (Found: $[M+H]^+$ 360.15789 $C_{15}H_{26}N_3O_5S$ requires M^+ 360.15877); *m/z* (ES^+) 360 ($[M+H]^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 8.542 min, Purity: 98.97%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Peak 1: Ret. Time = 19.990 min, 62.48%; Peak 2: Ret. Time = 26.366 min, 35.50%.

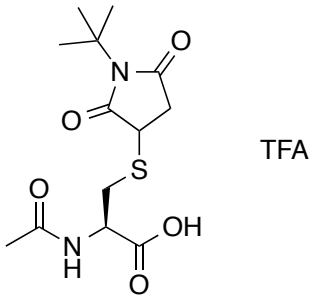
_S_-(_N-tert_-Butylsuccinimido) γ-L-glutamyl-L-cysteine 2,2,2-trifluoroacetate (134)



To a stirred solution of γ-L-glu-L-cys, **133**, (15.3 mg, 0.06 mmol, 1.0 eq.) and NaOH (2.4 mg, 0.06 mmol, 1.0 eq.) in H_2O (0.4 mL) was added *N-tert*-butyl-maleimide (9.4 mg, 8.8 µL, 0.06 mmol, 1.0 eq.). The resulting solution was stirred at RT for 4.5 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *S*-(*N-tert*-butylsuccinimido) γ-L-glutamyl-L-cysteine, **134**, (mixture of diastereoisomers) as a colourless solid (19.8 mg, 63%): R_f 0.8 (MeCN/H_2O 50:50 (RP)); $[\alpha]_D^{20}$ = −6.4 (*c* 1.1, H_2O); ν_{max}/cm^{-1} (neat) 1739 (s), 1368 (w), 1216 (w); ^{1}H NMR (500 MHz; D_2O): δ 4.71-4.63 (1H, m, cys-α-C*H*), 4.04 (1H, dd, *J* 7.0, 6.5, glu-α-C*H*), 3.92-3.85 (1H, m, succinimide-C*H*), 3.37 (0.5H, dd, *J* 14.1, 4.6, cys-β-CH_2), 3.27 (0.5H, dd, *J* 14.1, 4.6, cys-β-CH_2), 3.23-3.10 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 3.04 (0.5H, dd, *J* 14.1, 8.7, cys-β-CH_2), 2.64-2.52 (3H, m, glu-γ-CH_2, succinimide-CH_2), 2.31-2.15 (2H, m, glu-β-CH_2), 1.54 (4.5H, s, NC(CH_3)$_3$), 1.53 (4.5H, s, NC(CH_3)$_3$); ^{19}F NMR (376.6 MHz; D_2O): δ −75.7 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.1 (succinimide-*C*O), 180.0 (succinimide-*C*O), 179.1 (succinimide-*C*O), 179.0 (succinimide-*C*O), 174.3 (glu-*C*O), 174.3 (glu-*C*O), 173.5 (cys-*C*OOH), 173.5 (cys-*C*OOH), 172.0 (glu-*C*OOH), 163.0 (q, *J* 35.3, O*C*OCF$_3$), 116.3 (q, *J* 291.5, OCO*C*F$_3$), 59.2 (N*C*(CH$_3$)$_3$), 59.2 (N*C*(CH$_3$)$_3$), 52.6 (glu-α-*C*H), 52.2 (cys-α-*C*H), 40.9 (succinimide-*C*H), 39.9 (succinimide-*C*H), 36.3 (succinimide-*C*H$_2$), 35.8 (succinimide-*C*H$_2$), 32.3 (cys-β-*C*H$_2$), 31.8 (cys-β-*C*H$_2$), 31.0 (glu-γ-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$), 25.7 (glu-β-*C*H$_2$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 404.14836. $C_{16}H_{26}O_7N_3S$ requires M$^+$ 404.14860); *m/z* (ES$^+$) 404 ([M+H]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column;

95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.094 min, Purity: 98.57%.

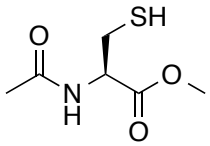
_N_-Acetyl-_S_-(_N-tert_-butylsuccinimido) L-cysteine 2,2,2-trifluoroacetate (135)



To a stirred solution of *N*-acetyl-L-cysteine, **126**, (20.0 mg, 0.12 mmol, 1.0 eq.) and NaOH (4.9 mg, 0.12 mmol, 1.0 eq.) in H_2O (0.9 mL) was added *N-tert*-butyl-maleimide (18.7 mg, 17.7 µL, 0.12 mmol, 1.0 eq.). The resulting solution was stirred at RT for 3.5 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *N*-acetyl-*S*-(*N-tert*-butylsuccinimido) L-cysteine 2,2,2-trifluoroacetate, **135**, (mixture of diastereoisomers) as a pale yellow solid (35.6 mg, 67%): R_f 0.8 (MeCN/H_2O 50:50 (RP)); $[\alpha]_D^{20}$ = −17.3 (*c* 1.0, MeOH); ν_{max}/cm^{-1} (neat) 3329 (w), 2970 (w), 1738 (s), 1719 (s), 1697 (s), 1612 (m), 1558 (m), 1423 (w), 1366 (m), 1348 (m), 1315 (w), 1263 (m), 1229 (m), 1217 (s), 1206 (m), 1161 (m), 1117 (w), 1038 (w); ^{1}H NMR (500 MHz; D_2O): δ 4.70-4.63 (1H, m, cys-α-C*H*), 3.92-3.85 (1H, m, succinimide-C*H*), 3.36 (0.2H, dd, *J* 14.1, 4.6, cys-β-CH_2), 3.29-3.10 (2.6H, m, cys-β-CH_2, succinimide-CH_2), 3.04 (0.2H, dd, *J* 14.1, 8.6, cys-β-CH_2), 2.61-2.51 (1H, m, succinimide-CH_2), 2.06 (0.4H, s, CH_3), 2.05 (2.6H, s, CH_3), 1.53 (9H, s, NC(CH_3)$_3$); ^{19}F NMR (376.6 MHz; D_2O): δ −75.7 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.2 (succinimide-*C*O), 180.0 (succinimide-*C*O), 179.1 (succinimide-*C*O), 179.0 (succinimide-*C*O), 174.2 (*C*OCH$_3$), 174.2 (*C*OCH$_3$), 173.5 (*C*OOH), 162.9 (q, *J* 35.3, O*C*OCF$_3$), 116.3 (q, *J* 291.5, OCO*C*F$_3$), 59.2 (N*C*(CH$_3$)$_3$), 52.6 (cys-α-*C*H), 52.1 (cys-α-*C*H), 40.9 (succinimide-*C*H), 39.9 (succinimide-*C*H), 36.2 (succinimide-*C*H$_2$), 35.8 (succinimide-*C*H$_2$), 32.4 (cys-β-*C*H$_2$),

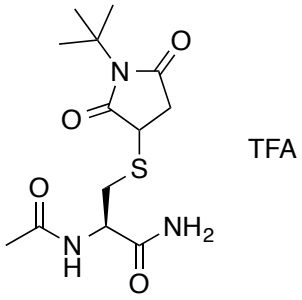
31.8 (cys-β-*C*H_2), 27.3 (NC(*C*H_3)$_3$), 21.6 (*C*H_3); HRMS *m/z* (ES^+) (Found: $[M+H]^+$ 317.11603. $C_{13}H_{21}O_5N_2S$ requires M^+ 317.11657); *m/z* (ES^+) 317 ($[M+H]^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.847 min, Purity: 98.48%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Ret. Time = 20.397 min, 99.99%.

_N_-Acetyl-L-cysteine methyl ester (136)



To a stirred solution of *N*-acetyl-L-cysteine, **126**, (1.00 g, 6.13 mmol, 1.0 eq.) in MeOH (22 mL) was added thionyl chloride (0.85 g, 0.52 mL, 7.11 mmol, 1.16 eq.) dropwise over 5 min and the reaction mixture was stirred under argon at RT for 1.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was partitioned between brine (50 mL) and EtOAc (50 mL). The aqueous layer was extracted with EtOAc (3 × 50 mL), and the organic layers were combined, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting crude material was adsorbed onto Celite® and purified by silica gel chromatography, eluting with MeOH and $CHCl_3$ (1:99 → 10:90; MeOH : $CHCl_3$) to yield *N*-acetyl-L-cysteine methyl ester, **136**, as a colourless solid (0.49 g, 45%): R_f 0.6 (MeOH/$CHCl_3$ 5:95); mp 79-81 °C (from MeOH/$CHCl_3$) [lit.[36] 78-80 °C (EtOAc)]; $[\alpha]_D^{20}$ = +75.5 (*c* 1.0, $CHCl_3$) {lit.[35,36] $[\alpha]_D$ = +71.0 (*c* 1.0, $CHCl_3$)}; ^{1}H NMR (400 MHz; $CDCl_3$): δ 6.68 (1H, d, *J* 5.8, N*H*), 4.86-4.78 (1H, m, cys-α-C*H*), 3.72 (3H, s, COOC*H*$_3$), 2.93 (2H, dd, *J* 9.0, 4.4, cys-β-C*H*$_2$), 2.01 (3H, s, COC*H*$_3$), 1.36 (1H, t, *J* 8.8, S*H*); *m/z* (ES^+) 178 ($[M+H]^+$, 100%). The data are in good agreement with literature values.[35,36]

_N_-Acetyl-_S_-(_N-tert_-butylsuccinimido) L-cysteinamide 2,2,2-trifluoroacetate (138)

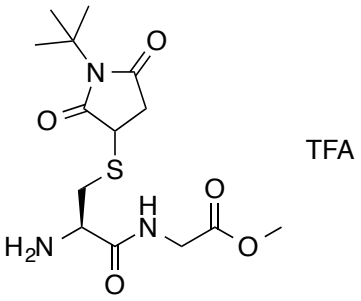


N-acetyl-L-cysteine methyl ester, **136**, (0.40 g, 2.26 mmol, 1.0 eq.) was stirred in a mixture of toluene (3 mL) and ammonium hydroxide (3 mL, 25% solution in H_2O) under argon at RT for 19 h. After this time the reaction mixture was concentrated *in vacuo* at 60 °C to afford a mixture of *N*-acetyl-L-cysteinamide and oxidised *N*-acetyl-L-cysteinamide (0.37 g) as a colourless solid, which was used without further purification.

The colourless solid (20.0 mg) containing *N*-acetyl-L-cysteinamide, **137a**, and oxidised *N*-acetyl-L-cysteinamide, **137b**, was stirred with triphenylphosphine, polymer-bound (79.0 mg, ~3 mmol/g triphenylphosphine loading - Sigma Aldrich) in DMF (0.5 mL) at RT for 4 h. After this time, the reaction mixture was filtered and washed through with DMF (0.5 mL). *N-tert*-Butyl-maleimide (9.2 mg, 8.7 µL, 0.06 mmol) and NEt_3 (8.4 µL) were added to the filtrate, which was stirred at RT for 144 h. After this time, the reaction mixture was concentrated *in vacuo* and the crude reaction mixture was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *N*-acetyl-*S*-(*N-tert*-butylsuccinimido) L-cysteinamide 2,2,2-trifluoroacetate, **138**, as a colourless solid (3.3 mg, 6% over two steps): $[\alpha]_D^{20}$ = −11.4 (*c* 0.33, H_2O); ν_{max}/cm^{-1} (neat); 3391 (w), 3316 (w), 2970 (w), 1738 (w), 1699 (s), 1636 (s), 1522 (m), 1458 (w), 1425 (m), 1398 (w), 1373 (m), 1346 (m), 1312 (w), 1281 (w), 1263 (m), 1229 (w), 1198 (m), 1167 (m), 1153 (m), 1115 (w), 1042 (w); ^{1}H NMR (500 MHz; D_2O): δ 4.61-4.54 (1H, m, cys-α-C*H*), 3.92-3.84 (1H, m, succinimide-C*H*), 3.36-3.10 (2.5H, m, cys-β-CH_2, succinimide-CH_2), 3.00 (0.5H, dd, *J* 14.1, 8.8, cys-β-CH_2),

2.58 (0.5H, dd, *J* 4.7, 4.3, succinimide-C*H*$_2$), 2.54 (0.5H, dd, *J* 4.7, 4.3, succinimide-C*H*$_2$), 2.07 (1.5H, s, C*H*$_3$), 2.07 (1.5H, s, C*H*$_3$), 1.54 (9H, s, NC(C*H*$_3$)$_3$); ^{19}F NMR (376.6 MHz; D_2O): δ −75.6 (OCOC*F*$_3$); ^{13}C NMR (126 MHz; D_2O): δ 180.2 (succinimide-*C*O), 180.2 (succinimide-*C*O), 179.1 (succinimide-*C*O), 179.0 (succinimide-*C*O), 174.6 (*C*ONH$_2$), 174.6 (*C*ONH$_2$), 174.4 (*C*OCH$_3$), 174.3 (*C*OCH$_3$), 59.2 (N*C*(CH$_3$)$_3$), 59.1 (N*C*(CH$_3$)$_3$), 53.0 (cys-α-*C*H), 52.6 (cys-α-*C*H), 40.9 (succinimide-*C*H), 40.0 (succinimide-*C*H), 36.2 (succinimide-*C*H$_2$), 35.9 (succinimide-*C*H$_2$), 32.7 (cys-β-*C*H$_2$), 32.3 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$), 21.7 (*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 316.13229. $C_{13}H_{22}N_3O_4S$ requires M^+ 316.13255); *m/z* (ES$^+$) 316 ([M+H]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.061 min, Purity: 96.00%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Diastereomer 1: Ret. Time = 24.185 min, 45.97%; Diastereomer 2: Ret. Time = 32.510 min, 53.84%.

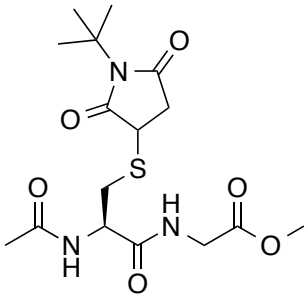
_S_-(_N-tert_-Butylsuccinimido) L-cysteinylglycine methyl ester 2,2,2-trifluoroacetate (140)



To a stirred solution of L-cys-gly, **103**, (24.6 mg, 0.14 mmol, 1.0 eq.) in MeOH (1.1 mL) was added thionyl chloride (19.1 mg, 11.6 µL, 0.16 mmol, 1.16 eq.) and the reaction mixture was stirred at RT for 2.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was lyophilized to afford L-cys-gly methyl ester, **139**, as a colourless solid (26.5 mg), which was used without further purification.

To a stirred solution of L-cys-gly methyl ester, **139**, (26.5 mg) in H_2O (1.1 mL) was added *N-tert*-butyl-maleimide (21.1 mg, 20.0 µL, 0.14 mmol, 1.0 eq. with respect to the previous step). The resulting solution was stirred at RT for 8.5 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 43:57:0.1; H_2O : MeCN : TFA) to afford *S*-(*N-tert*-butylsuccinimido) L-cysteinylglycine methyl ester 2,2,2-trifluoroacetate, **140**, (mixture of diastereoisomers) as a colourless solid (35.9 mg, 57% over two steps): $[\alpha]_D^{20}$ = +20.4 (*c* 1.0, $CHCl_3$); ν_{max}/cm^{-1} (neat) 2980 (w), 1748 (w), 1665 (s), 1460 (w), 1418 (w), 1368 (w), 1344 (m), 1265 (m), 1198 (s), 1165 (s), 1128 (s), 1028 (w); ^{1}H NMR (500 MHz; D_2O): δ 4.43-4.34 (1H, m, cys-α-C*H*), 4.16-4.05 (2H, m, gly-CH_2), 3.96-3.90 (1H, m, succinimide-C*H*), 3.77 (1.5H, s, COOCH_3), 3.76 (1.5H, s, COOCH_3), 3.48 (0.5H, dd, *J* 14.7, 5.9, cys-β-CH_2), 3.41-3.31 (1H, m, cys-β-CH_2), 3.24-3.12 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 2.58 (0.5H, dd, *J* 5.2, 4.6, succinimide-CH_2), 2.55 (0.5H, dd, *J* 5.2, 4.6, succinimide-CH_2), 1.54 (9H, s, NC(CH_3)$_3$); ^{19}F NMR (377 MHz; D_2O): δ −75.6 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.5 (succinimide-*C*O), 180.2 (succinimide-*C*O), 178.8 (succinimide-*C*O), 178.8 (succinimide-*C*O), 171.5 (*C*OOCH$_3$), 171.5 (*C*OOCH$_3$), 168.6 (cys-*C*O), 168.6 (cys-*C*O), 163.0 (q, *J* 35.1, O*C*OCF$_3$), 116.4 (q, *J* 291.5, OCO*C*F$_3$), 59.2 (N*C*(CH$_3$)$_3$), 52.9 (COO*C*H$_3$), 52.5 (cys-α-*C*H), 52.2 (cys-α-*C*H), 41.3 (gly-*C*H$_2$), 41.2 (gly-*C*H$_2$), 41.0 (succinimide-*C*H), 40.1 (succinimide-*C*H$_2$), 35.9 (succinimide-*C*H$_2$), 35.6 (succinimide-*C*H$_2$), 32.8 (cys-β-*C*H$_2$), 32.2 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 346.14322. $C_{14}H_{24}N_3O_5S$ requires M$^+$ 346.14312); *m/z* (ES$^+$) 346 ([M+H]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Diastereomer 1: Ret. Time = 8.812 min, Purity: 37.16%; Diastereomer 2: Ret. Time = 8.940 min, Purity: 59.20%;

_N_-Acetyl-_S_-(_N-tert_-butylsuccinimido) L-cysteinylglycine methyl ester (141)



Method A

To a stirred solution of _S_-(_N-tert_-butylsuccinimido) L-cysteinylglycine methyl ester 2,2,2-trifluoroacetate, **140**, (38.6 mg, 0.08 mmol, 1.0 eq.) and anhydrous NEt_3 (17.0 mg, 23.4 µL, 0.17 mmol 2.0 eq.) in anhydrous THF (0.9 mL) was added acetyl chloride (7.3 mg, 6.6 µL, 0.09 mmol, 1.1 eq.). The reaction mixture was stirred under nitrogen at RT for 4 h. After this time the reaction mixture was diluted with H_2O, lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5 → 5:95; H_2O : MeCN) to afford _N_-acetyl-_S_-(_N-tert_-butylsuccinimido) L-cysteinylglycine methyl ester, **141**, (mixture of diastereoisomers) as a colourless hygroscopic solid (20.8 mg, 64%). Data are shown below.

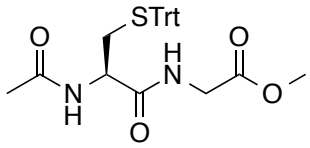
Method B

_N_-Acetyl-_S_-trityl-L-cysteinylglycine methyl ester, **142**, (0.10 g, 0.21 mmol, 1.0 eq.) and triethylsilane (0.07 g, 0.10 mL, 0.63 mmol, 3.0 eq.) were stirred under argon at RT in 20% TFA in CH_2Cl_2 (7.3 mL) for 3 h. After this time the reaction solution was concentrated _in vacuo_ and resulting residue was purified by RP C-18 silica gel chromatography, eluting with H_2O to afford _N_-acetyl-L-cysteinylglycine methyl ester trifluoroacetic acid salt, **143**, as an off-white sticky foam (25.5 mg), which was used without further purification.

To a stirred solution of _N_-acetyl-L-cysteinylglycine methyl ester trifluoroacetic acid salt, **143**, (25.5 mg) in H_2O (0.9 mL) was added _N-tert_-butylmaleimide (16.6 mg, 15.7 µL, 0.11 mmol). The

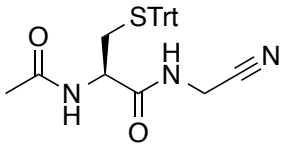
resulting solution was stirred at RT for 9 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5 → 5:95; H_2O : MeCN) to afford *N*-acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine methyl ester, **141**, (mixture of diastereoisomers) as a colourless solid (28.4 mg, 35% over two steps): $[\alpha]_D^{20}$ = −14.5 (*c* 0.55, H_2O); ν_{max}/cm^{-1} (neat) 3298 (w), 2977 (w), 1749 (m), 1696 (s), 1650 (m), 1529 (m), 1458 (m), 1437 (m), 1411 (m), 1369 (m), 1342 (m), 1263 (m), 1204 (m), 1163 (m), 1040 (m); ^{1}H NMR (500 MHz; D_2O): δ 4.68-4.61 (1H, m, cys-α-C*H*), 4.04 (2H, s, gly-CH_2), 3.93-3.85 (1H, m, succinimide-C*H*), 3.75 (3H, s, COOCH_3), 3.30 (0.5H, dd, *J* 14.1, 5.2, cys-β-CH_2), 3.24-3.10 (2H, m, cys-β-CH_2, succinimide-CH_2), 3.01 (0.5H, dd, *J* 14.1, 8.7, cys-β-CH_2), 2.58 (0.5H, dd, *J* 4.8, 4.1, succinimide-CH_2), 2.54 (0.5H, dd, *J* 4.8, 4.1, succinimide-CH_2), 2.08 (1.5H, s, COCH_3), 2.07 (1.5H, s, COCH_3), 1.54 (s, 9H, NC(CH_3)$_3$); ^{13}C NMR (126 MHz; D_2O): δ 180.2 (succinimide-*C*O), 180.1 (succinimide-*C*O), 179.1 (succinimide-*C*O), 179.0 (succinimide-*C*O), 174.3 (*C*OCH$_3$), 174.3 (*C*OCH$_3$), 172.6 (cys-*C*O), 172.5 (cys-*C*O), 171.7 (*C*OOCH$_3$), 171.7 (*C*OOCH$_3$), 59.2 (N*C*(CH$_3$)$_3$), 59.1 (N*C*(CH$_3$)$_3$), 53.2 (cys-α-*C*H), 52.8 (COO*C*H$_3$), 52.7 (cys-α-*C*H), 41.2 (gly-*C*H$_2$), 40.8 (succinimide-*C*H), 40.0 (succinimide-*C*H), 36.2 (succinimide-*C*H$_2$), 35.9 (succinimide-*C*H$_2$), 32.6 (cys-β-*C*H$_2$), 32.3 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$), 21.7 (CO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 388.15286. $C_{16}H_{26}N_3O_6S$ requires M$^+$ 388.15368); *m/z* (ES$^+$) 388 ([M+H]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.677 min, Purity: 97.16%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Diastereomer 1: Ret. Time = 37.366 min, 38.54%; Diastereomer 2: Ret. Time = 40.515 min, 61.46%.

_N_-Acetyl-_S_-trityl-L-cysteinylglycine methyl ester (142)



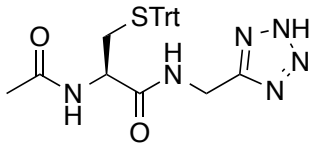
_N_-Acetyl-L-Cys(Trt)-OH, **127**, (0.30 g, 0.74 mmol, 1.05 eq.), glycine methyl ester hydrochloride (0.09 g, 0.70 mmol, 1.00 eq.) and HBTU (0.28 g, 0.74 mmol, 1.05 eq.) were stirred in anhydrous DMF (1.7 mL) for 5 min. After this time, DIPEA (0.18 g, 0.25 mL, 1.41 mmol, 2.0 eq.) was added and the reaction mixture was stirred under argon at RT for 24 h. The reaction mixture was partitioned between EtOAc (50 mL) and aqueous 0.5 M LiCl (50 mL) and the organic layer was collected, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting crude material was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether (60:40 → 100:0; EtOAc : petroleum ether) to yield _N_-acetyl-_S_-trityl-L-cysteinylglycine methyl ester, **142**, (0.28 g, 79%) as a colourless solid: R_f 0.4 (EtOAc); mp 59-63 °C (from $CHCl_3$); $[\alpha]_D^{20}$ = +1.93 (*c* 1.0, $CHCl_3$); ν_{max}/cm^{-1} (neat) 3289 (w), 3055 (w), 2970 (w), 2949 (w), 1742 (m), 1651 (m), 1593 (m), 1526 (m), 1489 (m), 1441 (m), 1371 (m), 1206 (m), 1180 (m), 1084 (m), 1034 (m), 1001 (m); ^{1}H NMR (500 MHz; $CDCl_3$): δ 7.49-7.18 (15H, m, Ar C*H*), 6.61 (1H, t, *J* 5.1, gly-N*H*), 5.80 (1H, d, *J* 7.6, cys-N*H*), 4.18-4.09 (1H, m, cys-α-C*H*), 4.02-3.86 (2H, m, gly-CH_2), 3.72 (3H, s, COOCH_3), 2.83-2.72 (1H, m, cys-β-CH_2), 2.61-2.54 (1H, m, cys-β-CH_2), 1.90 (3H, s, COCH_3), ^{13}C NMR (126 MHz; $CDCl_3$): δ 170.4 (cys-*C*O), 170.4 (NH*C*OCH_3), 169.8 (*C*OOCH_3), 144.5 (Ar *C*), 129.7 (Ar *C*H), 128.2 (Ar *C*H), 127.1 (Ar *C*H), 67.4 (S*C*Ph_3), 52.5 (COO*C*H_3), 52.1 (cys-α-*C*H), 41.3 (gly-*C*H_2), 33.2 (cys-β-*C*H_2), 23.2 (CO*C*H_3); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 499.16623 $C_{27}H_{28}N_2NaO_4S$ requires M^+ 499.16620); *m/z* (ES^+) 499 ($[M+Na]^+$, 100%); Analytical HPLC @ 220 nm Method 2 (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 13.032 min, Purity: 99.35%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Ret. Time = 26.638 min, 92.97%.

_N_-Acetyl-_S_-trityl-(_R_)-2-amino-_N_-(cyanomethyl)-3-mercaptopropanamide (144)



_N_-Acetyl-L-Cys(Trt)-OH, **127**, (0.40 g, 0.99 mmol, 1.05 eq.), aminoacetonitrile hydrochloride (86.9 mg, 0.94 mmol, 1.0 eq.) and HBTU (0.37 g, 0.99 mmol, 1.05 eq.) were stirred in anhydrous DMF (2.3 mL) for 5 min. After this time, DIPEA (0.24 g, 0.33 mL, 1.88 mmol, 2.0 eq.) was added and the reaction mixture was stirred under argon at RT for 23 h. The reaction mixture was partitioned between EtOAc (50 mL) and aqueous 0.5 M LiCl (50 mL) and the organic layer was collected, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting crude material was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether (60:40 → 100:0; EtOAc : petroleum ether) to yield _N_-acetyl-_S_-trityl-(_R_)-2-amino-_N_-(cyanomethyl)-3-mercaptopropanamide, **144**, (0.31 g, 72%) as a colourless solid: R_f 0.4 (100% EtOAc); mp 176-178 °C (from EtOAc/petroleum ether); $[\alpha]_D^{25}$ = −3.1 (_c_ 0.29, $CHCl_3$); ν_{max}/cm^{-1} (neat) 3315 (w), 3056 (w), 2981 (w), 2361 (w), 1637 (s), 1531 (m), 1489 (m), 1442 (m), 1416 (m), 1372 (w), 1345 (m), 1282 (w), 1241 (w), 1184 (w), 1082 (w), 1034 (w), 1001 (w); ^{1}H NMR (500 MHz; $CDCl_3$): δ 7.51-7.17 (16H, m, Ar C*H*, N*H*CH_2CN), 5.79 (1H, d, *J* 7.3, cys-N*H*), 4.21-4.12 (1H, m, cys-α-C*H*), 4.12-3.98 (2H, m, C*H*$_2$CN), 2.77 (1H, dd, *J* 13.1, 7.3, cys-β-C*H*$_2$), 2.62 (1H, dd, *J* 13.1, 5.4, cys-β-C*H*$_2$), 1.92 (3H, s, NHCOC*H*$_3$); ^{13}C NMR (126 MHz; $CDCl_3$): δ 171.1 (NH*C*OCH_3), 170.6 (cys-*C*O), 144.4 (Ar *C*), 129.6 (Ar *C*H), 128.3 (Ar *C*H), 127.2 (Ar *C*H), 115.7 (*C*N), 67.6 (S*C*Ph$_3$), 52.1 (cys-α-*C*H), 32.6 (cys-β-*C*H$_2$), 27.5 (*C*H$_2$CN), 23.2 (NHCO*C*H$_3$); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 466.15610 $C_{26}H_{25}N_3NaO_2S$ requires M^+ 466.15597); *m/z* (ES^-) 442 ($[M-H]^-$, 100%); Analytical HPLC @ 220 nm Method 2 (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 12.982 min, Purity: 97.85%.

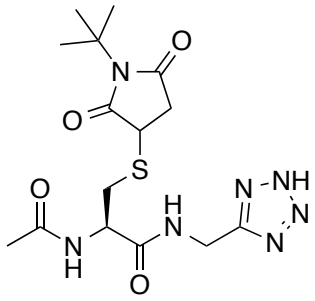
_N_-Acetyl-_S_-trityl-(_R_)-_N_-((2_H_-tetrazol-5-yl)methyl)-2-amino-3-mercaptopropanamide (145)



Sodium azide (58.6 mg, 0.90 mmol, 4.0 eq.) and zinc bromide (0.10 g, 0.45 mmol, 2.0 eq.) were stirred with *N*-acetyl-*S*-trityl-(*R*)-2-amino-*N*-(cyanomethyl)-3-mercaptopropanamide, **144**, (0.10 g, 0.23 mmol, 1.0 eq.) in THF (2.4 mL) and the reaction mixture was exposed to microwave irradiation (80 W) at 80 °C for 0.5 h. The reaction mixture was cooled and partitioned between aqueous 1M HCl (25 mL) and EtOAc (25 mL). The aqueous layer was extracted with EtOAc (2 × 25 mL), and the organic layers were combined, dried over anhydrous $MgSO_4$, filtered and concentrated *in vacuo*. The resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, MeOH and H_2O (100:0:0 → 70:15:15; EtOAc : MeOH : H_2O) to afford *N*-acetyl-*S*-trityl-(*R*)-*N*-((2*H*-tetrazol-5-yl)methyl)-2-amino-3-mercaptopropanamide, **145**, (27.1 mg, 25%) as a colourless solid: R_f 0.5 (6:1:1 EtOAc : MeOH : H_2O); mp 201-204 °C (from MeOH); $[\alpha]_D^{20}$ = +7.8 (*c* 1.0, MeOH); ν_{max}/cm^{-1} (neat) 3271 (w), 3020 (w), 2361 (m), 2342 (m), 1636 (m), 1541 (m), 1489 (m), 1443 (m), 1424 (m), 1374 (m), 1230 (m), 1184 (m), 1055 (m), 1034 (m), 1000 (m); ^{1}H NMR (500 MHz; CD_3OD): δ 7.41-7.16 (15H, m, Ar C*H*), 4.66 (1H, d, *J* 15.8, NHC$H_A$$H_B$-tetrazole), 4.55 (1H, d, *J* 15.8, NHC$H_A$$H_B$-tetrazole), 4.32-4.26 (1H, m, cys-α-C*H*), 2.60 (1H, dd, *J* 12.4, 5.8, cys-β-CH_2), 2.53 (1H, dd, *J* 12.4, 8.0, cys-β-CH_2), 1.96 (3H, s, NHCOCH_3), ^{13}C NMR (126 MHz; CD_3OD): δ 173.3 (NH*C*OCH_3), 172.6 (cys-*C*O), 157.4 (N*C*N), 145.9 (Ar *C*), 130.7 (Ar *C*H), 129.0 (Ar *C*H), 127.9 (Ar *C*H), 68.0 (S*C*Ph_3), 53.9 (cys-α-*C*H), 34.7 (cys-β-*C*H_2), 34.6 (NH*C*H_2-tetrazole), 22.5 (NHCO*C*H_3); HRMS *m/z* (ES^+) (Found: $[M+Na]^+$ 509.1731 $C_{26}H_{26}N_6NaO_2S$ requires M^+ 509.17302); *m/z* (ES^-) 485 ($[M-H]^-$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 12.195 min, Purity: 97.67%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column

(5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Ret. Time = 16.586 min, 100%.

_S_-(_N-tert_-Butylsuccinimido) (_R_)-_N_-((2_H_-tetrazol-5-yl)methyl)-2-acetamido-3-mercaptopropanamide (147)

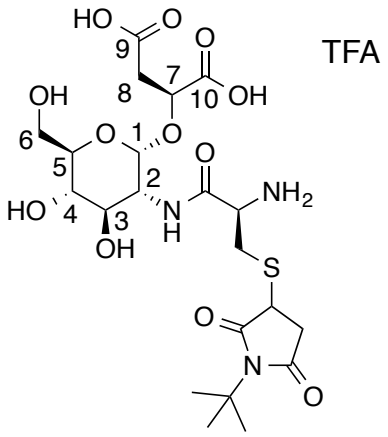


_N_-Acetyl-_S_-trityl-(_R_)-_N_-((2_H_-tetrazol-5-yl)methyl)-2-amino-3-mercaptopropanamide, **145**, (27.1 mg, 0.06 mmol, 1.0 eq.) and triethylsilane (19.4 mg, 26.7 µL, 0.17 mmol, 3.0 eq.) were stirred under argon at RT in 20% TFA in CH_2Cl_2 (2.1 mL) for 5 h. After this time the reaction solution was concentrated *in vacuo* and resulting residue was purified by RP C-18 silica gel chromatography, eluting with H_2O to afford (_R_)-_N_-((2_H_-tetrazol-5-yl)methyl)-2-acetamido-3-mercaptopropanamide, 2,2,2-trifluoroacetate salt, **146**, as a colourless solid (32.5 mg), which was used without further purification.

To a solution of (_R_)-_N_-((2_H_-tetrazol-5-yl)methyl)-2-acetamido-3-mercaptopropanamide, 2,2,2-trifluoroacetate salt, **146**, (32.5 mg) in H_2O (1.1 mL) was added _N-tert_-butylmaleimide (19.9 mg, 18.8 µL, 0.13 mmol). The resulting solution was stirred at RT for 7 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5 → 5:95; H_2O : MeCN) to afford _S_-(_N-tert_-butylsuccinimido) (_R_)-_N_-((2_H_-tetrazol-5-yl)methyl)-2-acetamido-3-mercaptopropanamide, **147**, (mixture of diastereoisomers) as a colourless solid (6.5 mg, 29% over two steps): $[\alpha]_D^{25}$ = −6.3 (_c_ 0.30, H_2O); ν_{max}/cm^{-1} (neat) 3286 (w), 3026 (w), 2937 (w),

2841 (w), 2361 (w), 2342 (w), 1771 (w), 1697 (s), 1655 (m), 1540 (m), 1457 (m), 1419 (m), 1347 (m), 1265 (m), 1232 (m), 1202 (m), 1166 (m), 1042 (m); ^{1}H NMR (500 MHz; D_2O): δ 4.76 (2H, s, NHC*H*$_2$-tetrazole), 4.68-4.60 (1H, m, cys-α-C*H*), 3.84-3.77 (1H, m, succinimide-C*H*), 3.26 (0.5H, dd, *J* 14.2, 5.3, cys-β-C*H*$_2$), 3.21-3.06 (2H, m, cys-β-C*H*$_2$, succinimide-C*H*$_2$), 3.00 (0.5H, dd, *J* 14.2, 8.4, cys-β-C*H*$_2$), 2.53 (0.5H, dd, *J* 6.0, 4.4, succinimide-C*H*$_2$), 2.49 (0.5H, dd, *J* 6.0, 4.4, succinimide-C*H*$_2$), 2.07 (1.5H, s, NHCOC*H*$_3$), 2.06 (1.5H, s, NHCOC*H*$_3$), 1.52 (4.5H, s, NC(C*H*$_3$)$_3$), 1.52 (4.5H, s, NC(C*H*$_3$)$_3$); ^{13}C NMR (126 MHz; D_2O): δ 180.1 (succinimide-*C*O), 180.1 (succinimide-*C*O), 179.0 (succinimide-*C*O), 178.9 (succinimide-*C*O), 174.4 (NH*C*OCH$_3$), 174.4 (NH*C*OCH$_3$), 172.5 (cys-*C*O), 172.5 (cys-*C*O), 154.5 (N*C*N), 59.2 (N*C*(CH$_3$)$_3$), 59.1 (N*C*(CH$_3$)$_3$), 53.2 (cys-α-*C*H), 52.8 (cys-α-*C*H), 40.6 (succinimide-*C*H), 40.0 (succinimide-*C*H), 36.1 (succinimide-*C*H$_2$), 35.9 (succinimide-*C*H$_2$), 33.0 (NH*C*H$_2$-tetrazole), 32.9 (NH*C*H$_2$-tetrazole), 32.5 (cys-β-*C*H$_2$), 32.2 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$), 21.7 (NHCO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+Na]$^+$ 420.14241 $C_{15}H_{23}N_7NaO_4S$ requires M$^+$ 420.14244); *m/z* (ES$^+$) 420 ([M+Na]$^+$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 9.183 min, Purity: 97.43%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Diastereomer 1: Ret. Time = 21.178 min, 54.09%; Diastereomer 2: Ret. Time = 30.192 min, 45.84%.

***S*-(*N*-*tert*-Butylsuccinimido) bacillithiol 2,2,2-trifluoroacetate (155)**



To a stirred solution of crude diallyl (*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-4,5-diacetoxy-6-(acetoxymethyl)-3-azidotetrahydro-2*H*-pyran-2-yl)oxy)succinate, **150**, (0.22 g), in glacial acetic acid (1.1 mL) was added zinc dust (0.27 g, 4.13 mmol). The resulting mixture was stirred at RT for 4 h. After this time, the reaction mixture was diluted with diethyl ether (25 mL), filtered through Celite® and the filtrate was washed with H_2O (10 mL) and aqueous saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with diethyl ether (2 × 50 mL), and the organic layers were combined, dried over anhydrous $MgSO_4$, filtered and concentrated *in vacuo* to afford diallyl (*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-4,5-diacetoxy-6-(acetoxymethyl)-3-aminotetrahydro-2*H*-pyran-2-yl)oxy)succinate, **151**, (0.17 g) as a pale yellow oil, which was used without further purification.

To a stirred mixture of diallyl (*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-4,5-diacetoxy-6-(acetoxymethyl)-3-aminotetrahydro-2*H*-pyran-2-yl)oxy)succinate, **151**, (0.17 g) and Boc(Trt)-L-Cys-OH (0.23 g, 0.49 mmol), in anhydrous DMF (4.9 mL) was added PyBOP (0.18 g, 0.35 mmol) and anhydrous DIPEA (0.13 g, 0.17 mL, 0.98 mmol). The resulting mixture was stirred under argon at RT for 15.5 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting oil was partitioned over EtOAc (50 mL) and H_2O (25 mL) and the organic layer was washed with H_2O (25 mL), aqueous dilute $NaHCO_3$ (10% *w/v*, 25 mL) and brine (25 mL). The organic layer was dried over anhydrous $MgSO_4$, filtered and concentrated *in vacuo* and the resulting oil was adsorbed

onto Celite® and purified by silica gel chromatography, eluting with EtOAc and petroleum ether (10:90 → 80:20; EtOAc : petroleum ether) to yield diallyl (*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-4,5-diacetoxy-6-(acetoxymethyl)-3-((*R*)-2-((*tert*-butoxycarbonyl)amino)-3-(tritylthio)propanamido)tetrahydro-2*H*-pyran-2-yl)oxy)succinate, **152**, (67.6 mg), which was used without further purification.

Diallyl (*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-4,5-diacetoxy-6-(acetoxymethyl)-3-((*R*)-2-((*tert*-butoxycarbonyl)amino)-3-(tritylthio)propanamido)tetrahydro-2*H*-pyran-2-yl)oxy)succinate, **152**, (67.6 mg) and imidazole (16.9 mg, 0.25 mmol) were azeotroped with toluene and dried under reduced pressure for 1 h. After this time triphenylphosphine (0.94 mg, 3.58 µmol), and tetrakis(triphenylphosphine)palladium(0) (1.44 mg, 1.25 µmol) dissolved in anhydrous CH_2Cl_2 (0.3 mL) were added and the resulting reaction mixture was stirred under argon at RT for 6 h. After this time, the reaction mixture was diluted with MeOH and filtered through Amberlite® IR120 hydrogen form. The filtrate was concentrated *in vacuo* to afford (*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-4,5-diacetoxy-6-(acetoxymethyl)-3-((*R*)-2-((*tert*-butoxycarbonyl)amino)-3-(tritylthio)propanamido)tetrahydro-2*H*-pyran-2-yl)oxy)succinic acid, **153**, (53.3 mg), which was used without further purification.

To a stirred solution of (*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-4,5-diacetoxy-6-(acetoxymethyl)-3-((*R*)-2-((*tert*-butoxycarbonyl)amino)-3-(tritylthio)propanamido)tetrahydro-2*H*-pyran-2-yl)oxy)succinic acid, **153**, (53.3 mg) in anhydrous MeOH (0.9 mL) at 0 °C was added sodium methoxide solution (73.0 µL, 25 wt. % in MeOH) and the reaction mixture was stirred under argon at 0 °C for 3 h. After this time, the reaction mixture was filtered through Amberlite® IR120 hydrogen form. The filtrate was concentrated *in vacuo* and the resulting crude product was adsorbed onto Celite® and purified by silica gel chromatography, eluting with H_2O, EtOAc and isopropyl alcohol (2:49:49 → 6:47:47; H_2O : EtOAc : isopropyl alcohol) to yield (*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-3-((*R*)-2-((*tert*-butoxycarbonyl)amino)-3-(tritylthio)propanamido)-4,5-dihydroxy-6-

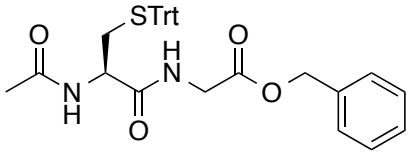
(hydroxymethyl)tetrahydro-2*H*-pyran-2-yl)oxy)succinic acid, **154**, (27.0 mg), which was used without further purification.

(*S*)-2-(((2*R*,3*R*,4*R*,5*S*,6*R*)-3-((*R*)-2-((*tert*-Butoxycarbonyl)amino)-3-(tritylthio)propanamido)-4,5-dihydroxy-6-(hydroxymethyl)tetrahydro-2*H*-pyran-2-yl)oxy)succinic acid, **154**, (27.0 mg) and triethylsilane (3.9 mg, 5.4 µL, 0.03 mmol) were stirred in TFA (0.3 mL), CH_2Cl_2 (1.0 mL) and H_2O (0.5 mL) at 0 °C for 0.75 h. After this time, the reaction mixture was concentrated *in vacuo* and the resulting residue was purified by RP C-18 silica gel chromatography, eluting with H_2O to afford slightly impure bacillithiol, **148**, (11.4 mg) as a colourless solid, which was used without further purification.

To a stirred solution of bacillithiol, **148**, (11.4 mg) in H_2O (0.5 mL) was added *N*-*tert*-butyl-maleimide (4.3 mg, 4.1 µL, 0.03 mmol) and the resulting solution was stirred at RT for 7 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by semi-preparative HPLC (Agilent ZORBAX 300SB-C18 column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) to afford *S*-(*N*-*tert*-butylsuccinimido) bacillithiol 2,2,2-trifluoroacetate, **155**, (mixture of diastereoisomers) as a colourless solid (1.5 mg): $[\alpha]_D^{20}$ = +25.1 (*c* 0.18, H_2O); ν_{max}/cm^{-1} (neat) 2980 (w), 1682 (s), 1558 (m), 1371 (m), 1348 (m), 1265 (m), 1136 (s), 1026 (s); ^{1}H NMR (500 MHz; D_2O): δ 5.11-5.07 (1H, m, $^1J_{C\text{-}H}$ = 173 Hz, C^1*H*), 4.47-4.42 (1H, m, C^7*H*), 4.31-4.25 (0.5H, m, cys-α-C*H*), 4.25-4.20 (0.5H, m, cys-α-C*H*), 3.98-3.70 (6H, m, succinimide-C*H,* C^2*H*, C^3*H*, C^4*H*, C^6*H*), 3.55-3.46 (1.5H, m, C^5*H*, cys-β-CH_2), 3.41-3.30 (1H, m, cys-β-CH_2), 3.22-3.12 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 2.94-2.86 (1H, m, C^8*H*), 2.85-2.77 (1H, m, C^8*H*), 2.59 (0.5H, dd, *J* 5.1, 4.6, succinimide-CH_2), 2.55 (0.5H, dd, 5.1, 4.6, succinimide-CH_2), 1.54 (4.5H, d, *J* 0.7, NC(CH_3)$_3$), 1.54 (4.5H, d, *J* 0.7, NC(CH_3)$_3$); ^{19}F NMR (377 MHz; D_2O): δ −75.6 (OCOCF_3); ^{13}C NMR (126 MHz; D_2O): δ 180.6 (succinimide-*C*O), 180.3 (succinimide-*C*O), 178.8 (succinimide-*C*O), 178.8 (succinimide-*C*O), 175.9 (*C*^{10}), 175.8 (*C*^{10}), 174.7 (*C*^{9}), 174.7 (*C*^{9}),

168.1 (cys-*C*O), 163.0 (q, *J* 36.8, O*C*OCF$_3$), 116.4 (q, *J* 291.1, OCO*C*F$_3$), 98.3 (*C*^{1}), 98.2 (*C*^{1}), 75.5 (*C*^{7}), 75.4 (*C*^{7}), 72.6 (*C*^{4}), 70.4 (*C*^{3}), 70.3 (*C*^{3}), 69.7 (*C*^{5}), 60.0 (*C*^{6}), 59.2 (N*C*(CH$_3$)$_3$), 53.9 (*C*^{2}), 52.8 (cys-α-*C*H), 52.5 (cys-α-*C*H), 41.2 (succinimide-*C*H), 40.0 (succinimide-*C*H), 37.7 (*C*^{8}), 36.0 (succinimide-*C*H$_2$), 35.7 (succinimide-*C*H$_2$), 33.0 (cys-β-*C*H$_2$), 32.3 (cys-β-*C*H$_2$), 27.3 (NC(*C*H$_3$)$_3$); HRMS *m/z* (ES$^+$) (Found: [M+H]$^+$ 552.18492. $C_{21}H_{34}N_3O_{12}S$ requires M$^+$ 552.18577); *m/z* (ES$^-$) 550 ([M−H]$^-$, 100%); Analytical HPLC @ 220 nm (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H$_2$O : MeCN : TFA) Ret. Time = 8.113 min, Purity: 94.96%.

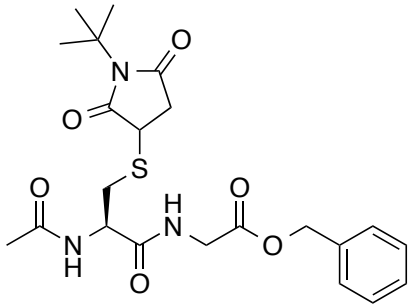
_N_-Acetyl-_S_-trityl-L-cysteinylglycine benzyl ester (156)



N-Acetyl-L-Cys(Trt)-OH, **127**, (0.30 g, 0.74 mmol, 1.05 eq.), glycine benzyl ester hydrochloride (0.14 g, 0.70 mmol, 1.00 eq.) and HBTU (0.28 g, 0.74 mmol, 1.05 eq.) were stirred in anhydrous DMF (1.7 mL) for 5 min. After this time, DIPEA (0.18 g, 0.25 mL, 1.41 mmol, 2.0 eq.) was added and the reaction mixture was stirred under argon at RT for 25.5 h. The reaction mixture was partitioned between EtOAc (50 mL) and aqueous 0.5 M LiCl (50 mL) and the organic layer was collected, dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The resulting crude material was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether (40:60 → 80:20; EtOAc : petroleum ether) to yield *N*-acetyl-*S*-trityl-L-cysteinylglycine benzyl ester, **156**, (0.32 g, 82%) as a colourless solid: R_f 0.6 (EtOAc); mp 63-66 °C (from EtOAc/petroleum ether); $[\alpha]_D^{25}$ = +0.9 (*c* 1.0, $CHCl_3$); ν_{max}/cm^{-1} (neat) 3282 (w), 3058 (w), 1747 (m), 1650 (m), 1531 (m), 1490 (m), 1444 (m), 1374 (w), 1185 (m), 1083 (w), 1033 (w), 1001 (w); ^{1}H NMR (500 MHz; $CDCl_3$): δ 7.53-7.20 (20H, m, Ar C*H*), 6.68 (1H, t, *J* 4.9, gly-N*H*), 6.03-5.90 (1H, m, N*H*COCH$_3$), 5.21-5.14 (2H, m, ArC*H*$_2$O), 4.20-4.11 (1H, m, cys-α-C*H*), 4.11-3.89 (2H, m, gly-C*H*$_2$), 2.84-2.74 (1H, m, cys-β-C*H*$_2$), 2.65-2.57 (1H, m, cys-β-C*H*$_2$),

1.90 (3H, s, NHCOC*H*$_3$); ^{13}C NMR (126 MHz; $CDCl_3$): δ 170.7 (NH*C*OCH$_3$), 170.4 (cys-*C*O), 169.2 (*C*OOBn), 144.5 (Ar *C*), 135.2 (Ar *C*), 129.7 (Ar *C*H), 128.8 (Ar *C*H), 128.7 (Ar *C*H), 128.5 (Ar *C*H), 128.2 (Ar *C*H), 127.1 (Ar *C*H), 67.4 (S*C*Ph$_3$), 67.3 (Ar*C*H$_2$O), 52.2 (cys-α-*C*H), 41.5 (gly-*C*H$_2$), 33.1 (cys-β-*C*H$_2$), 23.1 (NHCO*C*H$_3$); HRMS *m/z* (ES$^+$) (Found: [M+Na]$^+$ 575.19715 $C_{33}H_{32}N_2NaO_4S$ requires M$^+$ 575.19750); *m/z* (ES$^+$) 264 (100%), 575 ([M+Na]$^+$, 10%); Analytical HPLC @ 220 nm Method 2 (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 14.308 min, Purity: 97.88%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Ret. Time = 33.600 min, 98.66%.

_N_-Acetyl-_S_-_N_-_tert_-butylsuccinimido L-cysteinylglycine benzyl ester (158)



_N_-Acetyl-_S_-trityl-L-cysteinylglycine benzyl ester, **156**, (0.15 g, 0.28 mmol, 1.0 eq.) and triethylsilane (95.9 mg, 0.13 mL, 0.83 mmol, 3.0 eq.) were stirred under argon at RT in 20% TFA in CH_2Cl_2 (9.4 mL) for 5 h. After this time the reaction solution was concentrated *in vacuo* and resulting residue was adsorbed onto Celite® and purified by silica gel chromatography, eluting with EtOAc, petroleum ether (80:20 → 100:0; EtOAc : petroleum ether) to afford _N_-acetyl-L-cysteinylglycine benzyl ester trifluoroacetic acid salt, **157**, as a colourless solid (71.1 mg), which was used without further purification.

To a stirred solution of _N_-acetyl-L-cysteinylglycine benzyl ester trifluoroacetic acid salt, **157**, (71.1 mg) in H_2O/MeOH (1.9:1.3, 3.2 mL) was added _N-tert_-butyl-maleimide (35.1 mg, 33.1 µL,

0.23 mmol). The resulting solution was stirred at RT for 3 h. After this time, the reaction solution was lyophilized and the resulting solid was purified by preparative HPLC (SUPELCOSIL™ LC-Si column, 5 µM, 25 cm × 21.2 mm; 100:0 → 50:50; Hexane : IPA) to afford *N*-acetyl-*S*-*N*-*tert*-butylsuccinimido L-cysteinylglycine benzyl ester, **158**, (mixture of diastereoisomers) as a colourless solid (31.0 mg, 24% over two steps): $[\alpha]_D^{25}$ = −19.3 (*c* 1.0, MeCN); mp 42-46 °C (from hexane/IPA); ν_{max}/cm^{-1} (neat) 3298 (w), 2980 (w), 1747 (w), 1697 (s), 1650 (m), 1530 (m), 1456 (w), 1369 (m), 1342 (m), 1263 (m), 1232 (m), 1164 (s), 1039 (w); ^{1}H NMR (500 MHz; CD_3CN): δ 7.44-7.34 (5H, m, Ar C*H*), 7.31-7.23 (1H, m, gly-N*H*), 7.19 (0.5H, d, *J* 7.8, cys-N*H*), 6.95 (0.5H, d, *J* 7.8, cys-N*H*), 5.17 (2H, s, ArCH_2O), 4.68-4.58 (1H, m, cys-α-C*H*), 3.98 (2H, d, *J* 5.9, gly-CH_2), 3.79-3.74 (1H, m, succinimide-C*H*), 3.30-3.18 (1H, m, cys-β-CH_2), 3.09-2.96 (1.5H, m, cys-β-CH_2, succinimide-CH_2), 2.91 (0.5H, dd, *J* 14.2, 8.2, cys-β-CH_2), 2.40-2.30 (1H, m, succinimide-CH_2), 1.99-1.95 (3H, s, NHCOCH_3), 1.56 (4.5H, s, NC(CH_3)$_3$), 1.56 (4.5H, s, NC(CH_3)$_3$); ^{13}C NMR (126 MHz; CD_3CN): δ 179.6 (succinimide-*C*O), 179.2 (succinimide-*C*O), 176.8 (succinimide-*C*O), 176.7 (succinimide-*C*O), 171.7 (cys-*C*O), 171.7 (cys-*C*O), 171.2 (NH*C*OCH_3), 171.2 (NH*C*OCH_3), 170.4 (ArCH_2OO*C*), 170.4 (ArCH_2OO*C*), 137.1 (Ar*C*), 129.5 (Ar C*H*), 129.2 (Ar C*H*), 129.1 (Ar C*H*), 67.4 (Ar*C*H_2O), 59.0 (N*C*$(CH_3)_3$), 59.0 (N*C*$(CH_3)_3$), 53.8 (cys-α-*C*H), 53.2 (cys-α-*C*H), 41.9 (gly-*C*H_2), 41.7 (succinimide-*C*H), 40.7 (succinimide-*C*H), 37.2 (succinimide-*C*H_2), 36.9 (succinimide-*C*H_2), 34.5 (cys-β-*C*H_2), 34.2 (cys-β-*C*H_2), 28.4 (NC(*C*H_3)$_3$), 23.0 (NHCO*C*H_3); HRMS *m/z* (ES^+) (Found: $[M+H]^+$ 464.18498 $C_{22}H_{30}N_3O_6S$ requires M^+ 464.18498); *m/z* (ES^+) 464 ($[M+H]^+$, 100%); Analytical HPLC @ 220 nm Method 2 (Acclaim® 120 C18 RP LC Column; 95:5:0.1 → 5:95:0.1; H_2O : MeCN : TFA) Ret. Time = 11.460 min, Purity: 98.43%; Chiral HPLC @ 220 nm (ChiralPak® AD-H column (5 µm, 4.6 x 150 mm); Isocratic method: 75 : 25 Hexane : IPA, 0.3 mL / min) Ret. Time = 61.882 min, 96.60%.

7.6 Chapter 7 References:

[1] S. Miller, R. M. Douglas, P. Carter, I. R. Booth, *J. Biol. Chem.* **1997**, *272*, 24942–24947.
[2] L. S. Ness, I. R. Booth, *J. Biol. Chem.* **1999**, *274*, 9524–9530.
[3] J. Healy, S. Ekkerman, C. Pliotas, M. Richard, W. Bartlett, S. C. Grayer, G. M. Morris, S. Miller, I. R. Booth, S. J. Conway, et al., *Biochemistry* **2014**, *53*, 1982–1992.
[4] S. Miller, L. S. Ness, C. M. Wood, B. C. Fox, I. R. Booth, *J. Bacteriol.* **2000**, *182*, 6536–6540.
[5] F. H. Niesen, H. Berglund, M. Vedadi, *Nat Protoc* **2007**, DOI 10.1038/nprot.2007.321.
[6] L. Chen, J. Wang, Y.-Y. Zhang, S. F. Yan, D. Neumann, U. Schlattner, Z.-X. Wang, J.-W. Wu, *Nat. Struct. Mol. Biol.* **2012**, *19*, 716–718.
[7] F. D. L. Kondrat, W. B. Struwe, J. L. P. Benesch, *Methods Mol. Biol.* **2015**, *1261*, 349–371.
[8] J. A. Aguilar, M. Nilsson, G. Bodenhausen, G. A. Morris, *Chem. Commun.* **2012**, *48*, 811–813.
[9] P. S. C. Wu, G. Otting, *J. Magn. Reson.* **2005**, *176*, 115–119.
[10] C. Dalvit, *Concepts in Magnetic Resonance Part A* **2008**, *32A*, 341–372.
[11] *Clinical and laboratory standards institute* **2012**, *32*.
[12] A. B. Pangborn, M. A. Giardello, R. H. Grubbs, R. K. Rosen, F. J. Timmers, *Organometallics* **1996**, *15*, 1518–1520.
[13] W. L. F. Armarego, C. L. L. Chai, in *Purification of Laboratory Chemicals (Fifth Edition)* (Ed.: W.L.F.A.L.L. Chai), Butterworth-Heinemann, Burlington, **2003**, pp. xi–xii.
[14] K. C. Joshi, V. N. Pathak, S. Sharma, *Journal of Fluorine Chemistry* **1986**, *32*, 299–307.
[15] M. Hussain, V. Ahmed, B. Hill, Z. Ahmed, S. D. Taylor, *Bioorganic & Medicinal Chemistry* **2008**, *16*, 6764–6777.
[16] L. Pan, Y. Jiang, Z. Liu, X.-H. Liu, Z. Liu, G. Wang, Z.-M. Li, D. Wang, *Eur J Med Chem* **2012**, *50*, 18–26.
[17] V. Petrow, O. Stephenson, A. M. Wild, *J. Pharm. Pharmacol.* **1960**, *12*, 705–719.
[18] H. Li, S. Thota, D. Carroll, A. Argade, K. Tso, A. Sran, J. Clough, H. Keim, S. Bhamidipati, V. Taylor, et al., *Compositions and Methods for Inhibition of the Jak Pathway*, **2006**.
[19] L. Strekowski, S. E. Patterson, L. Janda, R. L. Wydra, D. B. Harden, M. Lipowska, M. T. Cegla, *J. Org. Chem.* **1992**, *57*, 196–201.
[20] L. Strekowski, A. S. Kiselyov, M. Hojjat, *J. Org. Chem.* **1994**, *59*, 5886–5890.
[21] L. S. Zhang, K. Chen, G. Chen, B. J. Li, S. Luo, Q. Y. Guo, J. B. Wei, Z. J. Shi, *Org. Lett.* **2013**, *15*, 10–13.
[22] Y. M. Lee, M. E. Moon, V. Vajpayee, V. D. Filimonov, K.-W. Chi, *Tetrahedron* **2010**, *66*, 7418–7422.
[23] X. Li, Z. Li, H. Deng, X. Zhou, *Tetrahedron Letters* **2013**, *54*, 2212–2216.
[24] D. Döpp, H. Weiler, *Chemische Berichte* **1979**, *112*, 3950–3954.
[25] K. Kondo, T. Iida, H. Fujita, T. Suzuki, K. Yamaguchi, Y. Murakami, *Tetrahedron* **2000**, *56*, 8883–8891.
[26] J. J. Neumann, S. Rakshit, T. Dröge, F. Glorius, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 6892–6895.
[27] J. H. Chan, G. A. Freeman, J. H. Tidwell, K. R. Romines, L. T. Schaller, J. R. Cowan, S. S. Gonzales, G. S. Lowell, C. W. Andrews, D. J. Reynolds, et al., *J. Med. Chem.* **2004**, *47*, 1175–1182.
[28] G. B. Jackman, V. Petrow, O. Stephenson, A. M. Wild, *Journal of Pharmacy and Pharmacology* **1963**, *15*, 202–211.
[29] P. A. Procopiou, V. J. Barrett, N. J. Bevan, K. Biggadike, P. C. Box, P. R. Butchers, D. M. Coe, R. Conroy, A. Emmons, A. J. Ford, et al., *J. Med. Chem.* **2010**, *53*, 4522–4530.
[30] H. Burton, P. F. Hu, *J. Chem. Soc.* **1948**, 601–603.

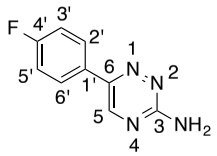
[31] S. J. Lin, W. J. Tsai, W. F. Chiou, T. H. Yang, L. M. Yang, *Bioorganic & Medicinal Chemistry* **2008**, *16*, 2697–2706.
[32] C. J. Zheng, S. M. Jiang, Z. H. Chen, B. J. Ye, H. R. Piao, *Archiv der Pharmazie* **2011**, *344*, 689–695.
[33] J. Healy, Studies on the Bacterial Potassium Efflux System KefC, and Its Ancillary Protein KefF, University of Aberdeen, **2010.**
[34] C. Loukou, P. Changenet-Barret, M.-N. Rager, P. Plaza, M. M. Martin, J.-M. Mallet, *Org. Biomol. Chem.* **2011**, *9*, 2209–2218.
[35] G. J. L. Bernardes, E. J. Grayson, S. Thompson, J. M. Chalker, J. C. Errey, F. El Oualid, T. D. W. Claridge, B. G. Davis, *Angewandte Chemie* **2008**, *120*, 2276–2279.
[36] C. E. Aroyan, A. Dermenci, S. J. Miller, *J. Org. Chem.* **2010**, *75*, 5784–5796.

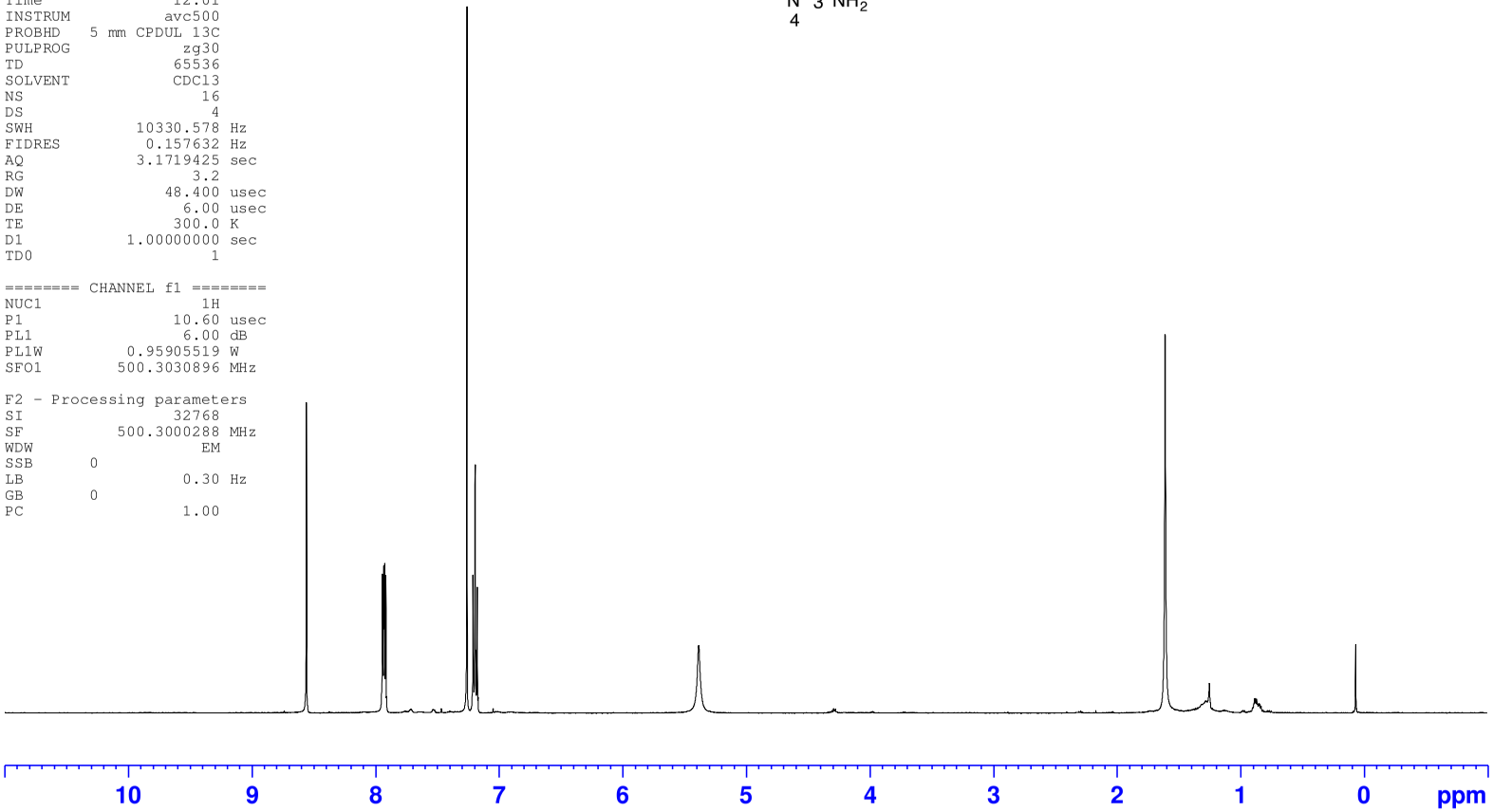
Selected NMR Spectra

^{1}H, ^{13}C and (where appropriate) ^{19}F NMR spectra are reported for all novel compounds, i.e. those for which this data has not previously been reported. Spectra are presented in the order in which the compounds appear in Chapter 7.

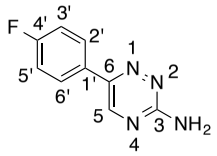
^{1}H NMR - 6-(4'-Fluorophenyl)-1,2,4-triazin-3-amine (49)

```
Current Data Parameters
NAME      sg53081309 -- SG53C
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130914
Time              12.01
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT            CDCl3
NS                    16
DS                     4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                   3.2
DW                48.400 usec
DE                 6.00 usec
TE                300.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                 10.60 usec
PL1                 6.00 dB
PL1W          0.95905519 W
SFO1         500.3030896 MHz

F2 - Processing parameters
SI                 32768
SF          500.3000288 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```

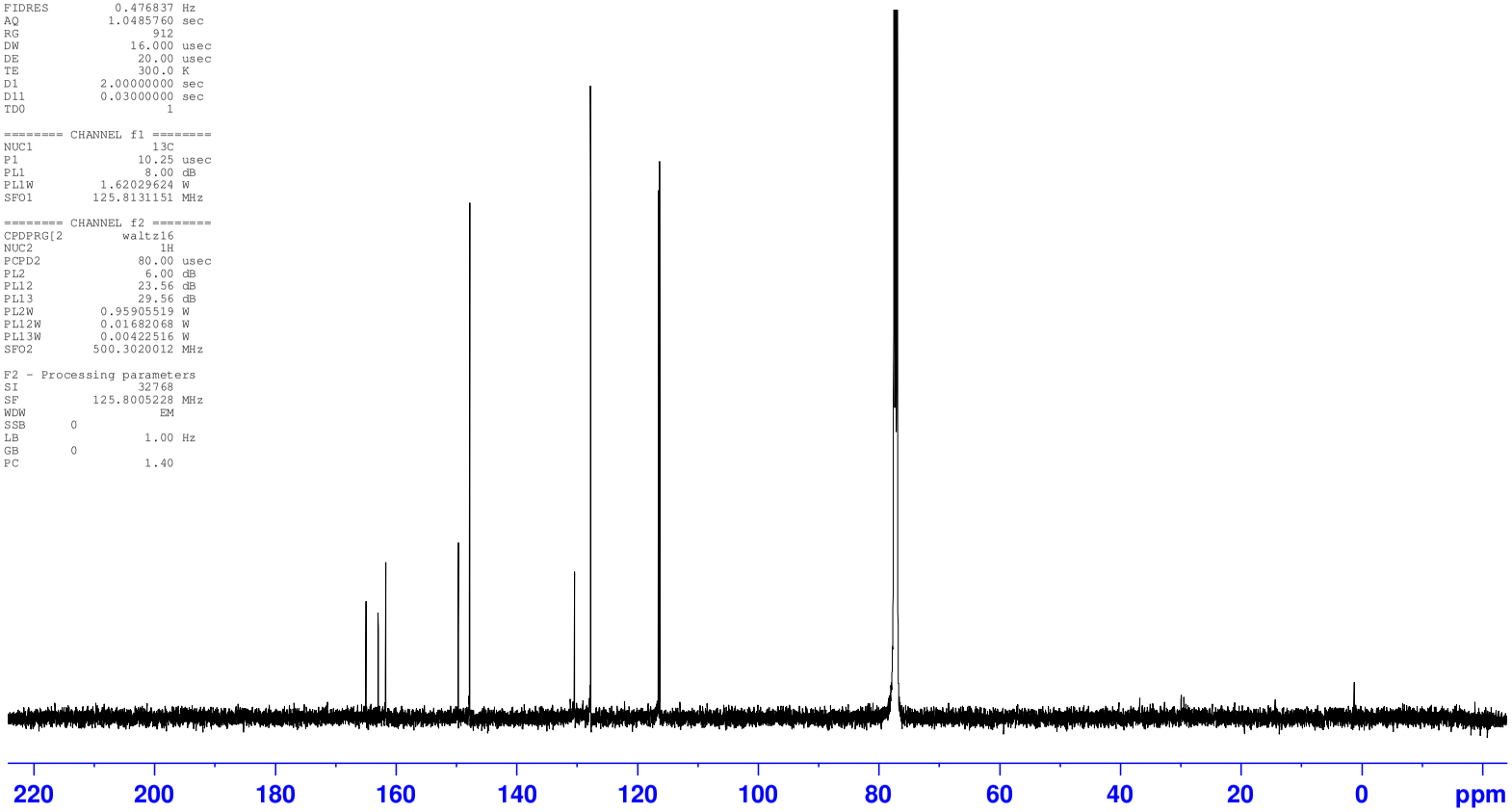




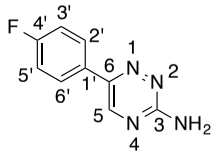
^{13}C NMR - 6-(4'-Fluorophenyl)-1,2,4-triazin-3-amine (49)



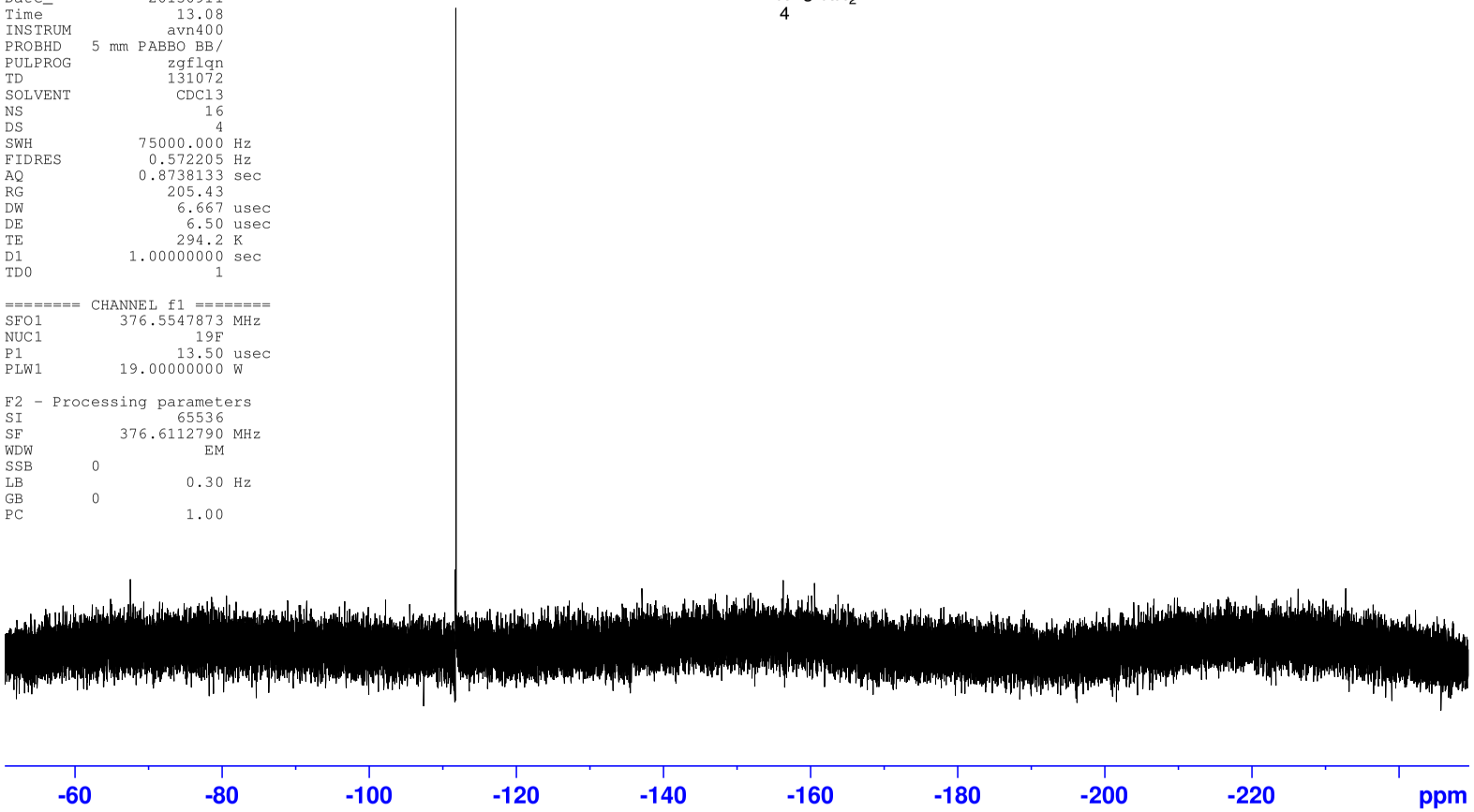

```
Current Data Parameters
NAME      sg53081309 -- SG53C
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130914
Time              12.55
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG           zgpg30
TD                65536
SOLVENT           CDCl3
NS                 1024
DS                    2
SWH           31250.000 Hz
FIDRES         0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               20.00 usec
TE                300.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                 13C
P1                10.25 usec
PL1                8.00 dB
PL1W         1.62029624 W
SFO1        125.8131151 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                 1H
PCPD2             80.00 usec
PL2                6.00 dB
PL12               23.56 dB
PL13               29.56 dB
PL2W          0.95905519 W
PL12W         0.01682068 W
PL13W         0.00422516 W
SFO2        500.3020012 MHz

F2 - Processing parameters
SI                32768
SP           125.8005228 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```



^{19}F NMR - 6-(4'-Fluorophenyl)-1,2,4-triazin-3-amine (49)

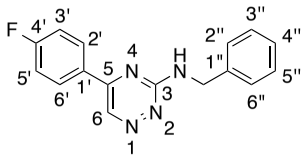



```
Current Data Parameters
NAME      Sep11-2013-24 -- SG53C
EXPNO                  5
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130911
Time              13.08
INSTRUM          avn400
PROBHD   5 mm PABBO BB/
PULPROG          zgflqn
TD               131072
SOLVENT            CDCl3
NS                    16
DS                     4
SWH            75000.000 Hz
FIDRES          0.572205 Hz
AQ             0.8738133 sec
RG                205.43
DW                 6.667 usec
DE                 6.50 usec
TE                294.2 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         376.5547873 MHz
NUC1                  19F
P1                  13.50 usec
PLW1        19.00000000 W

F2 - Processing parameters
SI                 65536
SF           376.6112790 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```

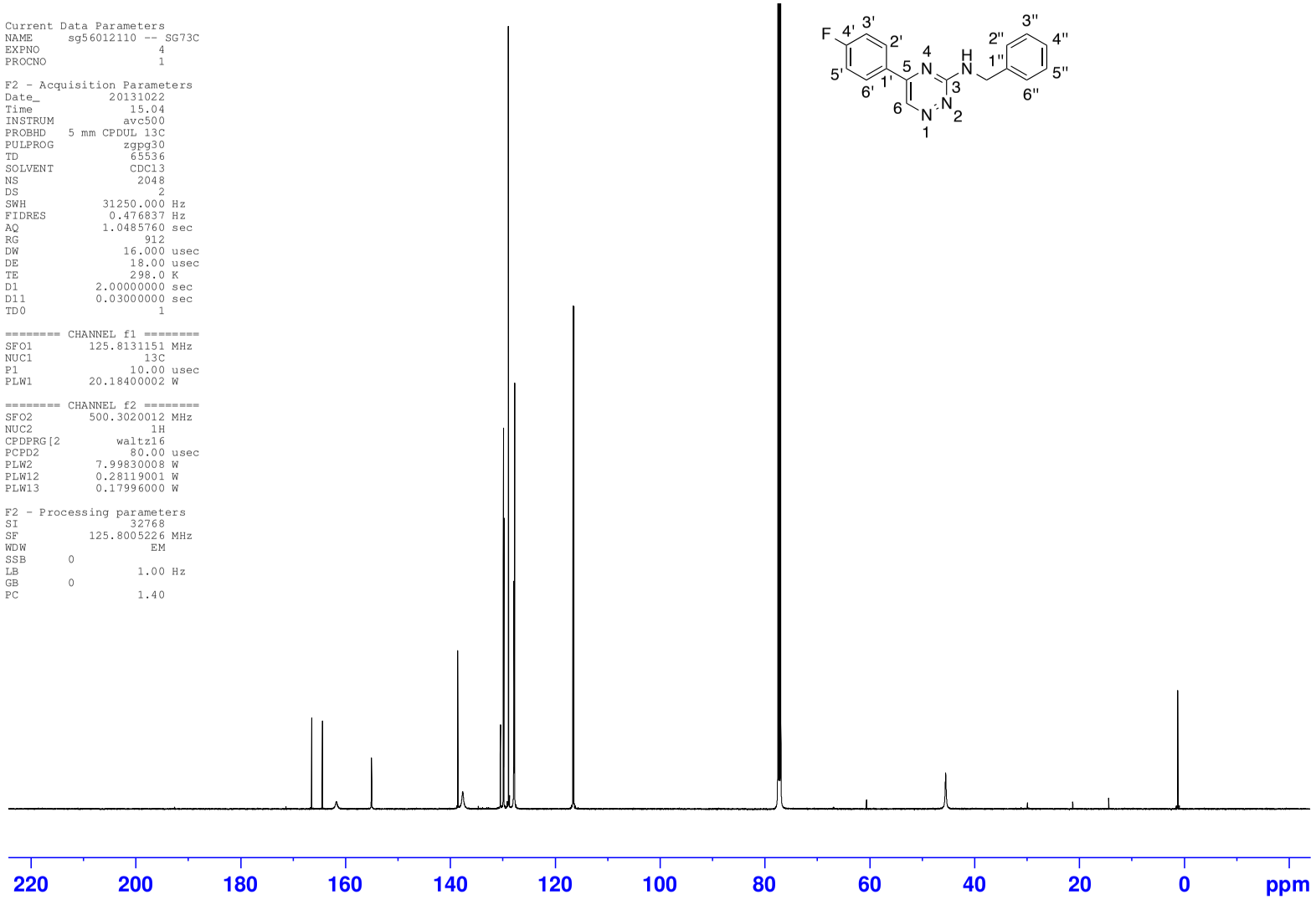


^{1}H NMR - *N*-Benzyl-5-(4'-fluorophenyl)-1,2,4-triazin-3-amine (53)

```
Current Data Parameters
NAME      sg56012110 -- SG73C
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20131022
Time              12.56
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT            CDCl3
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                 15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.3000132 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```

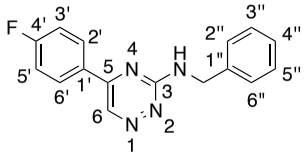


^{13}C NMR - *N*-Benzyl-5-(4'-fluorophenyl)-1,2,4-triazin-3-amine (53)



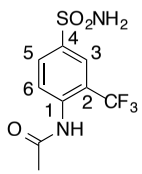
^{19}F NMR - *N*-Benzyl-5-(4'-fluorophenyl)-1,2,4-triazin-3-amine (53)

```
Current Data Parameters
NAME      Oct16-2013-46 -- SG73C 12-13
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20131017
Time               9.57
INSTRUM          avf400
PROBHD   5 mm PABBO BB/
PULPROG          zgflqn
TD               131072
SOLVENT            CDCl3
NS                    16
DS                     4
SWH            75000.000 Hz
FIDRES         0.572205 Hz
AQ             0.8738133 sec
RG                205.43
DW                6.667 usec
DE                6.50 usec
TE                294.5 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         376.5547873 MHz
NUC1                 19F
P1                 13.50 usec
PLW1         19.00000000 W

F2 - Processing parameters
SI                 65536
SF           376.6112790 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```



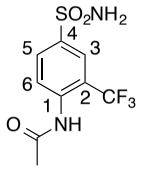
^{1}H NMR - *N*-[4-Sulfamoyl-2-(trifluoromethyl)phenyl]acetamide (67a)

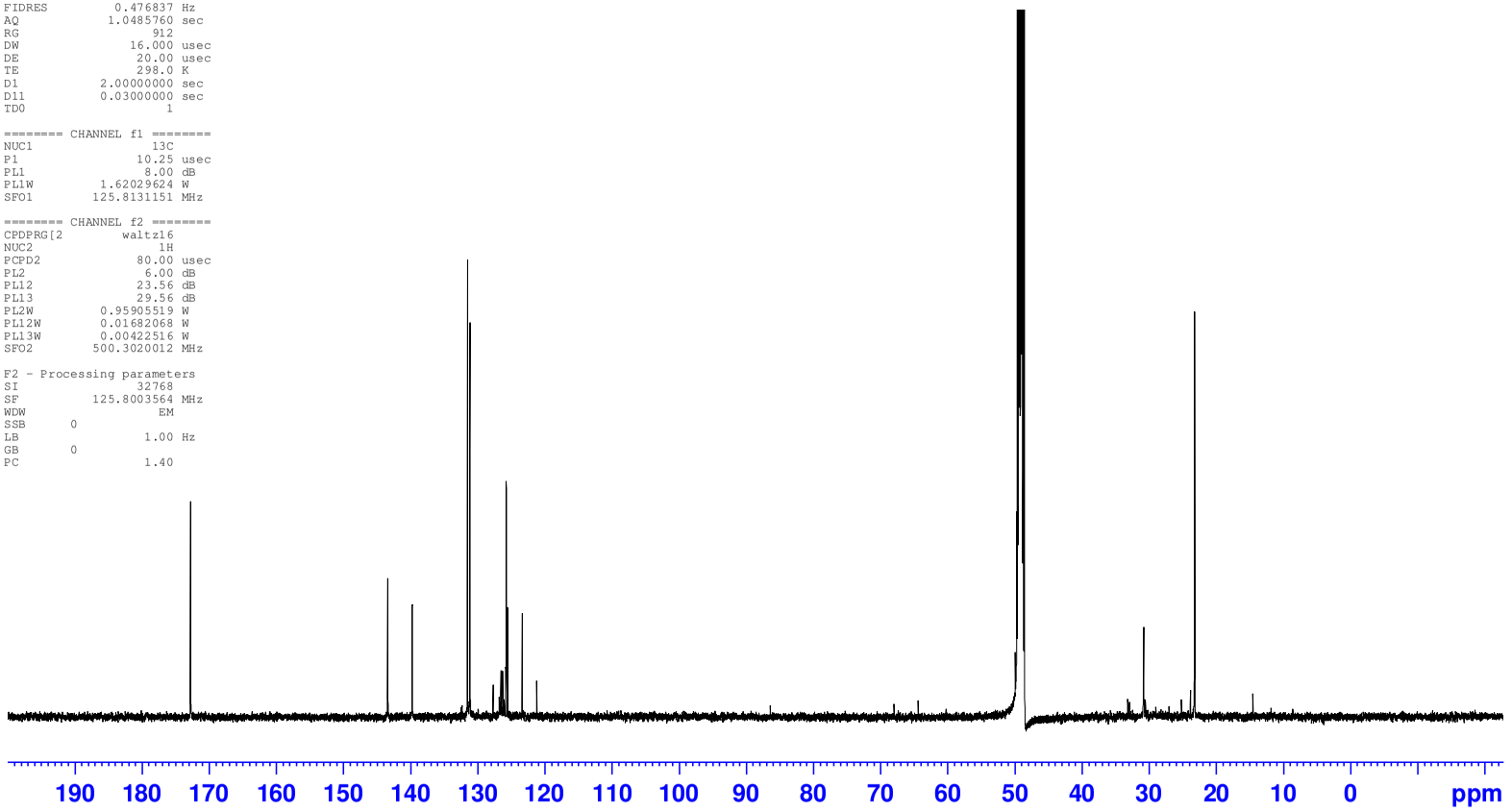
```
Current Data Parameters
NAME      Feb13-2013-4 -- SG56A
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130213
Time              15.55
INSTRUM           av400
PROBHD   5 mm QNP 1H/13
PULPROG            zg60
TD                65536
SOLVENT            MeOD
NS                   16
DS                    2
SWH            8278.146 Hz
FIDRES         0.126314 Hz
AQ             3.9583745 sec
RG                 71.8
DW                60.400 usec
DE                 7.50 usec
TE                300.0 K
D1           1.00000000 sec

======== CHANNEL f1 ========
NUC1                1H
P1                10.10 usec
PL1       0 dB
SFO1       400.2024714 MHz

F2 - Processing parameters
SI                32768
SF          400.2000044 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```



^{13}C NMR - *N*-[4-Sulfamoyl-2-(trifluoromethyl)phenyl]acetamide (67a)

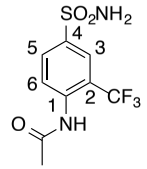
```
Current Data Parameters
NAME      sg27810402 -- SG52A
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_          20130205
Time              23.10
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT            MeOD
NS                 2048
DS                    2
SWH           31250.000 Hz
FIDRES            0.476837 Hz
AQ             1.0485760 sec
RG                  912
DW               16.000 usec
DE               20.00 usec
TE                298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                 13C
P1                10.25 usec
PL1                8.00 dB
PL1W         1.62029624 W
SFO1        125.8131151 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                 1H
PCPD2             80.00 usec
PL2                6.00 dB
PL12              23.56 dB
PL13              29.56 dB
PL2W         0.95905519 W
PL12W        0.01682068 W
PL13W        0.00422516 W
SFO2        500.3020012 MHz

F2 - Processing parameters
SI                32768
SP           125.8003564 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

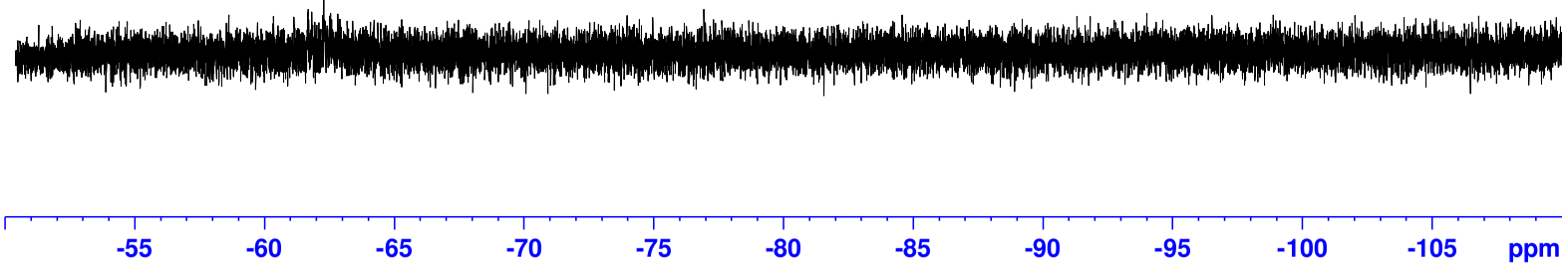




^{19}F NMR - *N*-[4-Sulfamoyl-2-(trifluoromethyl)phenyl]acetamide (67a)

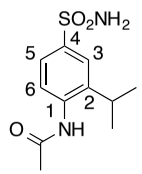
```
Current Data Parameters
NAME      Mar14-2013-28 -- SG52A
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130315
Time              19.24
INSTRUM          avn400
PROBHD   5 mm PABBO BB/
PULPROG         zgfhigqn
TD                131072
SOLVENT             MeOD
NS                    16
DS                     4
SWH            75000.000 Hz
FIDRES         0.572205 Hz
AQ             0.8738133 sec
RG                205.43
DW                6.667 usec
DE                 6.50 usec
TE                294.7 K
D1           1.00000000 sec
D11          0.03000000 sec
D12          0.00002000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         376.5547873 MHz
NUC1                 19F
P1                 13.50 usec
PLW1        19.00000000 W

======== CHANNEL f2 ========
SFO2         400.2516010 MHz
NUC2                 1H
CPDPRG[2          waltz16
PCPD2              90.00 usec
PLW2        16.70000076 W
PLW12        0.32991999 W

F2 - Processing parameters
SI                 65536
SF           376.6112790 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```





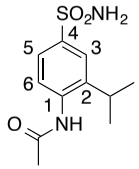
^{1}H NMR - *N*-[(2-Propan-2-yl)-4-sulfamoylphenyl)acetamide (67c)

```
Current Data Parameters
NAME      sg57950611 -- SG89C R1R2R3
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_           20131108
Time                0.30
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT             MeOD
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ              3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                  1H
P1                 15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.3000107 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```



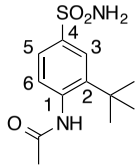
^{13}C NMR - *N*-[(2-Propan-2-yl)-4-sulfamoylphenyl)acetamide (67c)

```
Current Data Parameters
NAME      sg57950611 -- SG89C R1R2R3
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_           20131108
Time               1.44
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             MeOD
NS                 1024
DS                    2
SWH           31250.000 Hz
FIDRES         0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               18.00 usec
TE               298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        125.8131151 MHz
NUC1               13C
P1               10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2            80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8003589 MHz
WDW                EM
SSB      0
LB                1.00 Hz
GB       0
PC                1.40
```

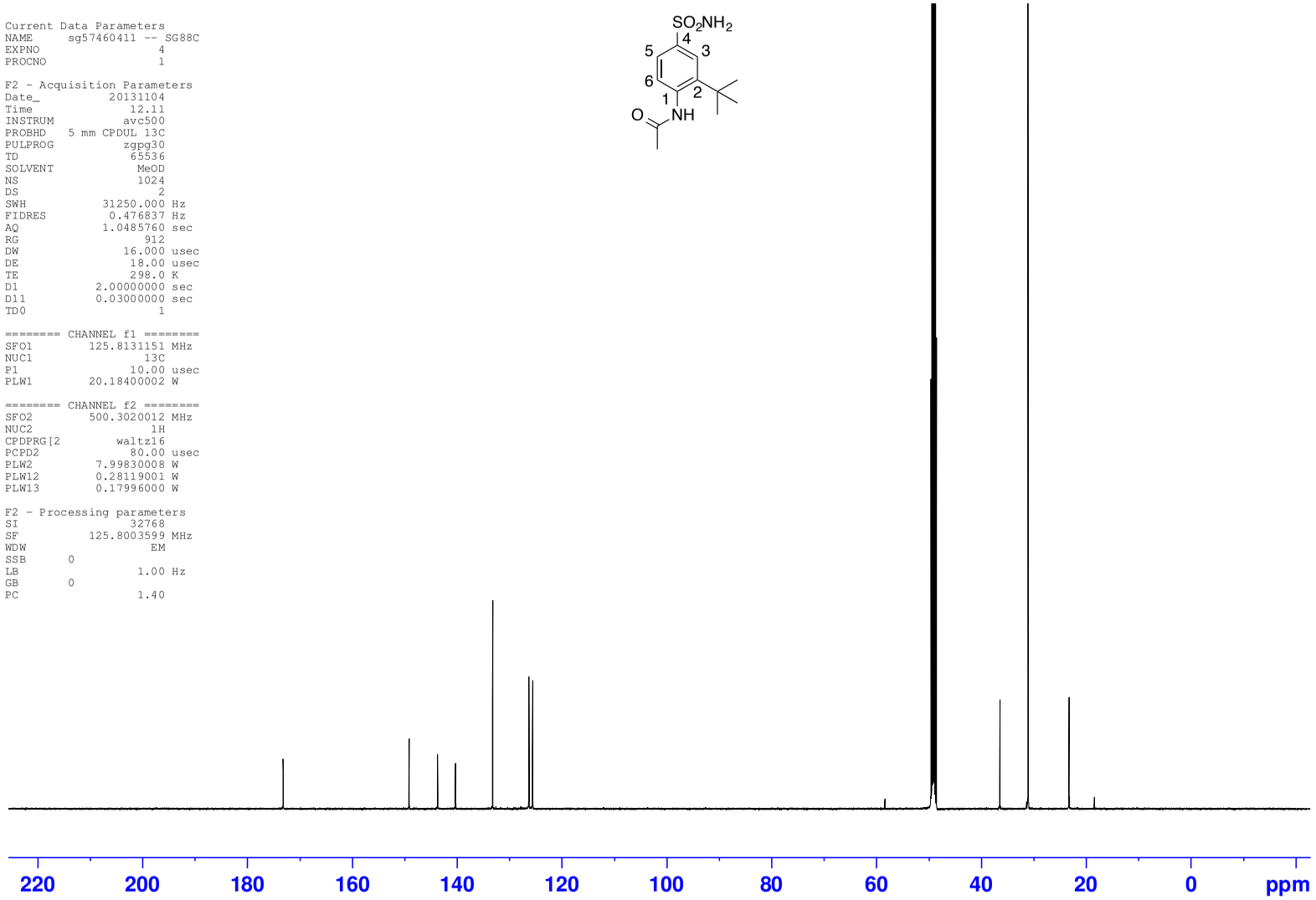


^{1}H NMR - *N*-(2-*tert*-Butyl-4-sulfamoylphenyl)acetamide (67d)

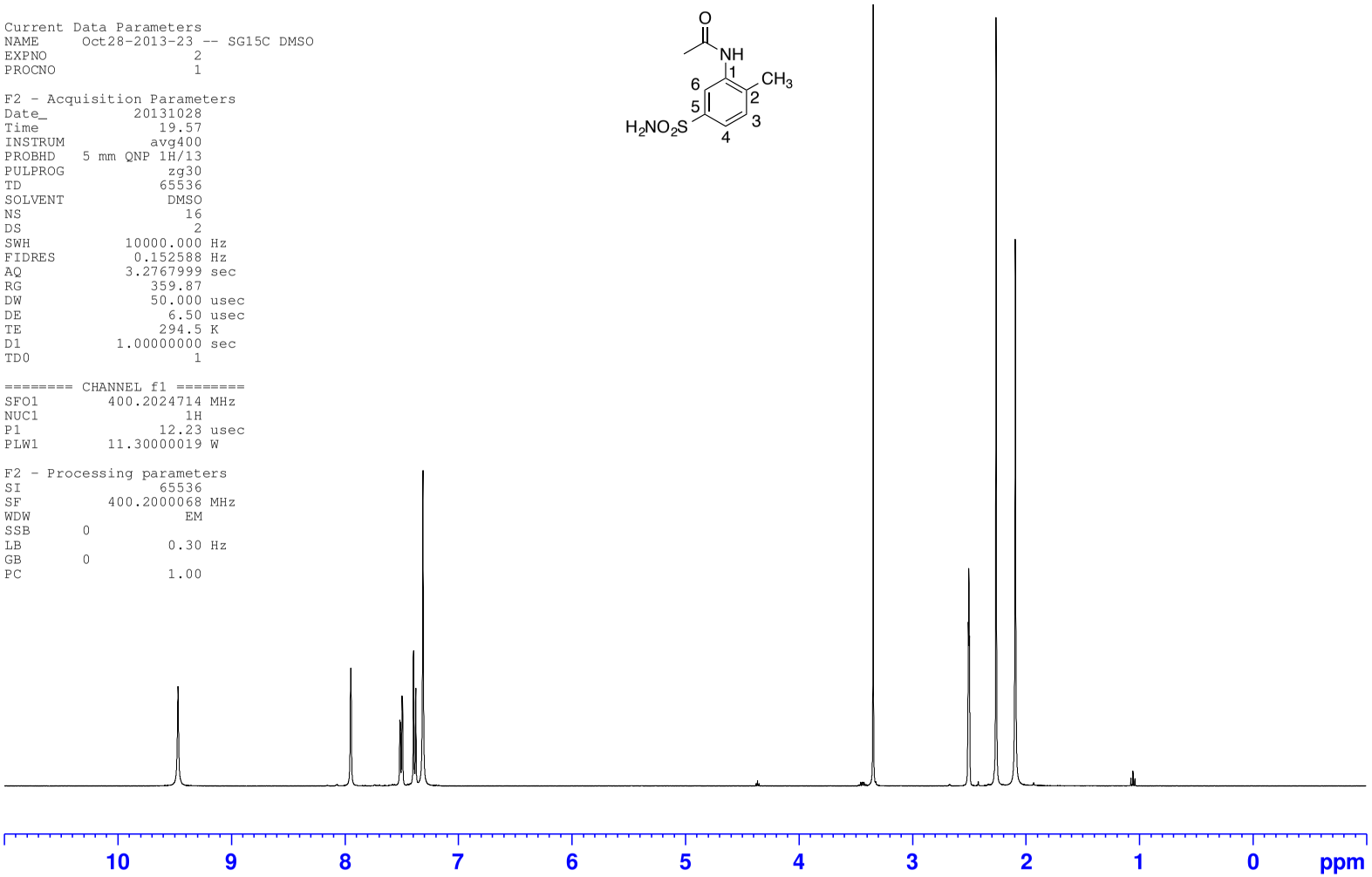
```
Current Data Parameters
NAME      sg57460411 -- SG88C
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20131104
Time              10.03
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG            zg30
TD                65536
SOLVENT            MeOD
NS                   16
DS                    4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                 2.56
DW               48.400 usec
DE               10.00 usec
TE               298.0 K
D1           1.00000000 sec
TD0                  1

======== CHANNEL f1 ========
SFO1        500.3030896 MHz
NUC1                1H
P1                15.00 usec
PLW1         7.99830008 W

F2 - Processing parameters
SI                65536
SF          500.3000108 MHz
WDW                 EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```



^{13}C NMR - *N*-(2-*tert*-Butyl-4-sulfamoylphenyl)acetamide (67d)

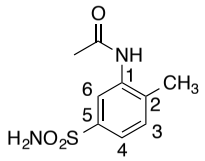


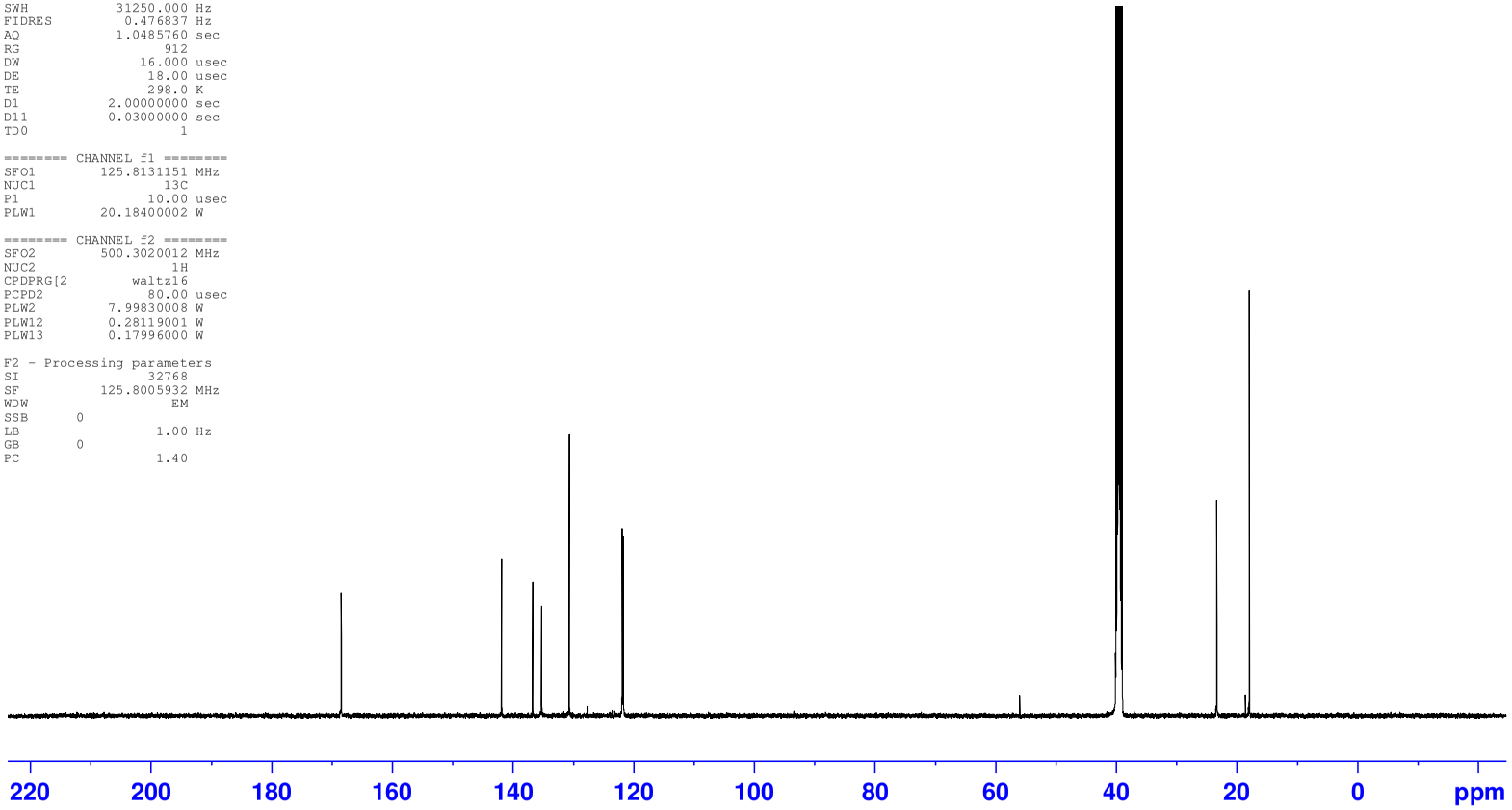
^{1}H NMR - *N*-(2-Methyl-5-sulfamoylphenyl)acetamide (72b)



^{13}C NMR - *N*-(2-Methyl-5-sulfamoylphenyl)acetamide (72b)

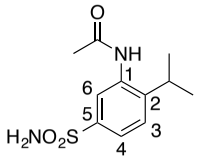
```
Current Data Parameters
NAME      sg54232609 -- S15C DMSO
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_           20130927
Time              15.02
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             DMSO
NS                  512
DS                    2
SWH           31250.000 Hz
FIDRES         0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               18.00 usec
TE               298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        125.8131151 MHz
NUC1               13C
P1               10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2            80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8005932 MHz
WDW                EM
SSB       0
LB               1.00 Hz
GB        0
PC               1.40
```





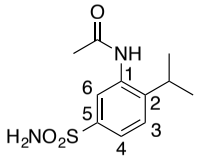
^{1}H NMR - *N*-[2-(Propan-2-yl)-5-sulfamoylphenyl]acetamide (72c)

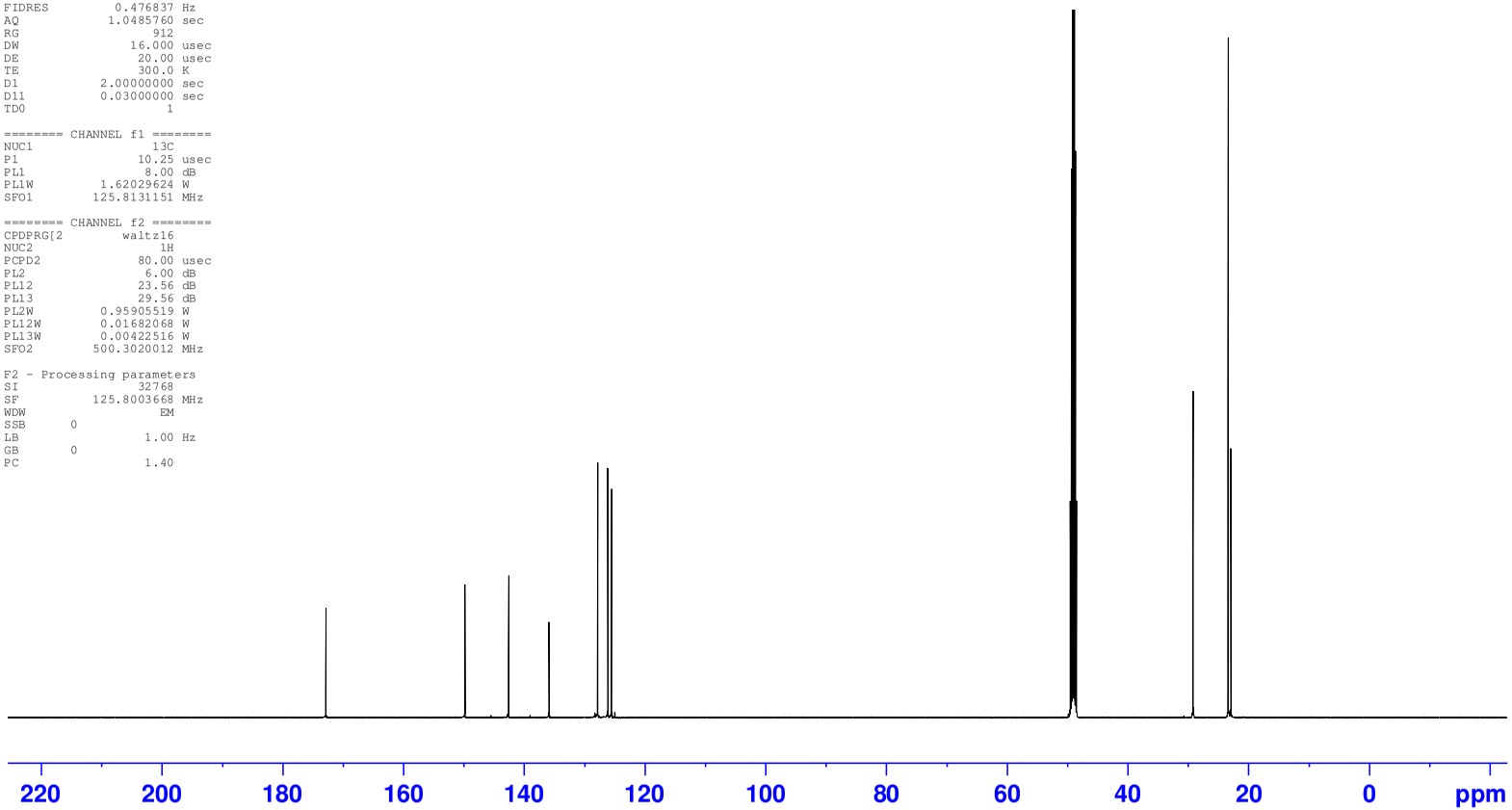
```
Current Data Parameters
NAME      sg26302301 -- SG41A
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130124
Time               7.14
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT             MeOD
NS                    16
DS                     2
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                 2.56
DW                48.400 usec
DE                 6.00 usec
TE                300.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                10.60 usec
PL1                6.00 dB
PL1W         0.95905519 W
SFO1        500.3030896 MHz

F2 - Processing parameters
SI                 32768
SF          500.3000106 MHz
WDW                   EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```



^{13}C NMR - *N*-[2-(Propan-2-yl)-5-sulfamoylphenyl]acetamide (72c)

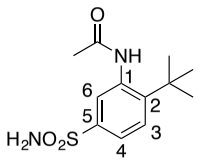
```
Current Data Parameters
NAME      sg26302301 -- SG41A
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_          20130124
Time               7.50
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             MeOD
NS                   256
DS                     2
SWH            31250.000 Hz
FIDRES         0.476837 Hz
AQ             1.0485760 sec
RG                   912
DW                16.000 usec
DE                20.00 usec
TE                300.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                  1

======== CHANNEL f1 ========
NUC1                13C
P1                10.25 usec
PL1                8.00 dB
PL1W         1.62029624 W
SFO1        125.8131151 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                1H
PCPD2              80.00 usec
PL2                6.00 dB
PL12               23.56 dB
PL13               29.56 dB
PL2W          0.95905519 W
PL12W         0.01682068 W
PL13W         0.00422516 W
SFO2         500.3020012 MHz

F2 - Processing parameters
SI                 32768
SF          125.8003668 MHz
WDW                EM
SSB      0
LB                 1.00 Hz
GB       0
PC                 1.40
```





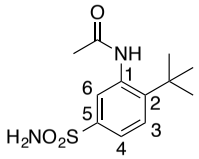
^{1}H NMR - *N*-(2-*tert*-Butyl-5-sulfamoylphenyl)acetamide (72d)

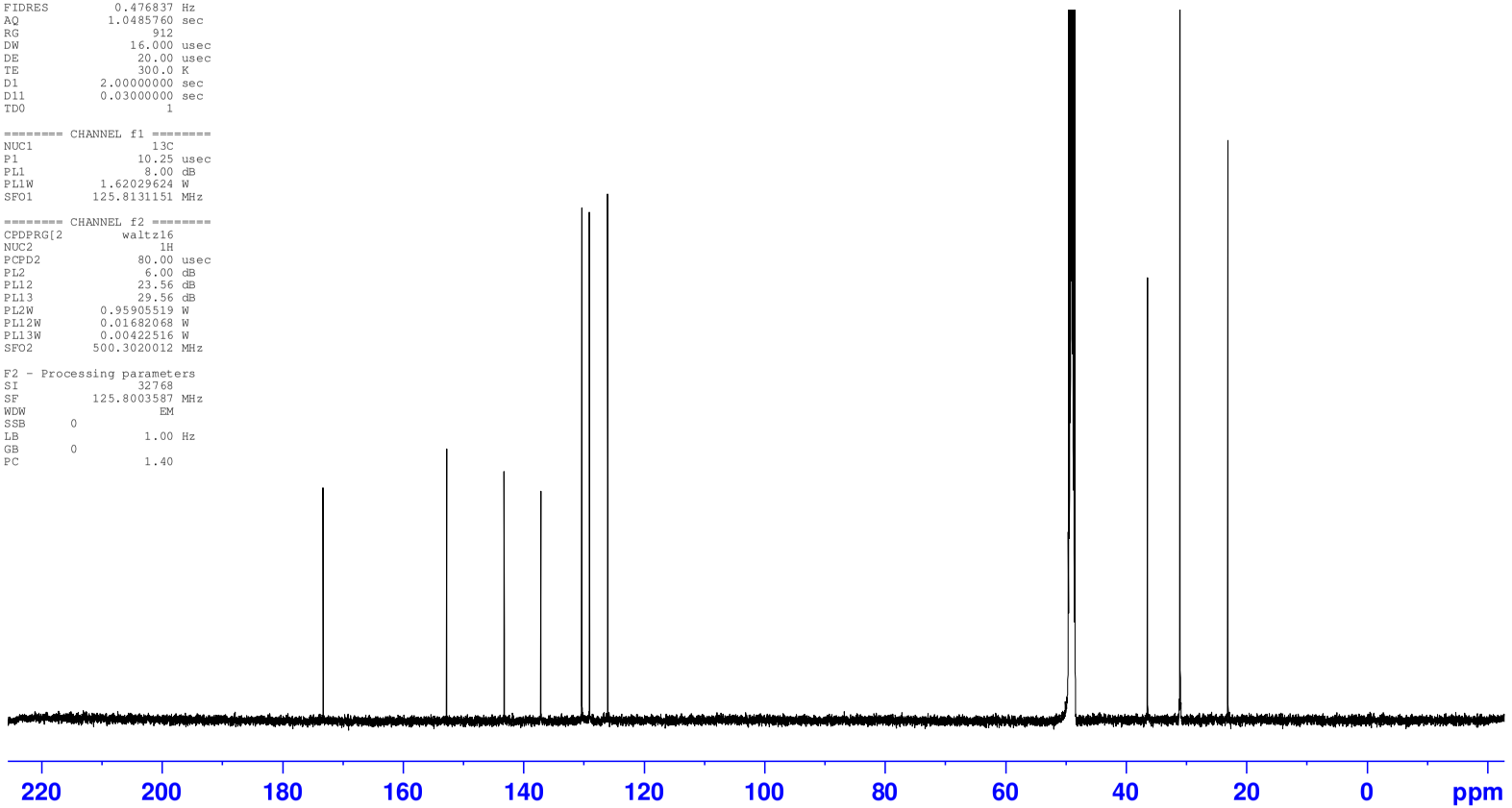
```
Current Data Parameters
NAME      sg26312301 -- SG37A
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130124
Time               8.28
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT             MeOD
NS                    16
DS                     2
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                   3.2
DW                48.400 usec
DE                 6.00 usec
TE                300.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                10.60 usec
PL1                6.00 dB
PL1W         0.95905519 W
SFO1        500.3030896 MHz

F2 - Processing parameters
SI                 32768
SF          500.3000106 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```



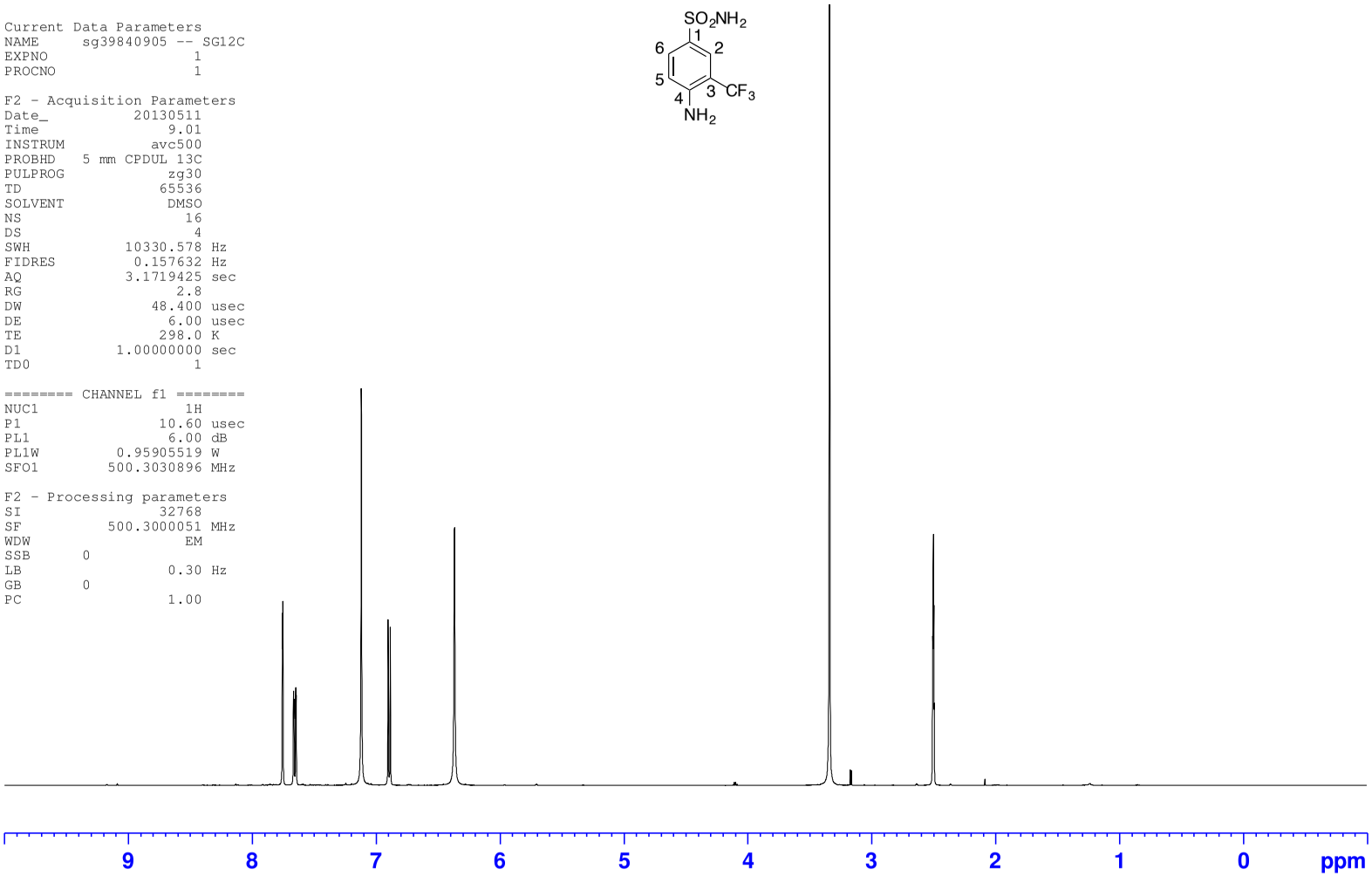
^{13}C NMR - *N*-(2-*tert*-Butyl-5-sulfamoylphenyl)acetamide (72d)

```
Current Data Parameters
NAME      sg26312301 -- SG37A
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_          20130124
Time             13.09
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT            MeOD
NS                 1024
DS                    2
SWH           31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               20.00 usec
TE               300.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                 13C
P1                10.25 usec
PL1                8.00 dB
PL1W         1.62029624 W
SFO1        125.81311151 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                 1H
PCPD2             80.00 usec
PL2                6.00 dB
PL12               23.56 dB
PL13               29.56 dB
PL2W          0.95905519 W
PL12W         0.01682068 W
PL13W         0.00422516 W
SFO2        500.3020012 MHz

F2 - Processing parameters
SI                32768
SF          125.8003587 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```



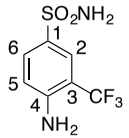


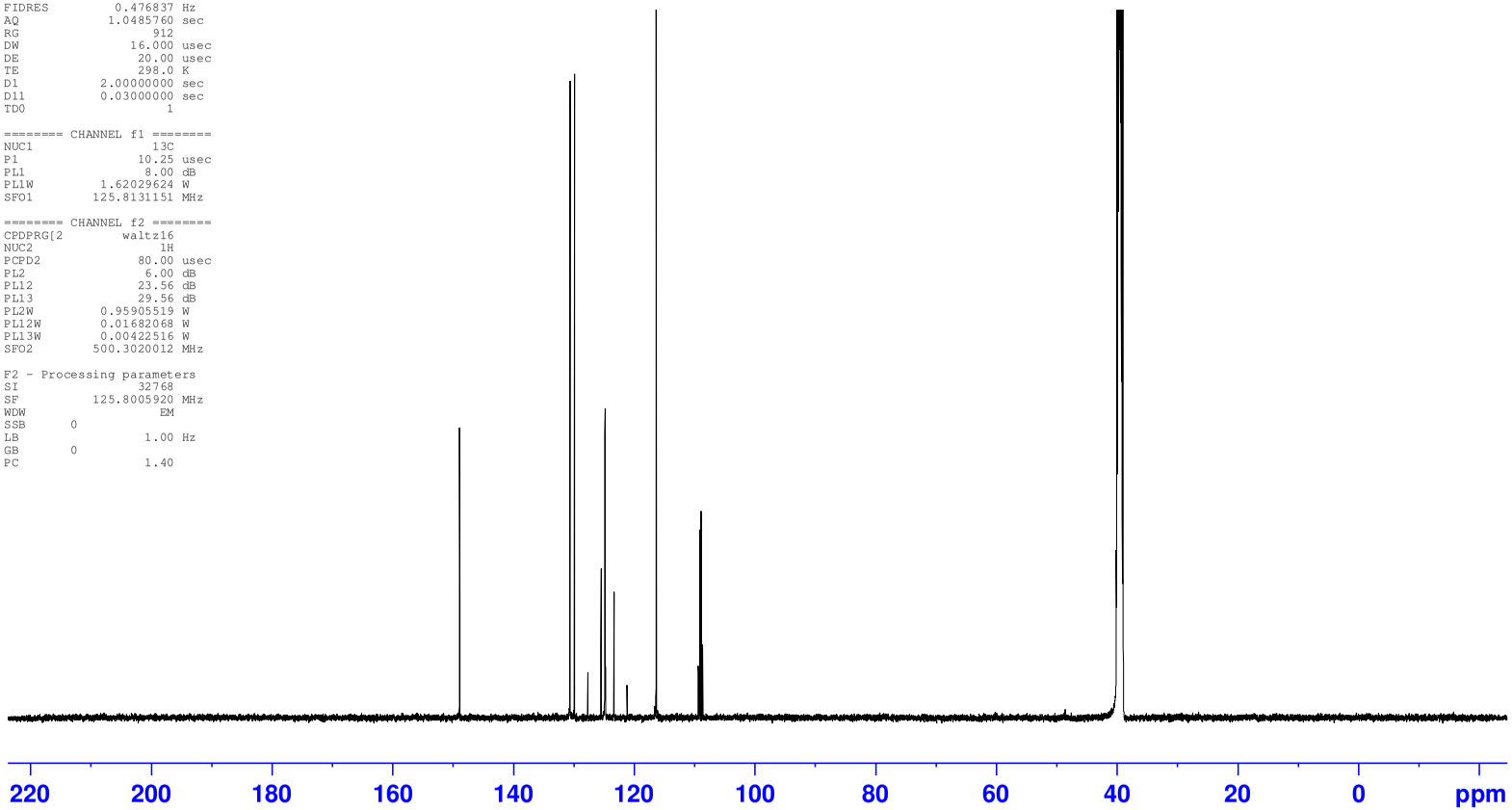
^{1}H NMR - 4-Amino-3-(trifluoromethyl)benzene-1-sulfonamide (74)



^{13}C NMR - 4-Amino-3-(trifluoromethyl)benzene-1-sulfonamide (74)

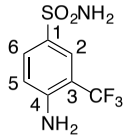
```
Current Data Parameters
NAME      sg39840905 -- SG12C
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_          20130511
Time              11.10
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT            DMSO
NS                 2048
DS                    2
SWH           31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               20.00 usec
TE               298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                 13C
P1                10.25 usec
PL1                8.00 dB
PL1W         1.62029624 W
SFO1        125.8131151 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                 1H
PCPD2             80.00 usec
PL2                6.00 dB
PL12               23.56 dB
PL13               29.56 dB
PL2W          0.95905519 W
PL12W         0.01682068 W
PL13W         0.00422516 W
SFO2        500.3020012 MHz

F2 - Processing parameters
SI                32768
SF          125.8005920 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```





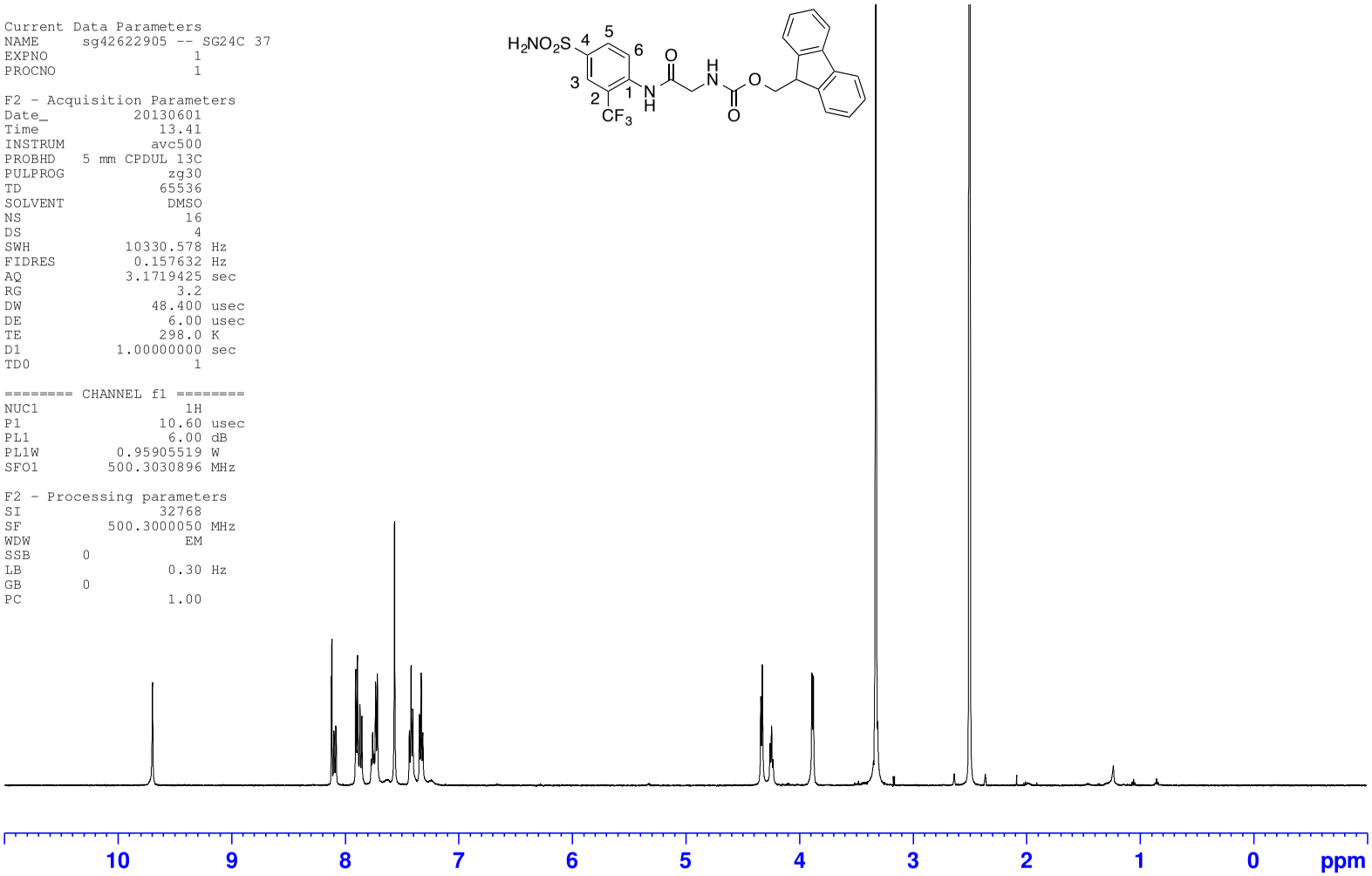
^{19}F NMR - 4-Amino-3-(trifluoromethyl)benzene-1-sulfonamide (74)

```
Current Data Parameters
NAME      May13-2013-12 -- SG12C
EXPNO                    2
PROCNO                   1

F2 - Acquisition Parameters
Date_           20130514
Time              11.34
INSTRUM          avn400
PROBHD   5 mm PABBO BB/
PULPROG          zgflqn
TD               131072
SOLVENT            DMSO
NS                   16
DS                    4
SWH            75000.000 Hz
FIDRES          0.572205 Hz
AQ             0.8738133 sec
RG                205.43
DW                 6.667 usec
DE                 6.50 usec
TE                294.7 K
D1           1.00000000 sec
TD0                  1

======== CHANNEL f1 ========
SFO1         376.5547873 MHz
NUC1                19F
P1                13.50 usec
PLW1        19.00000000 W

F2 - Processing parameters
SI                 65536
SF           376.6112790 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```



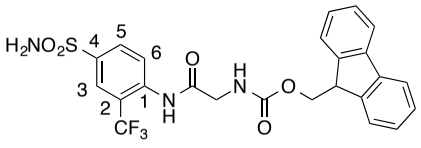
^{1}H NMR - (9*H*-Fluoren-9-yl)methyl (2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethyl)carbamate (75)

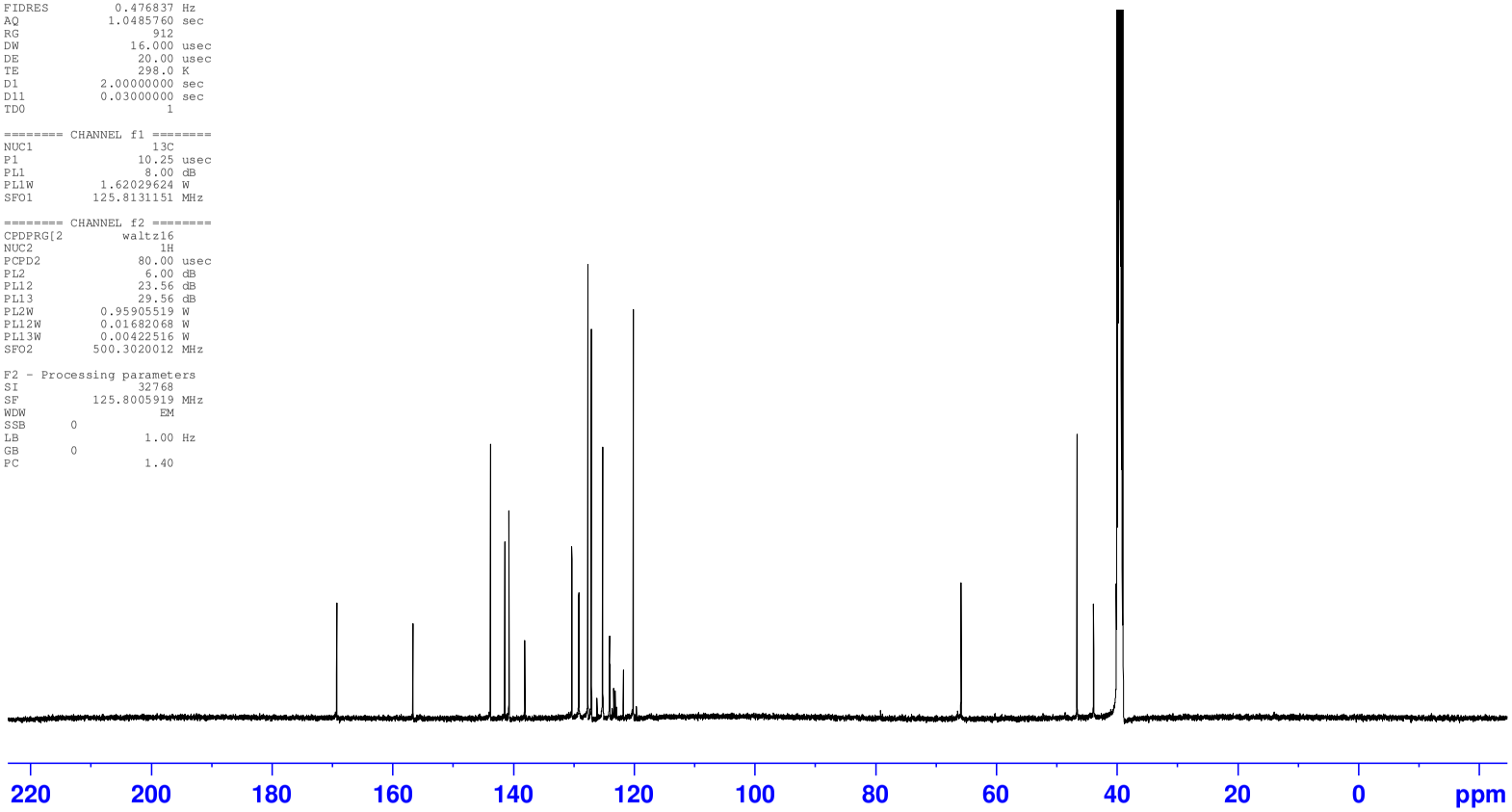
```
Current Data Parameters
NAME      sg42622905 -- SG24C 37
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130601
Time              13.41
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT             DMSO
NS                    16
DS                     4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                   3.2
DW                48.400 usec
DE                 6.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                 10.60 usec
PL1                 6.00 dB
PL1W          0.95905519 W
SFO1         500.3030896 MHz

F2 - Processing parameters
SI                 32768
SF           500.3000050 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```



^{13}C NMR - (9*H*-Fluoren-9-yl)methyl (2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethyl)carbamate (75)

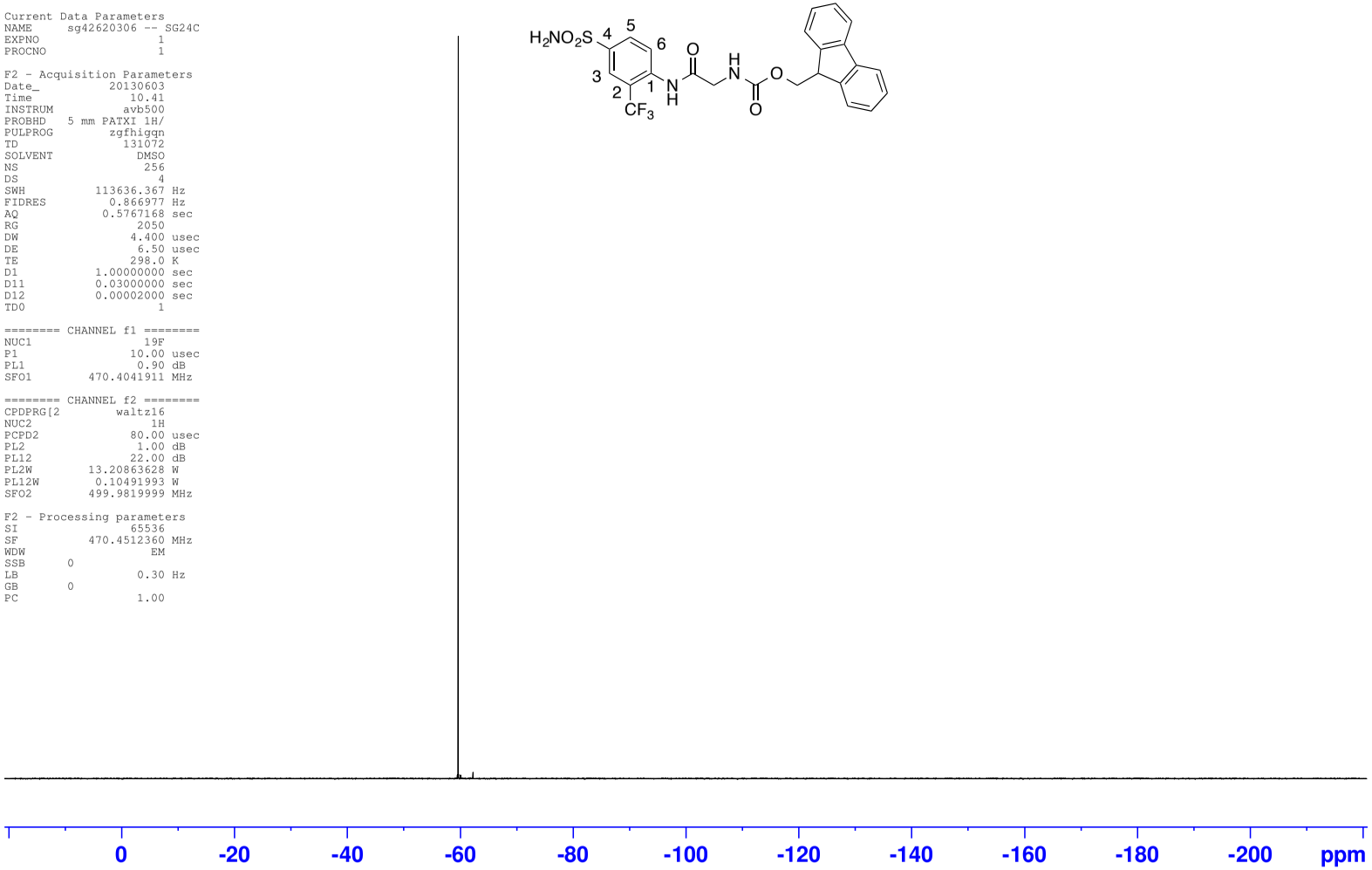
```
Current Data Parameters
NAME      sg43700606_2 -- SG24C
EXPNO                3
PROCNO               1

F2 - Acquisition Parameters
Date_         20130610
Time              9.08
INSTRUM         avc500
PROBHD   5 mm CPDUL 13C
PULPROG         zgpg30
TD               65536
SOLVENT           DMSO
NS                2048
DS                   2
SWH          31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                 912
DW              16.000 usec
DE              20.00 usec
TE              298.0 K
D1          2.00000000 sec
D11         0.03000000 sec
TD0                  1

======== CHANNEL f1 ========
NUC1                13C
P1                10.25 usec
PL1                8.00 dB
PL1W       1.62029624 W
SFO1       125.8131151 MHz

======== CHANNEL f2 ========
CPDPRG[2        waltz16
NUC2                1H
PCPD2             80.00 usec
PL2                6.00 dB
PL12              23.56 dB
PL13              29.56 dB
PL2W       0.95905519 W
PL12W      0.01682068 W
PL13W      0.00422516 W
SFO2       500.3020012 MHz

F2 - Processing parameters
SI                32768
SP         125.8005919 MHz
WDW                EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```



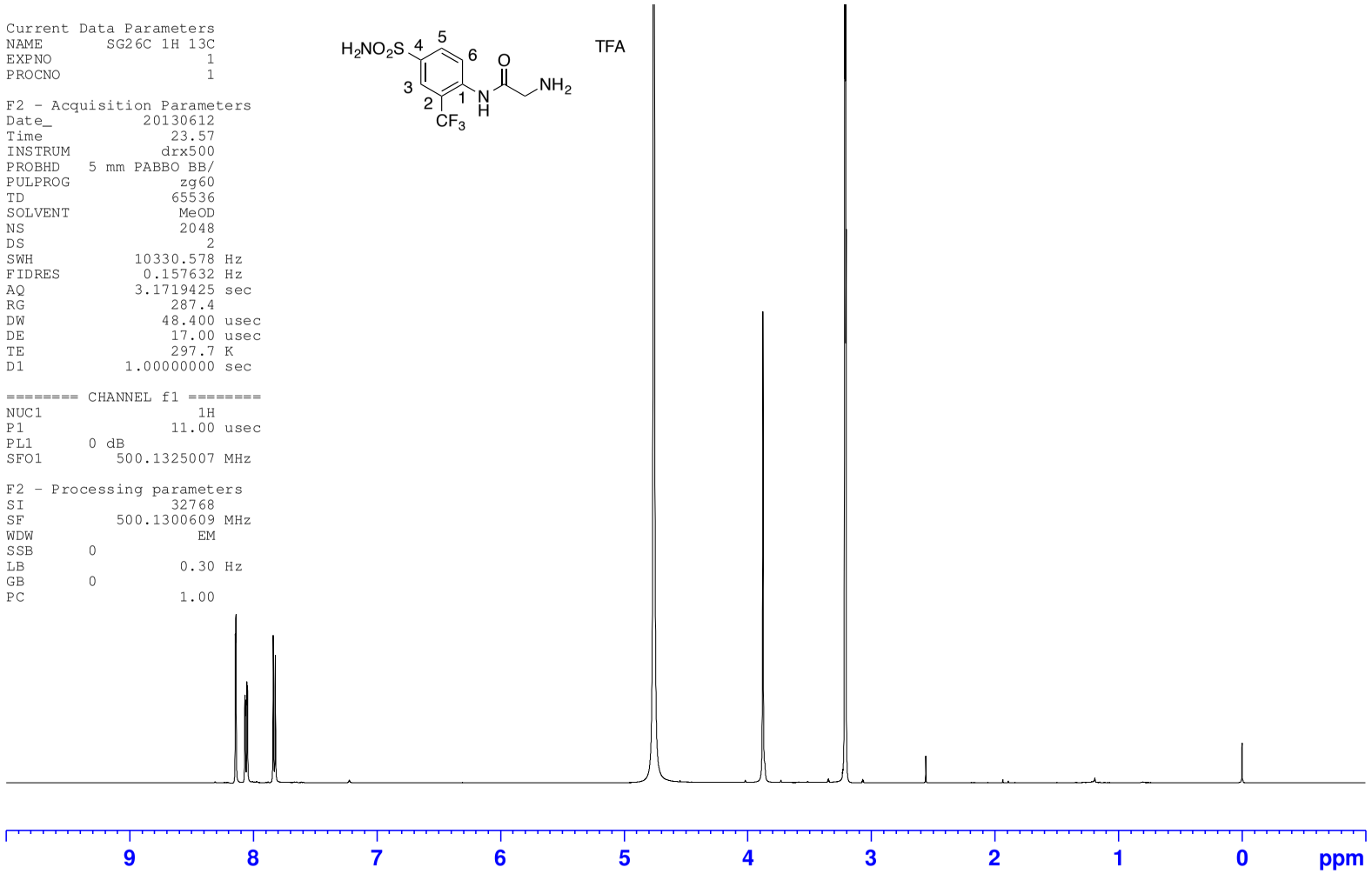


^{19}F NMR - (9*H*-Fluoren-9-yl)methyl (2-oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethyl)carbamate (75)

```
Current Data Parameters
NAME      sg42620306 -- SG24C
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130603
Time              10.41
INSTRUM           avb500
PROBHD   5 mm PATXI 1H/
PULPROG         zgfhigqn
TD               131072
SOLVENT             DMSO
NS                   256
DS                     4
SWH           113636.367 Hz
FIDRES          0.866977 Hz
AQ             0.5767168 sec
RG                  2050
DW                 4.400 usec
DE                 6.50 usec
TE                298.0 K
D1           1.00000000 sec
D11          0.03000000 sec
D12          0.00002000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                 19F
P1                 10.00 usec
PL1                 0.90 dB
SFO1        470.4041911 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                 1H
PCPD2              80.00 usec
PL2                 1.00 dB
PL12               22.00 dB
PL2W        13.20863628 W
PL12W        0.10491993 W
SFO2        499.9819999 MHz

F2 - Processing parameters
SI                 65536
SF          470.4512360 MHz
WDW                  EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```

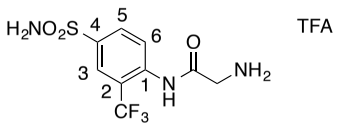


^{1}H NMR - 2-Oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethan-1-aminium 2,2,2-trifluoroacetate (76)

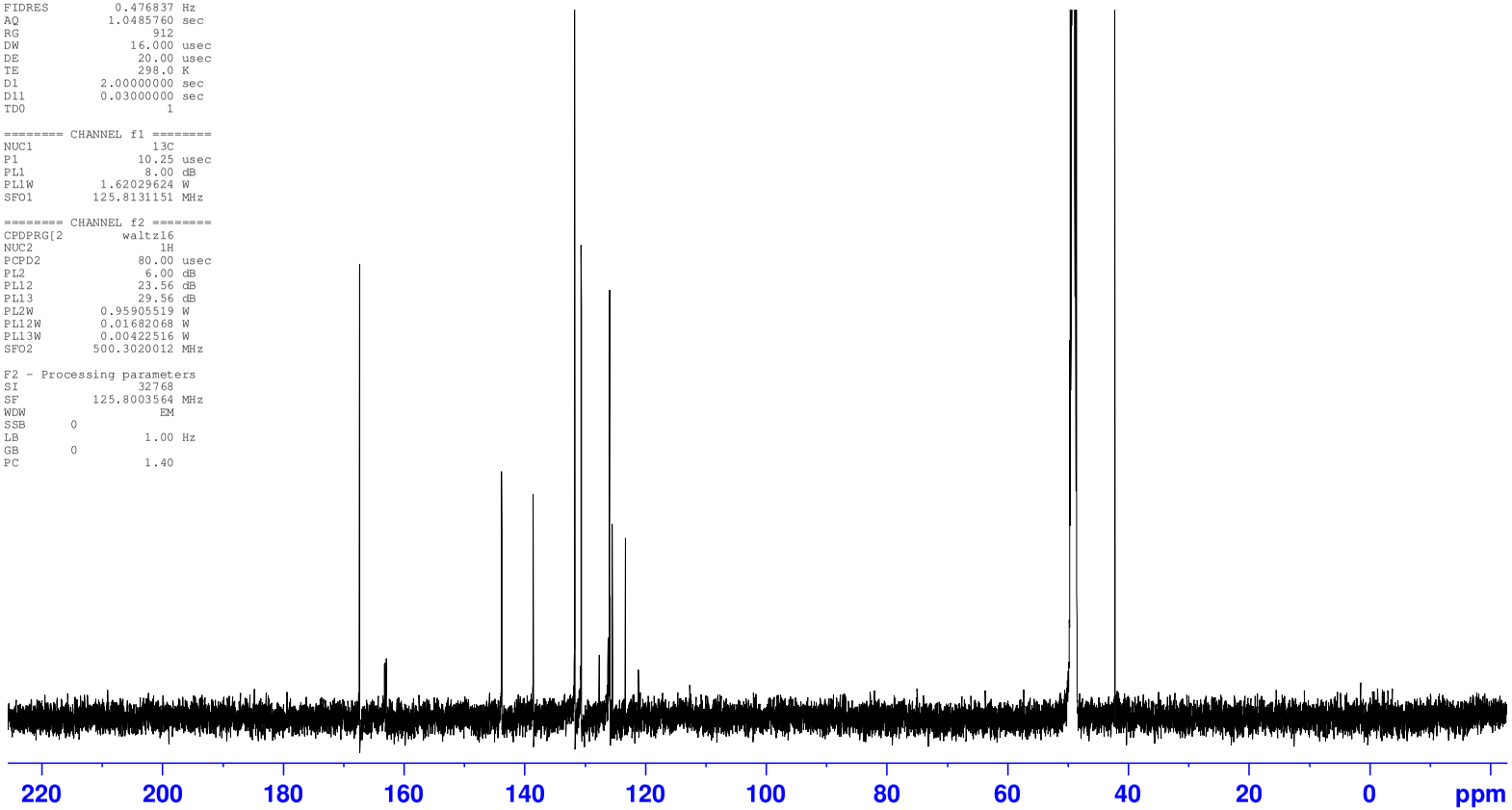


^{13}C NMR - 2-Oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethan-1-aminium 2,2,2-trifluoroacetate (76)

```
Current Data Parameters
NAME      sg43921006_2 SG26C
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_          20130610
Time             21.25
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT            MeOD
NS                 3072
DS                    2
SWH           31250.000 Hz
FIDRES         0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               20.00 usec
TE               298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                 13C
P1                10.25 usec
PL1                8.00 dB
PL1W         1.62029624 W
SFO1        125.81311151 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                 1H
PCPD2             80.00 usec
PL2                6.00 dB
PL12               23.56 dB
PL13               29.56 dB
PL2W          0.95905519 W
PL12W         0.01682068 W
PL13W         0.00422516 W
SFO2        500.3020012 MHz

F2 - Processing parameters
SI                32768
SF           125.8003564 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

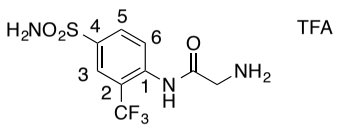


TFA



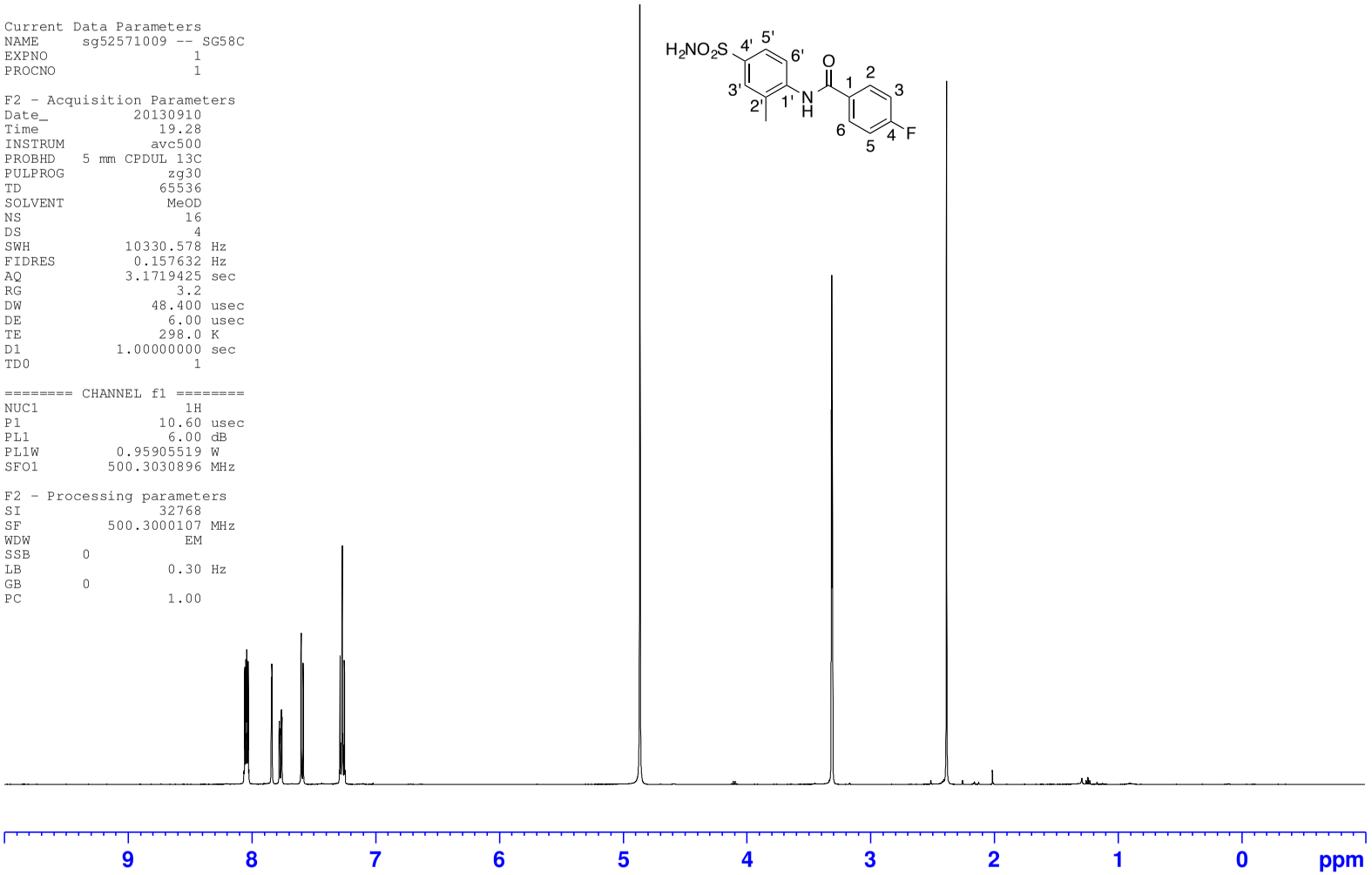
^{19}F 2-Oxo-2-((4-sulfamoyl-2-(trifluoromethyl)phenyl)amino)ethan-1-aminium 2,2,2-trifluoroacetate (76)

```
Current Data Parameters
NAME      sg43921106 -- SG26C
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130611
Time              10.45
INSTRUM           avb500
PROBHD   5 mm PATXI 1H/
PULPROG         zgfhigqn
TD               131072
SOLVENT             MeOD
NS                   128
DS                     4
SWH           113636.367 Hz
FIDRES          0.866977 Hz
AQ             0.5767168 sec
RG                  2050
DW                 4.400 usec
DE                 6.50 usec
TE                298.0 K
D1           1.00000000 sec
D11          0.03000000 sec
D12          0.00002000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                 19F
P1                 10.00 usec
PL1                 0.90 dB
SFO1         470.4041911 MHz

======== CHANNEL f2 ========
CPDPRG[2          waltz16
NUC2                 1H
PCPD2              80.00 usec
PL2                 1.00 dB
PL12               22.00 dB
PL2W         13.20863628 W
PL12W         0.10491993 W
SFO2         499.9819999 MHz

F2 - Processing parameters
SI                 65536
SF           470.4512360 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                10.00
```



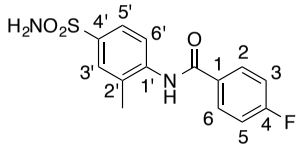
TFA

^{1}H NMR - 4-Fluoro-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide (78a)



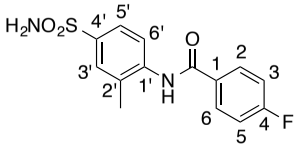
^{13}C NMR - 4-Fluoro-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide (78a)

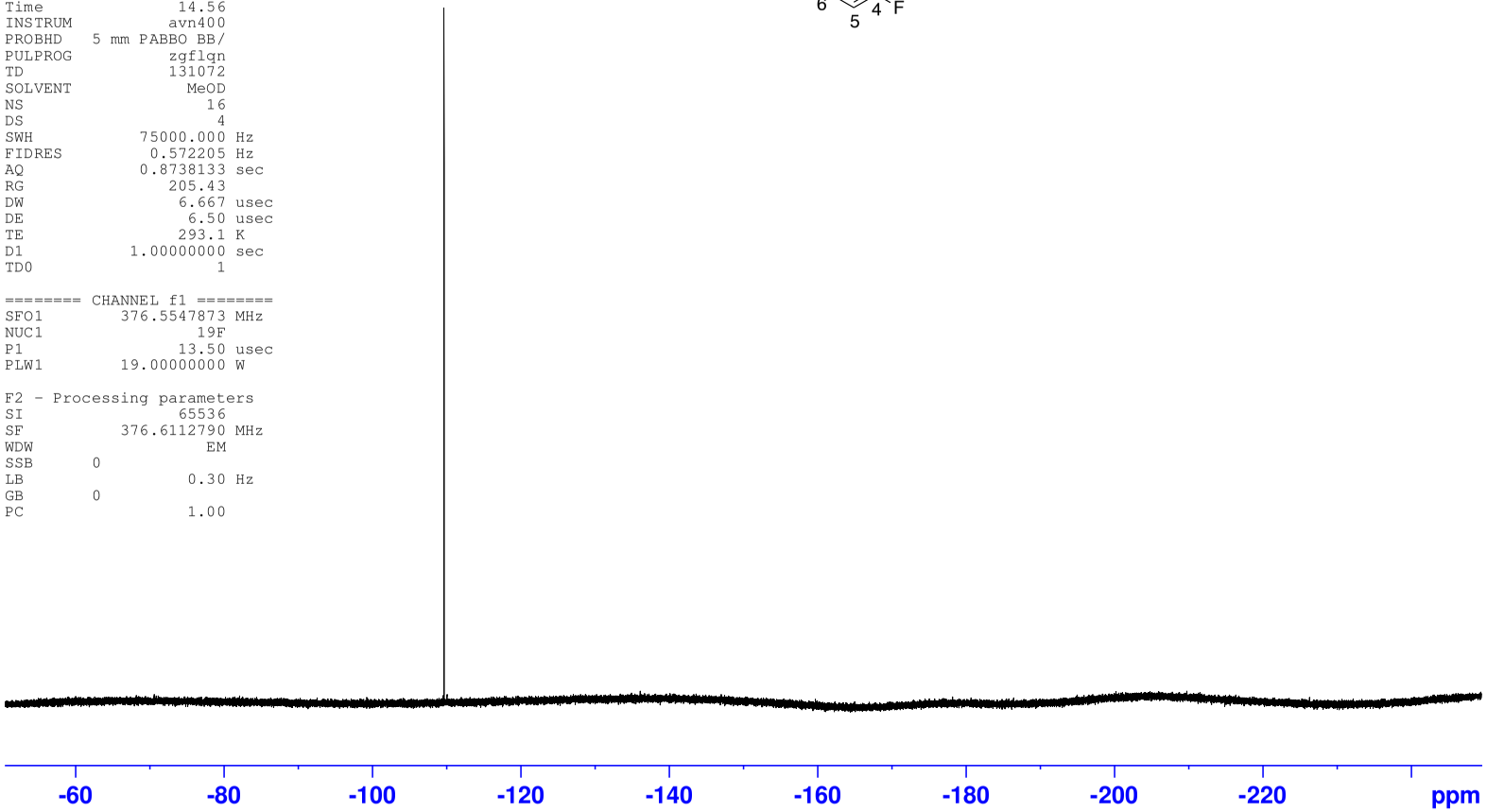
```
Current Data Parameters
NAME      sg52571009 -- SG58C
EXPNO                  4
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130910
Time              21.37
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG           zgpg30
TD                65536
SOLVENT             MeOD
NS                  2048
DS                     2
SWH           31250.000 Hz
FIDRES         0.476837 Hz
AQ            1.0485760 sec
RG                   912
DW               16.000 usec
DE               20.00 usec
TE               298.1 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                 13C
P1                10.25 usec
PL1                8.00 dB
PL1W         1.62029624 W
SFO1        125.8131151 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                 1H
PCPD2             80.00 usec
PL2                6.00 dB
PL12               23.56 dB
PL13               29.56 dB
PL2W         0.95905519 W
PL12W        0.01682068 W
PL13W        0.00422516 W
SFO2        500.3020012 MHz

F2 - Processing parameters
SI                 32768
SP          125.8003593 MHz
WDW                   EM
SSB       0
LB                 1.00 Hz
GB        0
PC                 1.40
```



^{19}F NMR - 4-Fluoro-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide (78a)

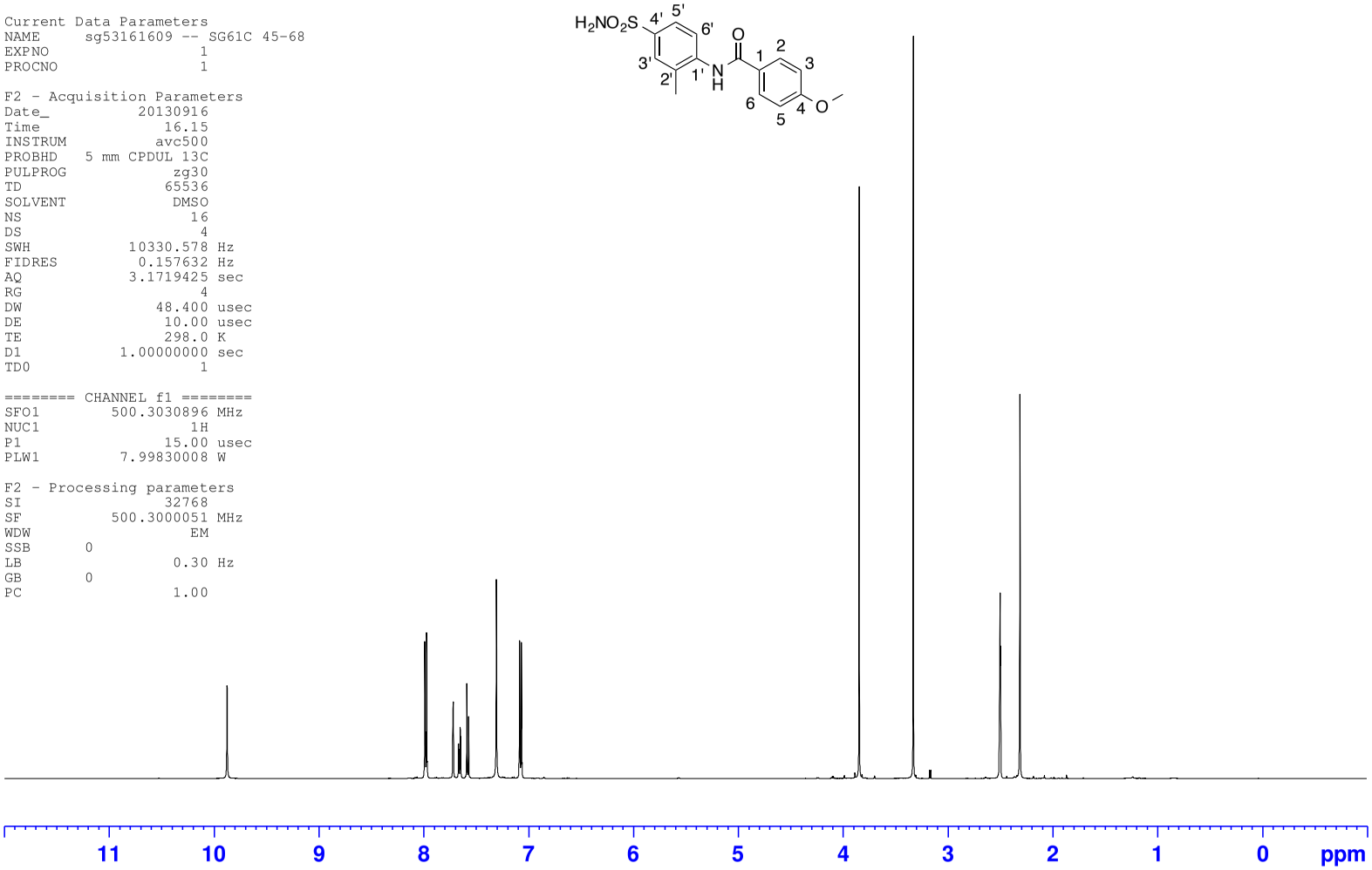
```
Current Data Parameters
NAME      Sep12-2013-42 -- SG58C
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130912
Time              14.56
INSTRUM          avn400
PROBHD   5 mm PABBO BB/
PULPROG          zgflqn
TD               131072
SOLVENT            MeOD
NS                   16
DS                    4
SWH            75000.000 Hz
FIDRES          0.572205 Hz
AQ             0.8738133 sec
RG                205.43
DW                 6.667 usec
DE                 6.50 usec
TE                293.1 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         376.5547873 MHz
NUC1                19F
P1                13.50 usec
PLW1        19.00000000 W

F2 - Processing parameters
SI                 65536
SF           376.6112790 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```





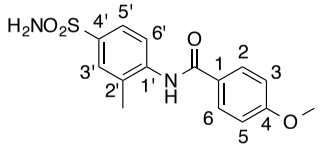
^{1}H NMR - 4-Methoxy-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide (78b)

```
Current Data Parameters
NAME      sg53161609 -- SG61C 45-68
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_           20130916
Time              16.15
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT             DMSO
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                 15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 32768
SF           500.3000051 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```

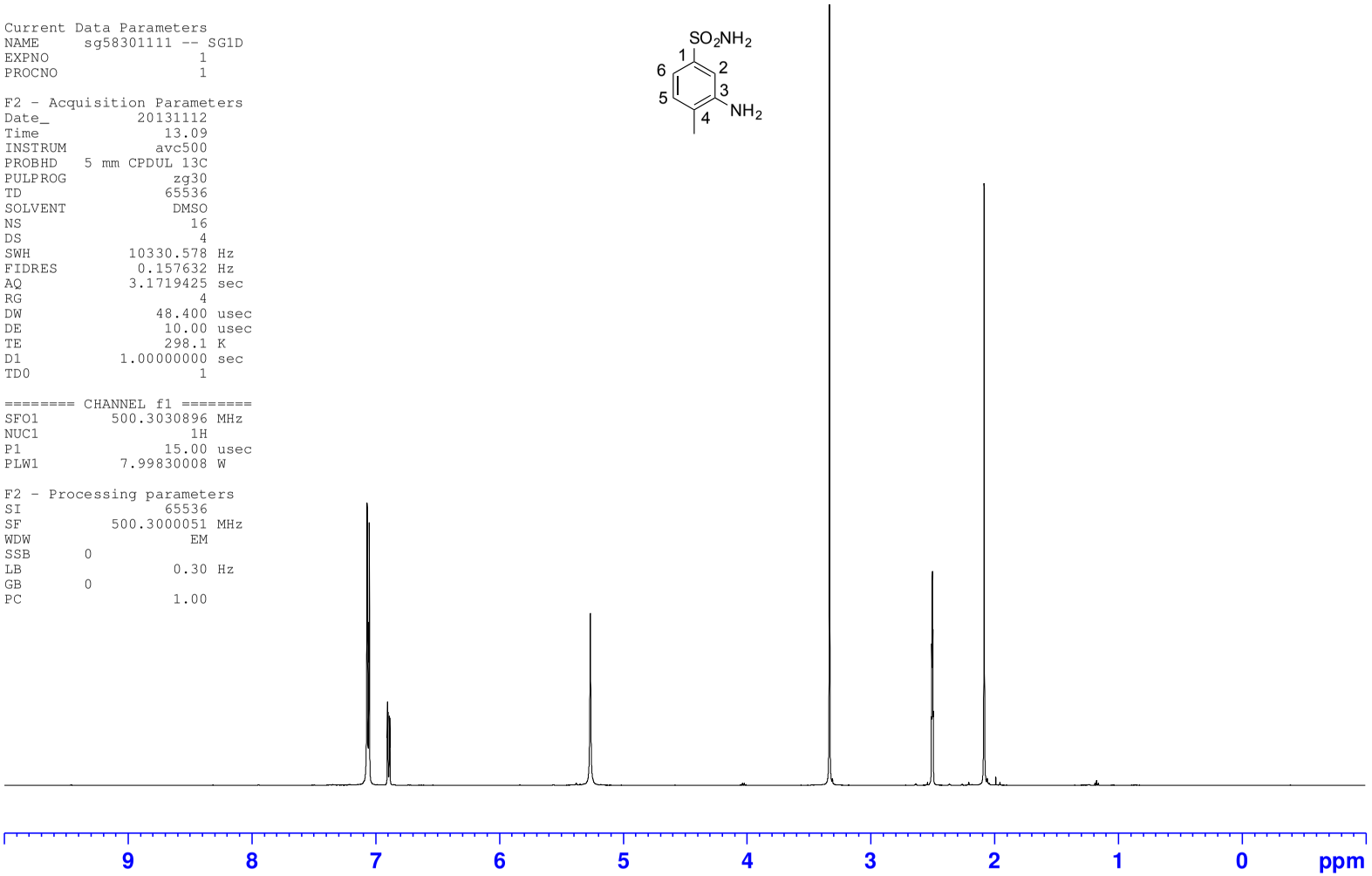


^{13}C NMR - 4-Methoxy-*N*-(2'-methyl-4'-sulfamoylphenyl)benzamide (78b)

```
Current Data Parameters
NAME      sg53161609 -- SG61C 45-68
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_           20130916
Time              17.03
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             DMSO
NS                  512
DS                    2
SWH           31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               18.00 usec
TE                298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        125.8131151 MHz
NUC1                13C
P1               10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2                1H
CPDPRG[2         waltz16
PCPD2              80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8005945 MHz
WDW                 EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

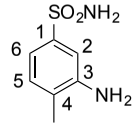


^{1}H NMR - 3-Amino-4-methylbenzene-1-sulfonamide (80)



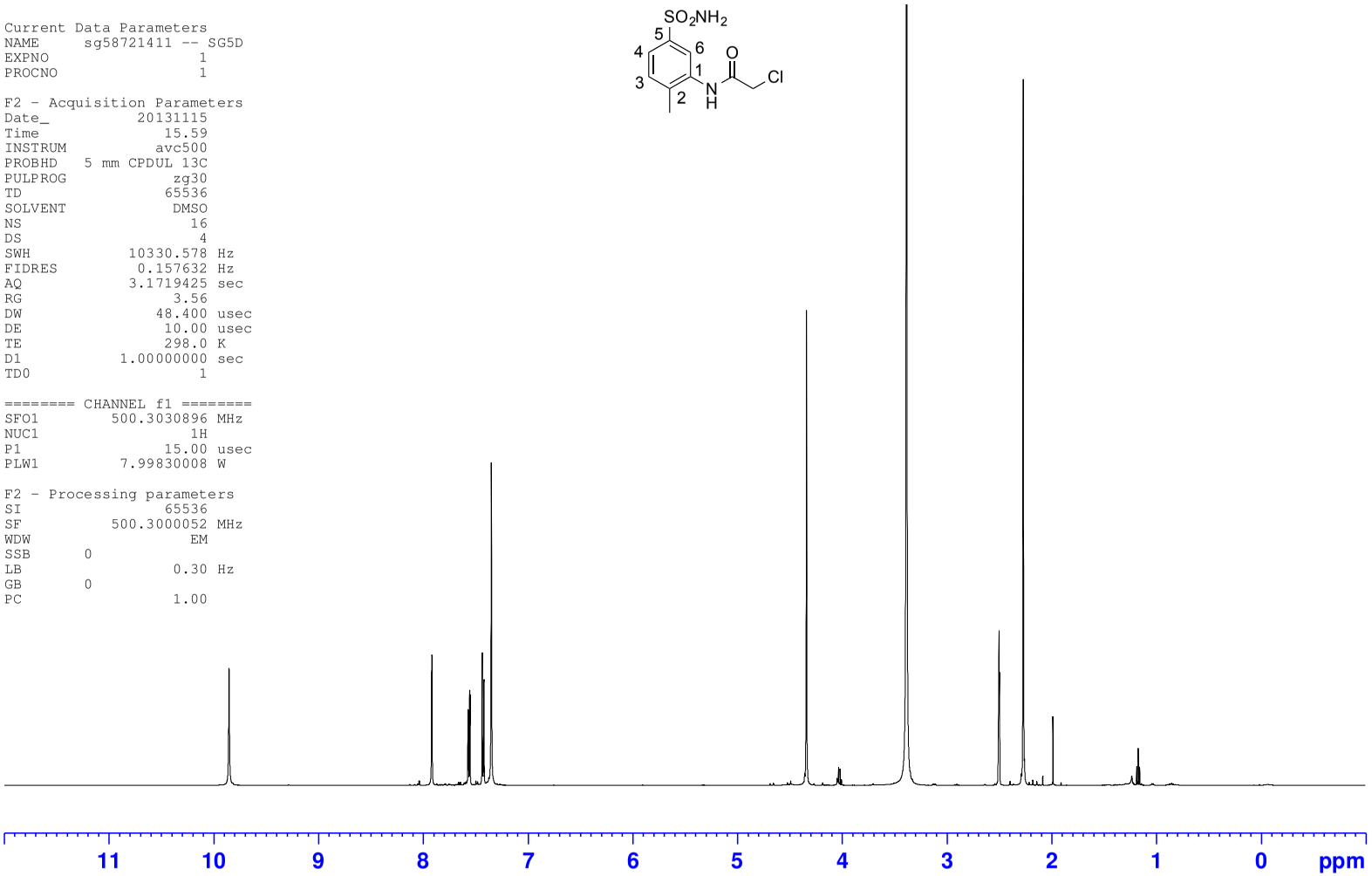
^{13}C NMR - 3-Amino-4-methylbenzene-1-sulfonamide (80)

```
Current Data Parameters
NAME      sg58301111 -- SG1D
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_           20131112
Time              13.59
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             DMSO
NS                    584
DS                      2
SWH           31250.000 Hz
FIDRES         0.476837 Hz
AQ             1.0485760 sec
RG                  912
DW                16.000 usec
DE                18.00 usec
TE                298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        125.8131151 MHz
NUC1                13C
P1                10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2                1H
CPDPRG[2         waltz16
PCPD2             80.00 usec
PLW2          7.99830008 W
PLW12         0.28119001 W
PLW13         0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8005938 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```



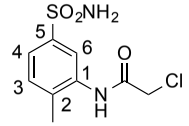
^{1}H NMR - 2-Chloro-*N*-(2-methyl-5-sulfamoylphenyl)acetamide (81)

```
Current Data Parameters
NAME      sg58721411 -- SG5D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20131115
Time               15.59
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT             DMSO
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                  3.56
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                  15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.3000052 MHz
WDW                   EM
SSB      0
LB                  0.30 Hz
GB       0
PC                  1.00
```



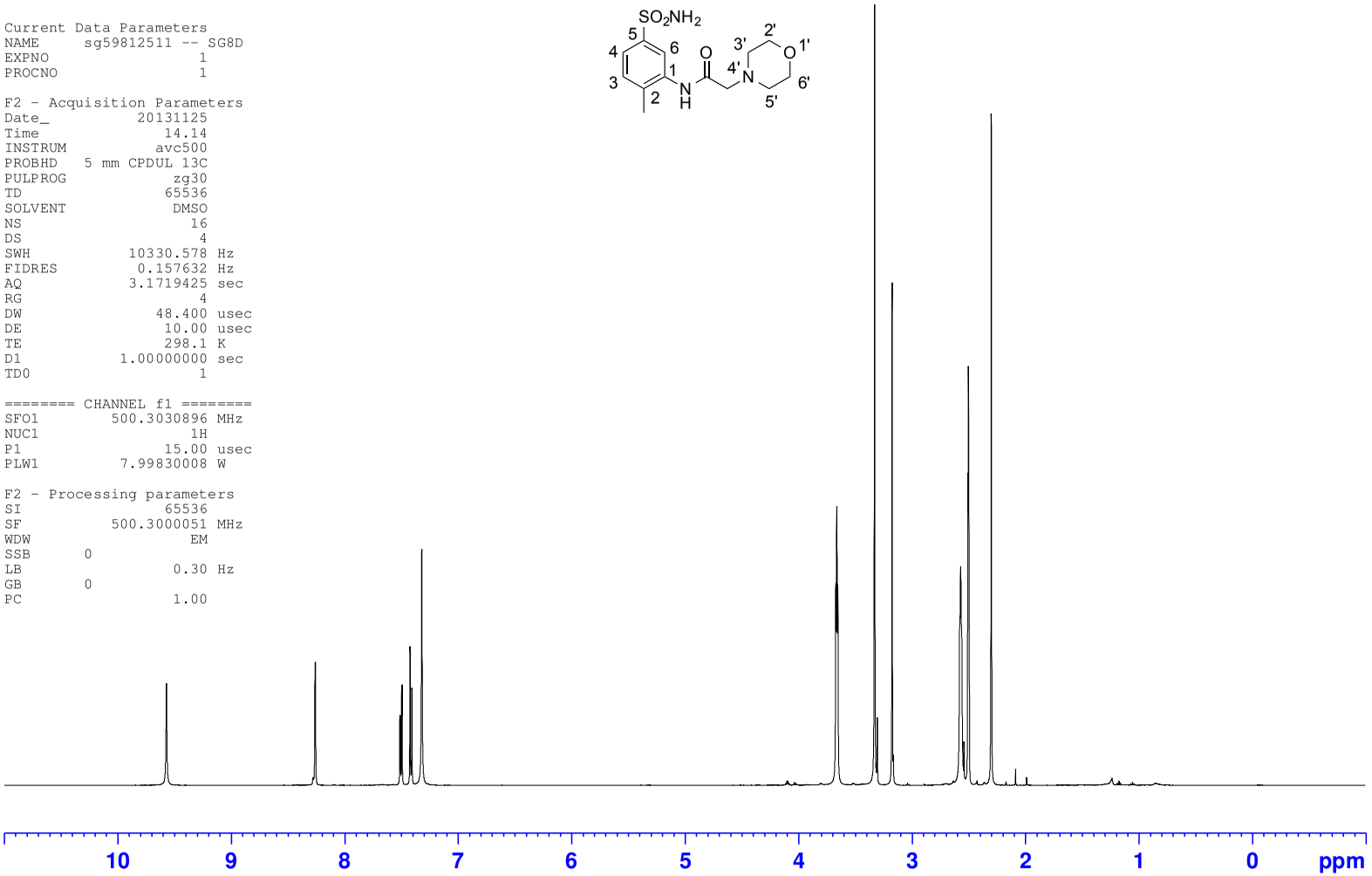
^{13}C NMR - 2-Chloro-*N*-(2-methyl-5-sulfamoylphenyl)acetamide (81)

```
Current Data Parameters
NAME      sg58721411 -- SG5D
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_           20131115
Time              17.14
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             DMSO
NS                 1024
DS                    2
SWH           31250.000 Hz
FIDRES         0.476837 Hz
AQ             1.0485760 sec
RG                  912
DW               16.000 usec
DE               18.00 usec
TE               298.0 K
D1            2.00000000 sec
D11           0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         125.8131151 MHz
NUC1                13C
P1                10.00 usec
PLW1         20.18400002 W

======== CHANNEL f2 ========
SFO2         500.3020012 MHz
NUC2                1H
CPDPRG[2         waltz16
PCPD2              80.00 usec
PLW2          7.99830008 W
PLW12         0.28119001 W
PLW13         0.17996000 W

F2 - Processing parameters
SI                 32768
SF           125.8005867 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```



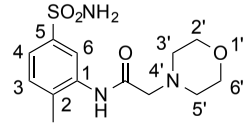
^{1}H NMR - *N*-(2-Methyl-5-sulfamoylphenyl)-2-(morpholin-4'-yl)acetamide (82)

```
Current Data Parameters
NAME      sg59812511 -- SG8D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20131125
Time               14.14
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT             DMSO
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.1 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                  15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.3000051 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```

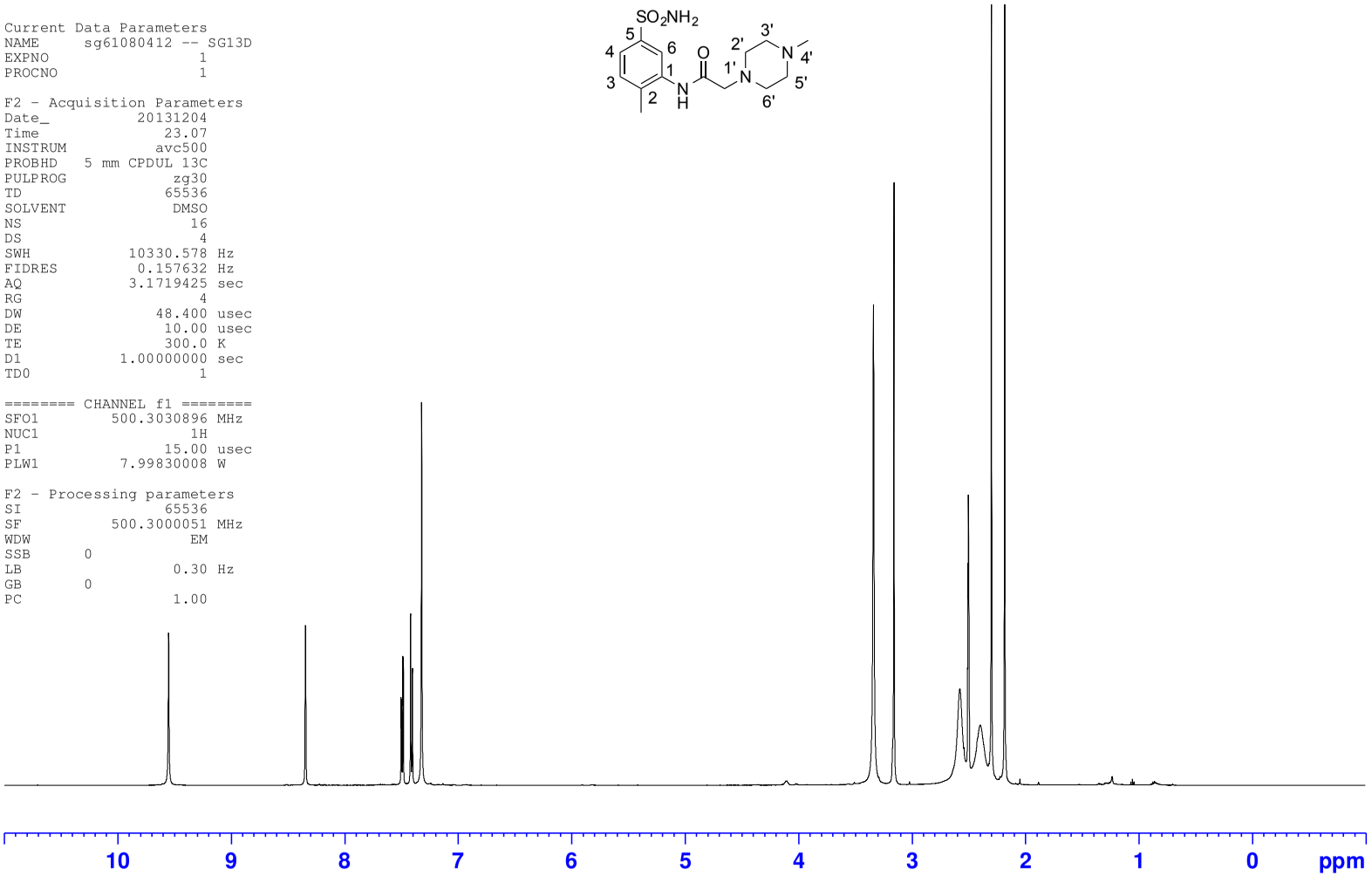


^{13}C NMR - *N*-(2-Methyl-5-sulfamoylphenyl)-2-(morpholin-4'-yl)acetamide (82)

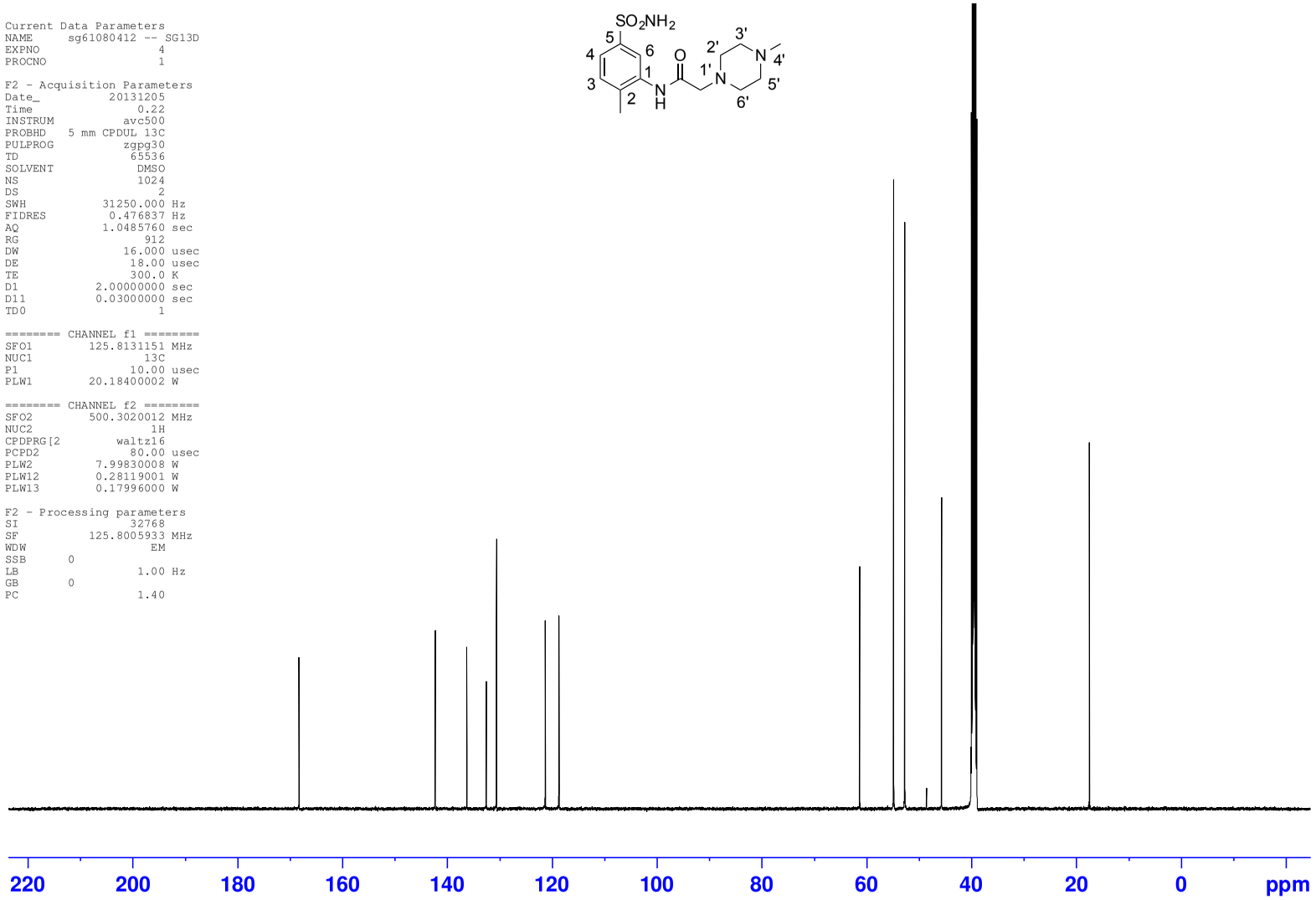
```
Current Data Parameters
NAME      sg59812511 -- SG8D
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_           20131125
Time              15.13
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             DMSO
NS                    352
DS                      2
SWH           31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               18.00 usec
TE               298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                  1

======== CHANNEL f1 ========
SFO1        125.8131151 MHz
NUC1               13C
P1               10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2               1H
CPDPRG[2         waltz16
PCPD2             80.00 usec
PLW2          7.99830008 W
PLW12         0.28119001 W
PLW13         0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8005938 MHz
WDW                EM
SSB      0
LB                1.00 Hz
GB       0
PC                1.40
```



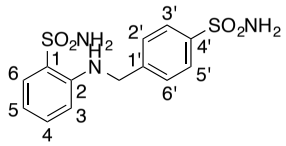
^{1}H NMR - *N*-(2-Methyl-5-sulfamoylphenyl)-2-(4'-methylpiperazin-1'-yl)acetamide (83)



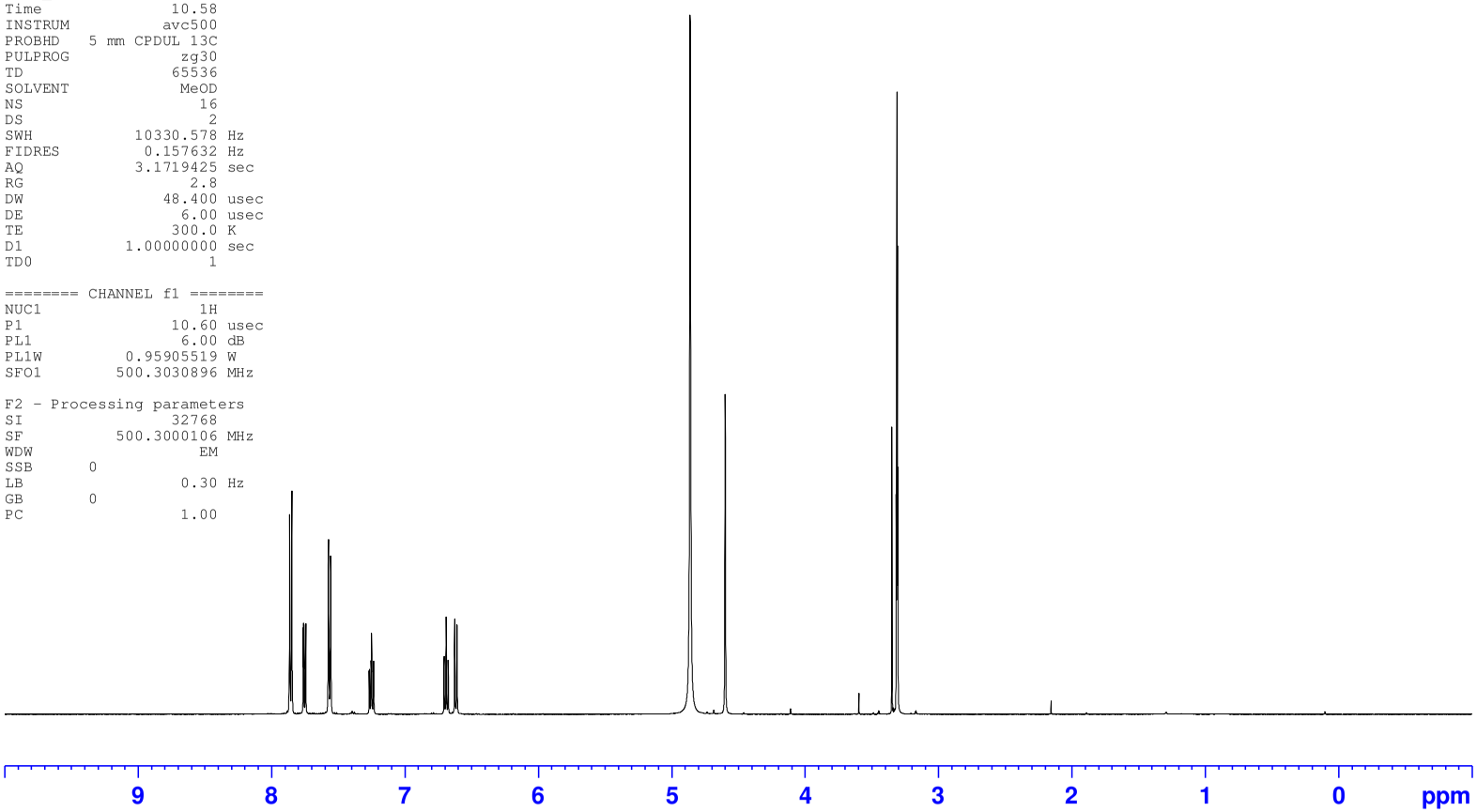
^{13}C NMR - *N*-(2-Methyl-5-sulfamoylphenyl)-2-(4'-methylpiperazin-1'-yl)acetamide (83)



^{1}H NMR - 2-((4’-Sulfamoylbenzyl)amino)benzene-1-sulfonamide (85)

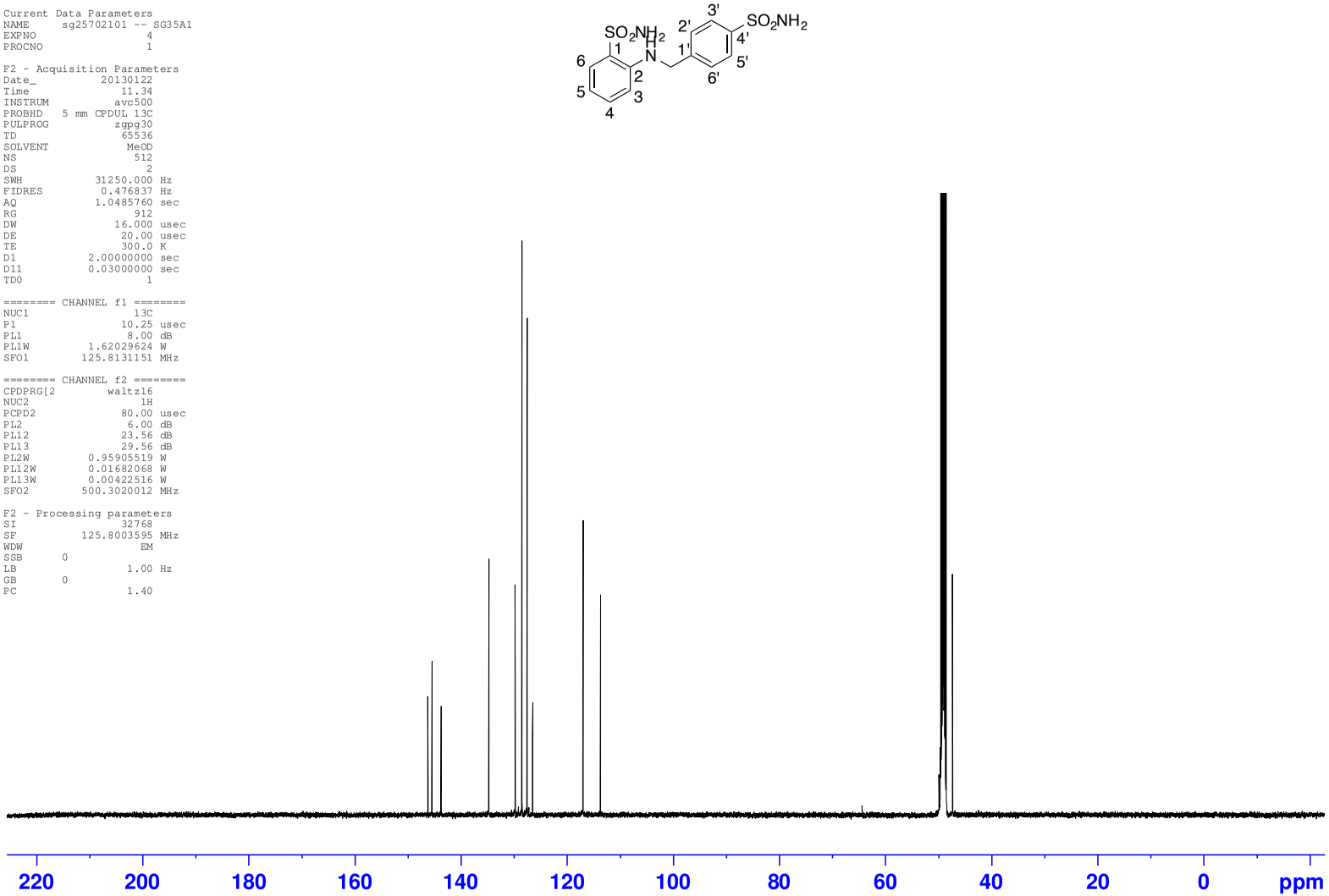



```
Current Data Parameters
NAME      sg25702101 -- SG35A1
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20130122
Time              10.58
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                65536
SOLVENT            MeOD
NS                   16
DS                    2
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                  2.8
DW                48.400 usec
DE                 6.00 usec
TE                300.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                  1H
P1                10.60 usec
PL1                6.00 dB
PL1W         0.95905519 W
SFO1        500.3030896 MHz

F2 - Processing parameters
SI                32768
SF          500.3000106 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```



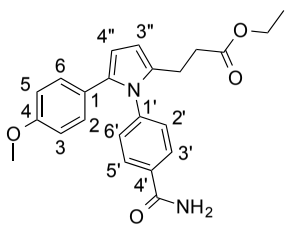
^{13}C NMR - 2-((4’-Sulfamoylbenzyl)amino)benzene-1-sulfonamide (85)

```
Current Data Parameters
NAME      sg25702101 -- SG35A1
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_           20130122
Time              11.34
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             MeOD
NS                  512
DS                    2
SWH           31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               20.00 usec
TE               300.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                 13C
P1                10.25 usec
PL1                8.00 dB
PL1W         1.62029624 W
SFO1        125.8131151 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                 1H
PCPD2             80.00 usec
PL2                6.00 dB
PL12               23.56 dB
PL13               29.56 dB
PL2W         0.95905519 W
PL12W        0.01682068 W
PL13W        0.00422516 W
SFO2        500.3020012 MHz

F2 - Processing parameters
SI                32768
SF          125.8003595 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

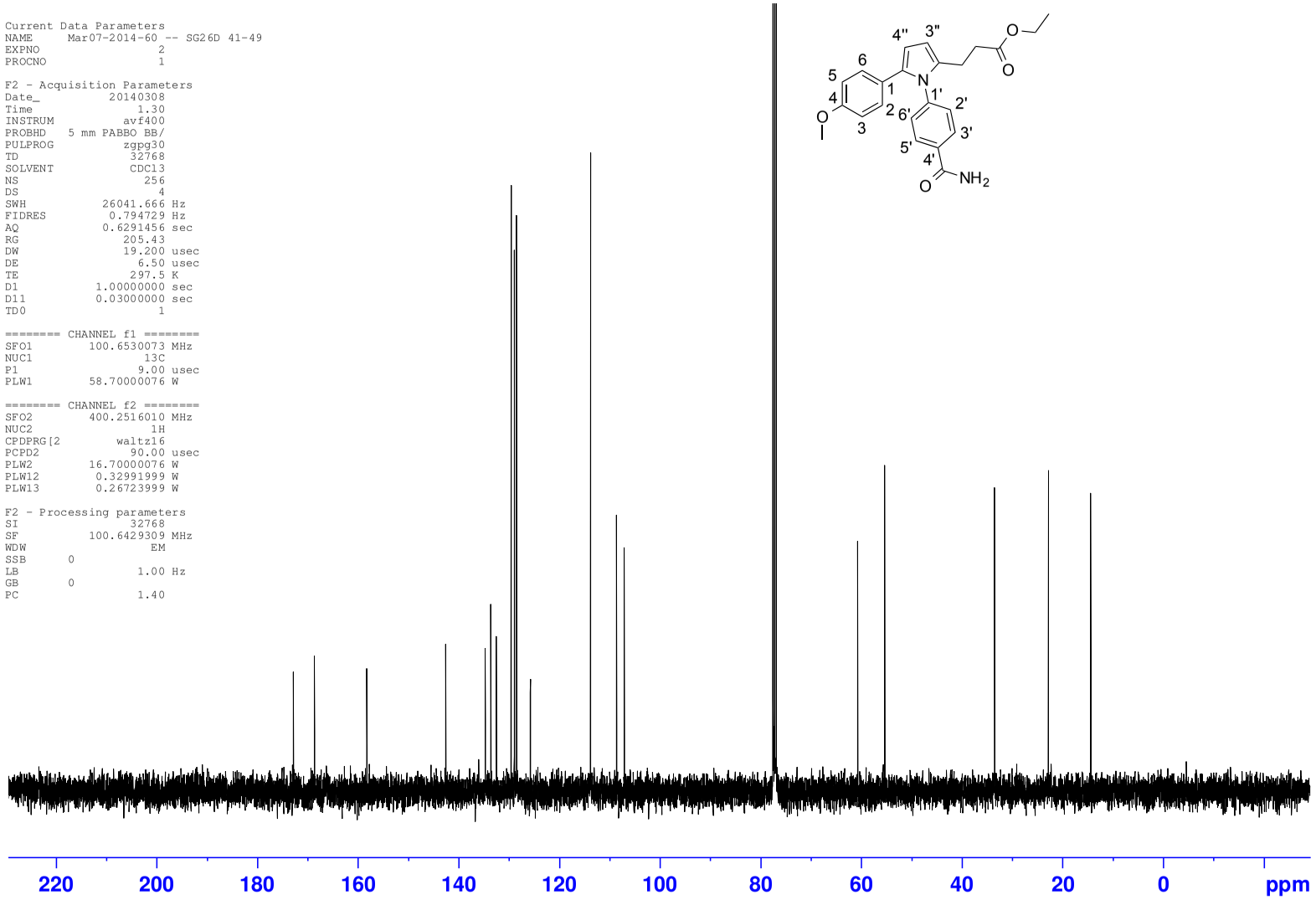


^{1}H NMR - Ethyl 3''-[1-(4'-carbamoylphenyl-5''-(4-methoxyphenyl)-1*H*-pyrrol-2''-yl]propanoate (93)

```
Current Data Parameters
NAME      Mar07-2014-60 -- SG26D 41-49
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140307
Time               15.36
INSTRUM          avf400
PROBHD   5 mm PABBO BB/
PULPROG             zg60
TD                 65536
SOLVENT            CDCl3
NS                    16
DS                     2
SWH             8012.820 Hz
FIDRES          0.122266 Hz
AQ              4.0894465 sec
RG                    93
DW                62.400 usec
DE                 6.50 usec
TE                295.5 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         400.2524015 MHz
NUC1                   1H
P1                 12.65 usec
PLW1         16.70000076 W

F2 - Processing parameters
SI                 32768
SF           400.2500118 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```

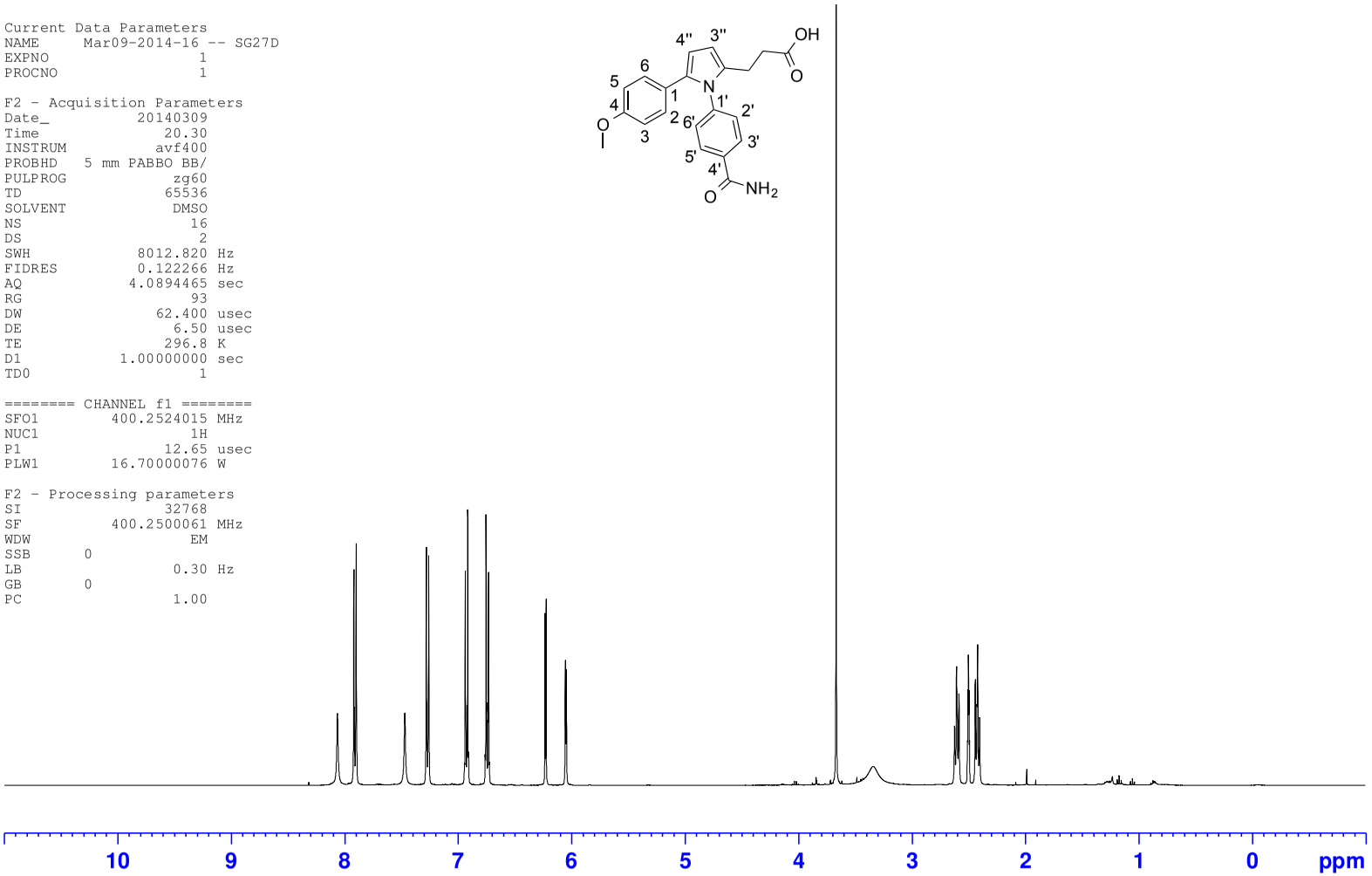


^{13}C NMR - Ethyl 3''-[1-(4'-carbamoylphenyl-5''-(4-methoxyphenyl)-1*H*-pyrrol-2''-yl]propanoate (93)

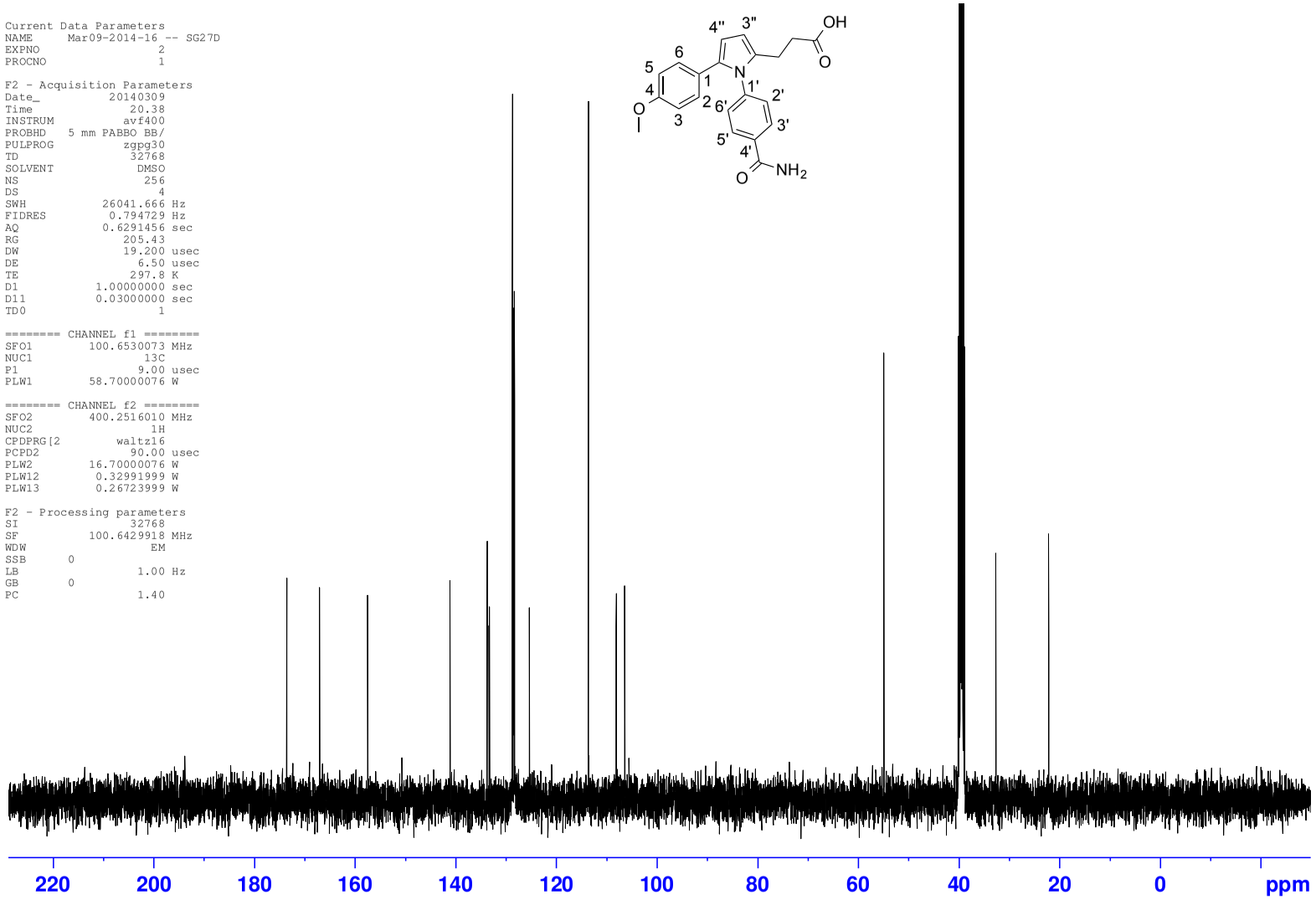


^{1}H NMR - 3’’-[1-(4’-Carbamoylphenyl)-5’’-(4-methoxyphenyl)-1*H*-pyrrol-2’’-yl]propanoic acid (94)

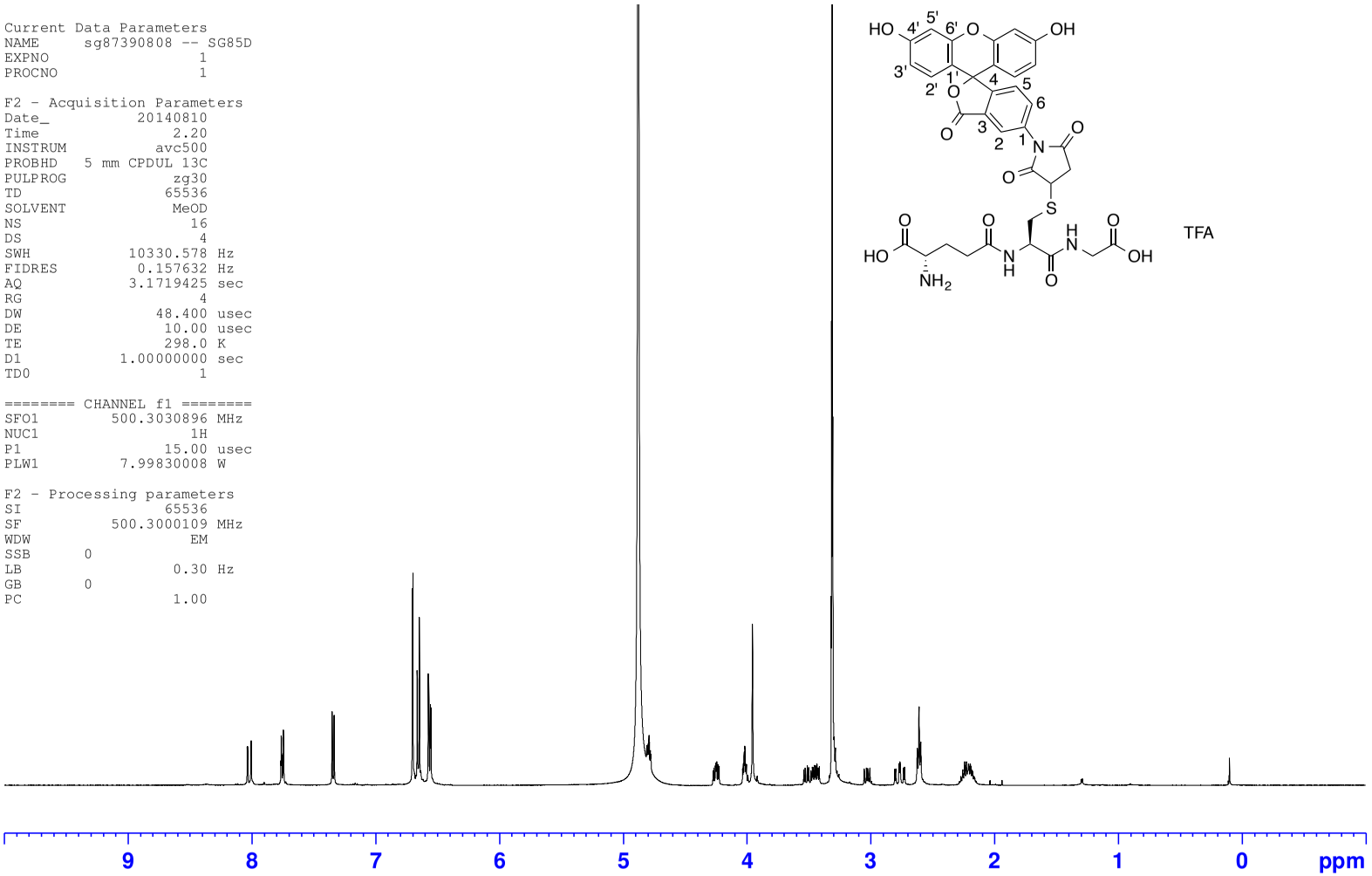
```
Current Data Parameters
NAME      Mar09-2014-16 -- SG27D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140309
Time              20.30
INSTRUM          avf400
PROBHD   5 mm PABBO BB/
PULPROG            zg60
TD                65536
SOLVENT            DMSO
NS                   16
DS                    2
SWH             8012.820 Hz
FIDRES          0.122266 Hz
AQ              4.0894465 sec
RG                   93
DW                62.400 usec
DE                 6.50 usec
TE                296.8 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         400.2524015 MHz
NUC1                1H
P1                12.65 usec
PLW1         16.70000076 W

F2 - Processing parameters
SI                32768
SF           400.2500061 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```



^{13}C NMR - 3''-[1-(4'-Carbamoylphenyl)-5''-(4-methoxyphenyl)-1*H*-pyrrol-2''-yl]propanoic acid (94)

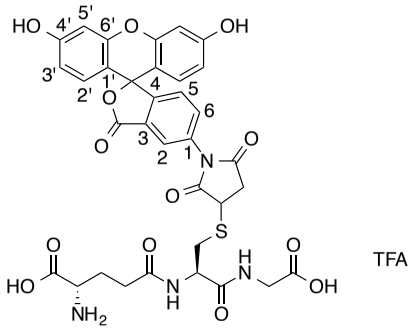


^{1}H NMR - *S*-(*N*-(5-Fluoresceinyl)succinimido) L-glutathione 2,2,2-trifluoroacetate (95)



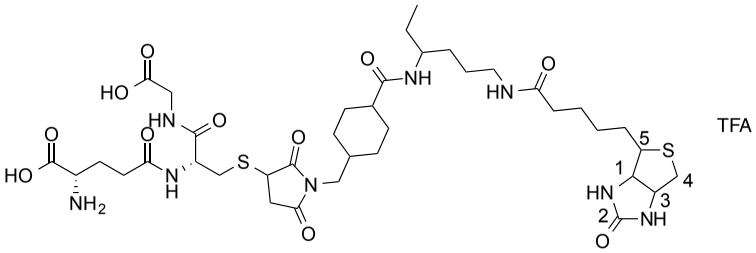
^{13}C NMR - *S*-(*N*-(5-Fluoresceinyl)succinimido) L-glutathione 2,2,2-trifluoroacetate (95)

```
Current Data Parameters
NAME      sg87390808 -- SG85D
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_           20140810
Time                3.35
INSTRUM           avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             MeOD
NS                  1024
DS                     2
SWH            31250.000 Hz
FIDRES         0.476837 Hz
AQ             1.0485760 sec
RG                   912
DW                16.000 usec
DE                18.00 usec
TE                298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         125.8131152 MHz
NUC1                 13C
P1                 10.00 usec
PLW1         20.18400002 W

======== CHANNEL f2 ========
SFO2         500.3020012 MHz
NUC2                 1H
CPDPRG[2          waltz16
PCPD2              80.00 usec
PLW2          7.99830008 W
PLW12         0.28119001 W
PLW13         0.17996000 W

F2 - Processing parameters
SI                  32768
SF           125.8003591 MHz
WDW                    EM
SSB       0
LB                  1.00 Hz
GB        0
PC                  1.40
```



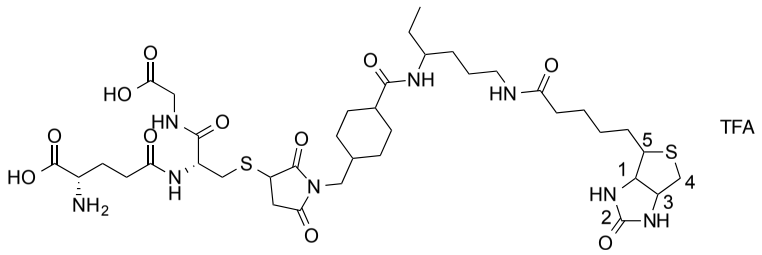
^{1}H NMR - *S*-((1-Biotinamido)-4-[4'-(succinimidomethyl)cyclohexanecarboxamido]hexane) L-glutathione 2,2,2-trifluoroacetate (96)

```
Current Data Parameters
NAME      sg68082102 -- SG22D mix
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140221
Time               10.45
INSTRUM            av700
PROBHD   5 mm CPTCI 1H-
PULPROG             zg30
TD                 65536
SOLVENT             MeOD
NS                    16
DS                     2
SWH            14097.744 Hz
FIDRES          0.215115 Hz
AQ             2.3243434 sec
RG                   203
DW                35.467 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         699.9943227 MHz
NUC1                   1H
P1                 10.34 usec
PLW1          5.48680019 W

F2 - Processing parameters
SI                 65536
SF           699.9900141 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```



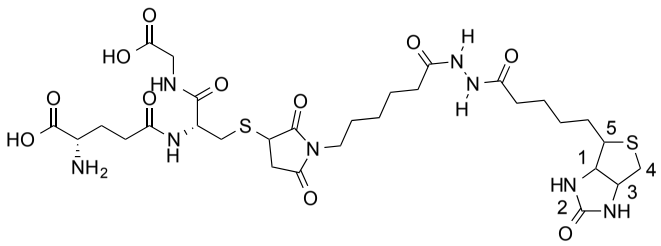
^{13}C NMR - *S*-((1-Biotinamido)-4-[4'-(succinimidomethyl)cyclohexanecarboxamido]hexane) L-glutathione 2,2,2-trifluoroacetate (96)

```
Current Data Parameters
NAME      sg68082102 -- SG22D mix
EXPNO                6
PROCNO               1

F2 - Acquisition Parameters
Date_           20140221
Time              16.25
INSTRUM           av700
PROBHD   5 mm CPTCI 1H-
PULPROG         zgpg30
TD               65536
SOLVENT            MeOD
NS                 1107
DS                    4
SWH           41666.668 Hz
FIDRES         0.635783 Hz
AQ            0.7864320 sec
RG                 2050
DW               12.000 usec
DE               18.00 usec
TE               298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                  1

======== CHANNEL f1 ========
SFO1        176.0302268 MHz
NUC1               13C
P1               15.00 usec
PLW1        56.11700058 W

======== CHANNEL f2 ========
SFO2        699.9928000 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2            65.00 usec
PLW2         5.49539995 W
PLW12        0.14177001 W
PLW13        0.05989600 W

F2 - Processing parameters
SI               65536
SF          176.0123772 MHz
WDW                EM
SSB      0
LB               1.00 Hz
GB       0
PC               1.40
```

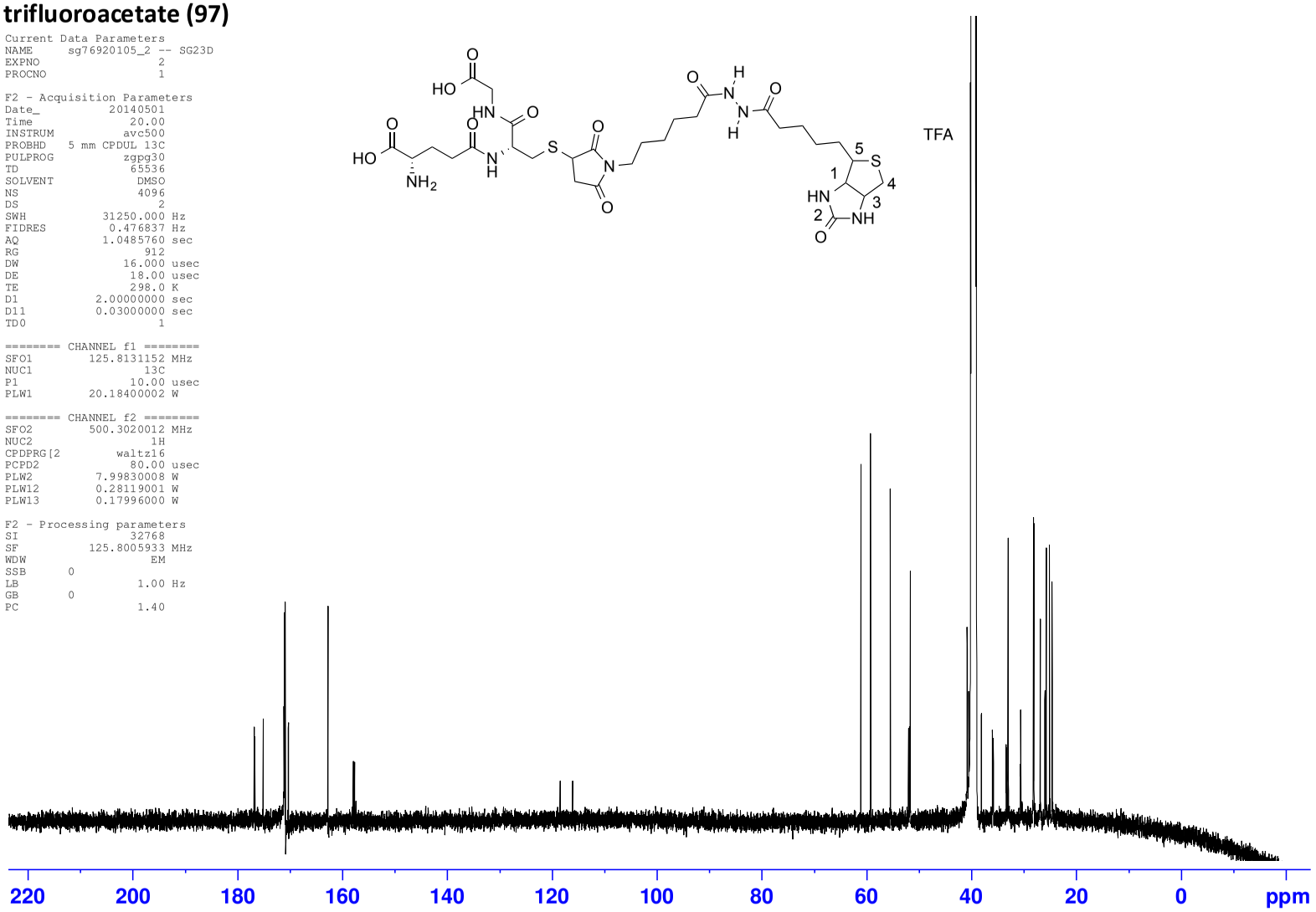


^{1}H NMR - (1*S*)-1-Carboxy-4-(((2*R*)-1-((carboxymethyl)amino)-3-((2,5-dioxo-1-(6-oxo-6-(2-(5-(2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanoyl)hydrazinyl)hexyl)pyrrolidin-3-yl)thio)-1-oxopropan-2-yl)amino)-4-oxobutan-1-aminium 2,2,2-trifluoroacetate (97)

```
Current Data Parameters
NAME      sg76920105 -- SG23D
EXPNO                1
PROCNO               1

F2 - Acquisition Parameters
Date_           20140501
Time              14.28
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG            zg30
TD                65536
SOLVENT            DMSO
NS                   16
DS                    4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                    4
DW               48.400 usec
DE               10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                1H
P1                15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                65536
SF           500.3000051 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```

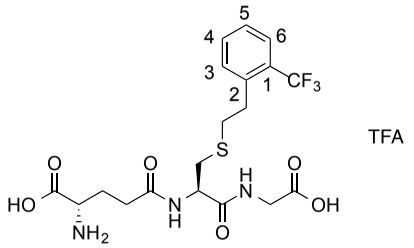


^{13}C NMR - (1*S*)-1-Carboxy-4-(((2*R*)-1-((carboxymethyl)amino)-3-((2,5-dioxo-1-(6-oxo-6-(2-(5-(2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanoyl)hydrazinyl)hexyl)pyrrolidin-3-yl)thio)-1-oxopropan-2-yl)amino)-4-oxobutan-1-aminium 2,2,2-trifluoroacetate (97)

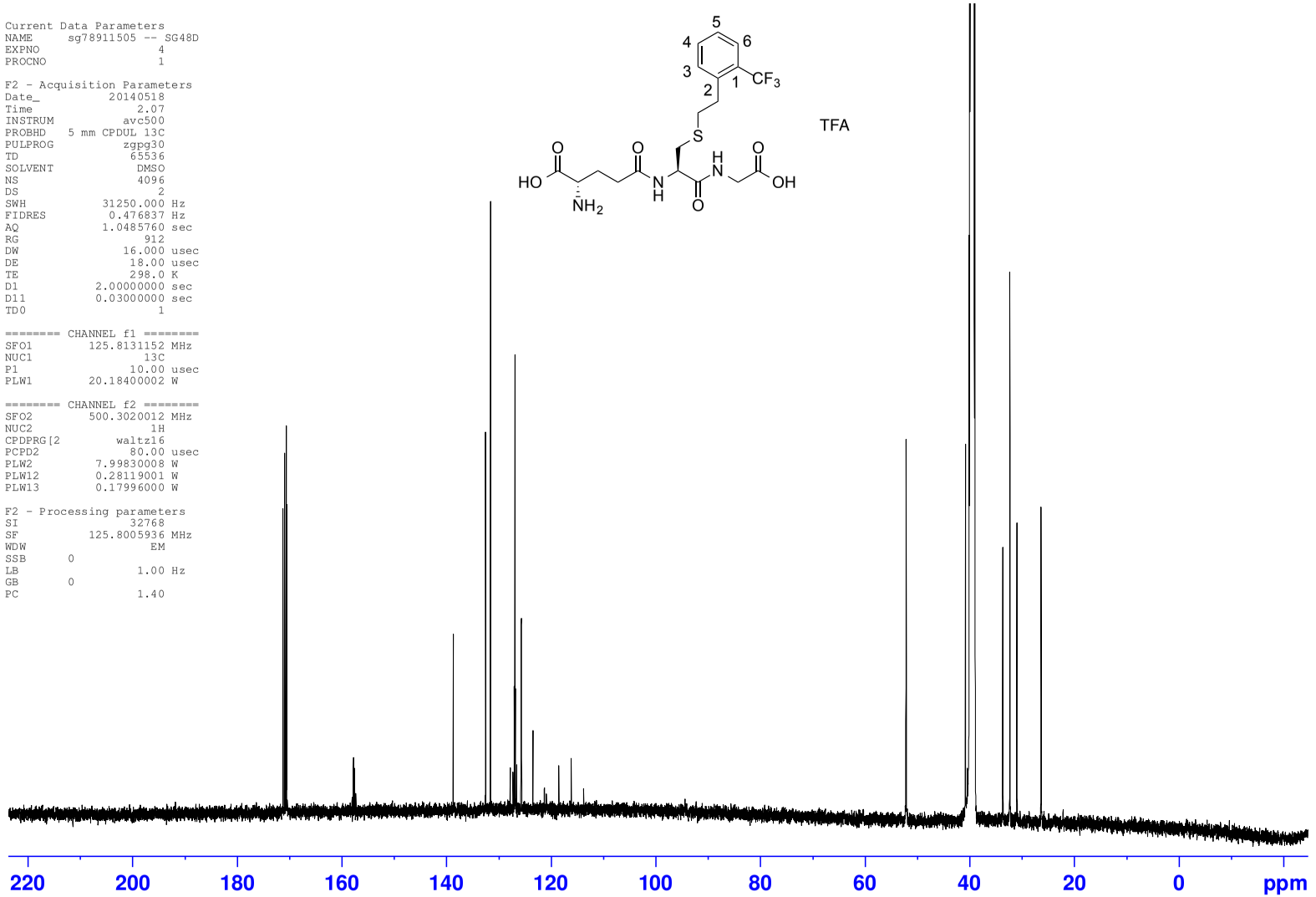


^{1}H NMR - *S*-2-(2-(Trifluoromethyl)phenyl)ethyl L-glutathione 2,2,2-trifluoroacetate (98)

```
Current Data Parameters
NAME      sg78911505 -- SG48D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140517
Time              21.45
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG            zg30
TD                65536
SOLVENT            DMSO
NS                   16
DS                    4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                    4
DW               48.400 usec
DE               10.00 usec
TE               298.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                1H
P1                15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                65536
SF           500.3000052 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```

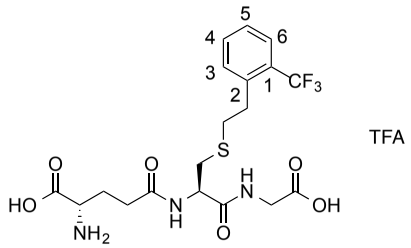


^{13}C NMR - *S*-2-(2-(Trifluoromethyl)phenyl)ethyl L-glutathione 2,2,2-trifluoroacetate (98)



^{19}F NMR - *S*-2-(2-(Trifluoromethyl)phenyl)ethyl L-glutathione 2,2,2-trifluoroacetate (98)

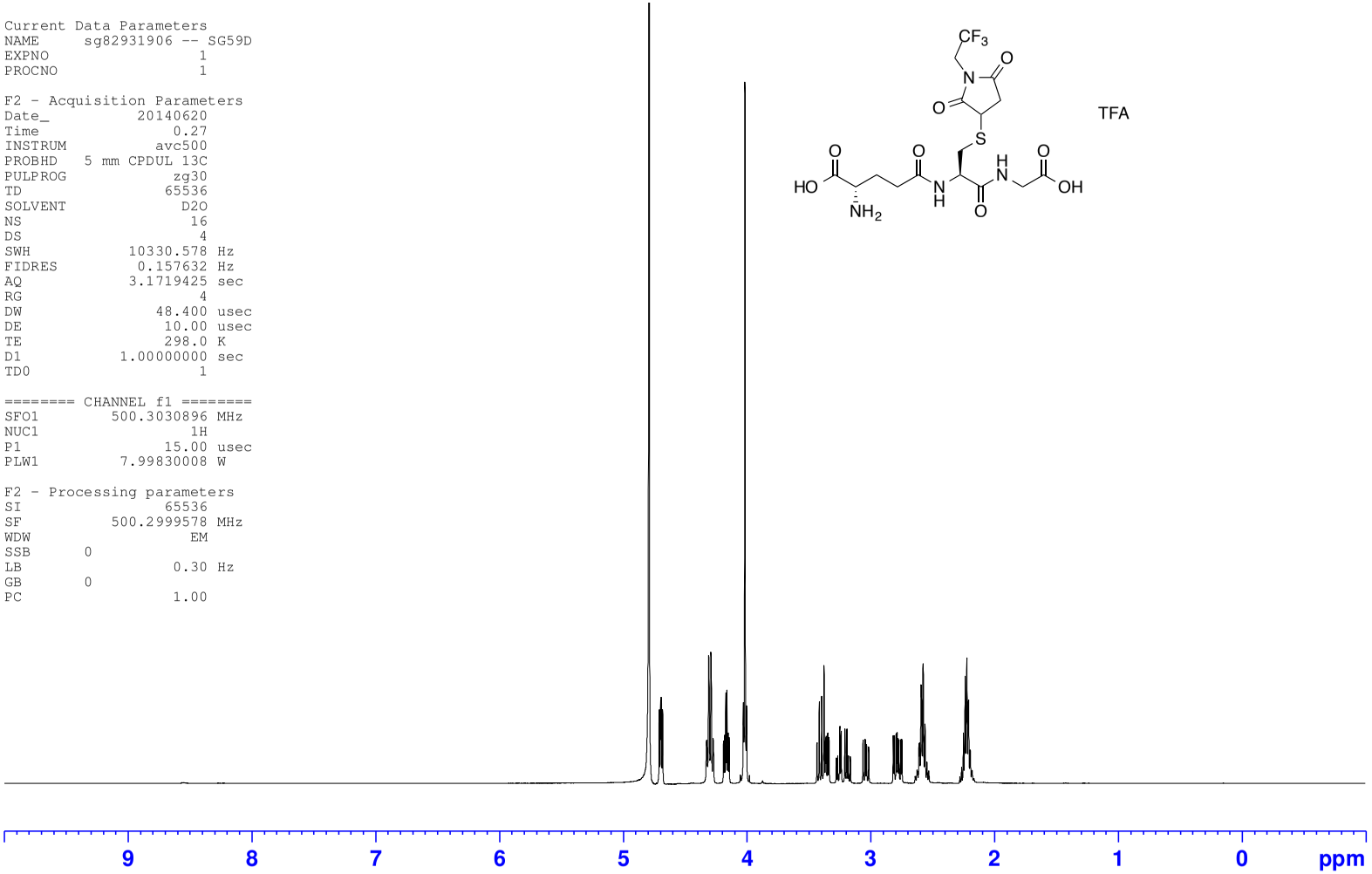
```
Current Data Parameters
NAME      May15-2014-55 -- SG48D
EXPNO                    2
PROCNO                   1

F2 - Acquisition Parameters
Date_           20140515
Time                9.55
INSTRUM          avf400
PROBHD   5 mm PABBO BB/
PULPROG          zgflqn
TD               131072
SOLVENT             DMSO
NS                    16
DS                     4
SWH            75000.000 Hz
FIDRES          0.572205 Hz
AQ             0.8738133 sec
RG                205.43
DW                 6.667 usec
DE                 6.50 usec
TE                296.2 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         376.5547873 MHz
NUC1                19F
P1                13.50 usec
PLW1        19.00000000 W

F2 - Processing parameters
SI                 65536
SF           376.6112790 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```



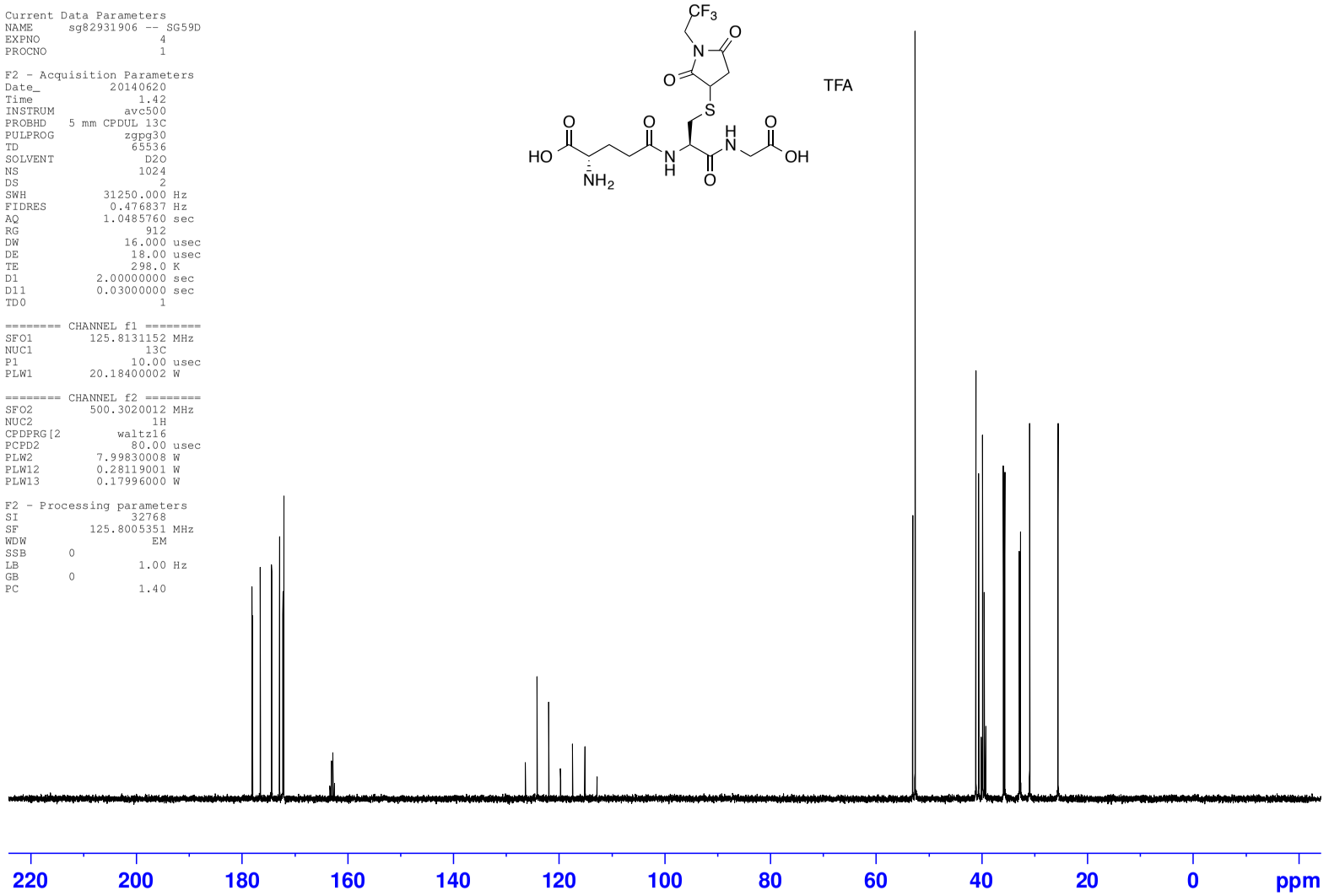
TFA

^{1}H NMR - *S*-(*N*-2,2,2-Trifluoroethylsuccinimido) L-glutathione 2,2,2-trifluoroacetate (99)

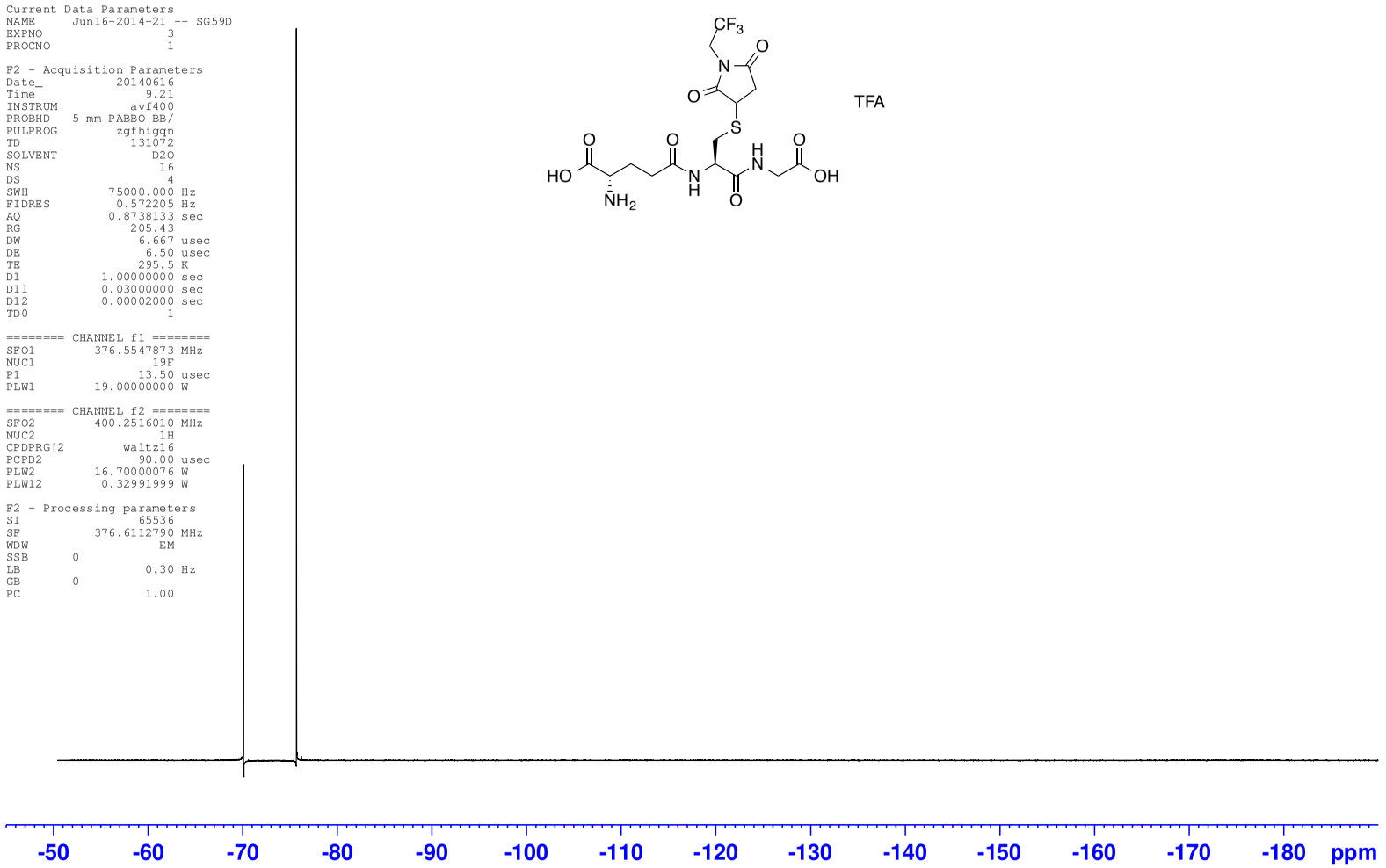
```
Current Data Parameters
NAME      sg82931906 -- SG59D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140620
Time                0.27
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                  15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999578 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```



^{13}C NMR - *S*-(*N*-2,2,2-Trifluoroethylsuccinimido) L-glutathione 2,2,2-trifluoroacetate (99)

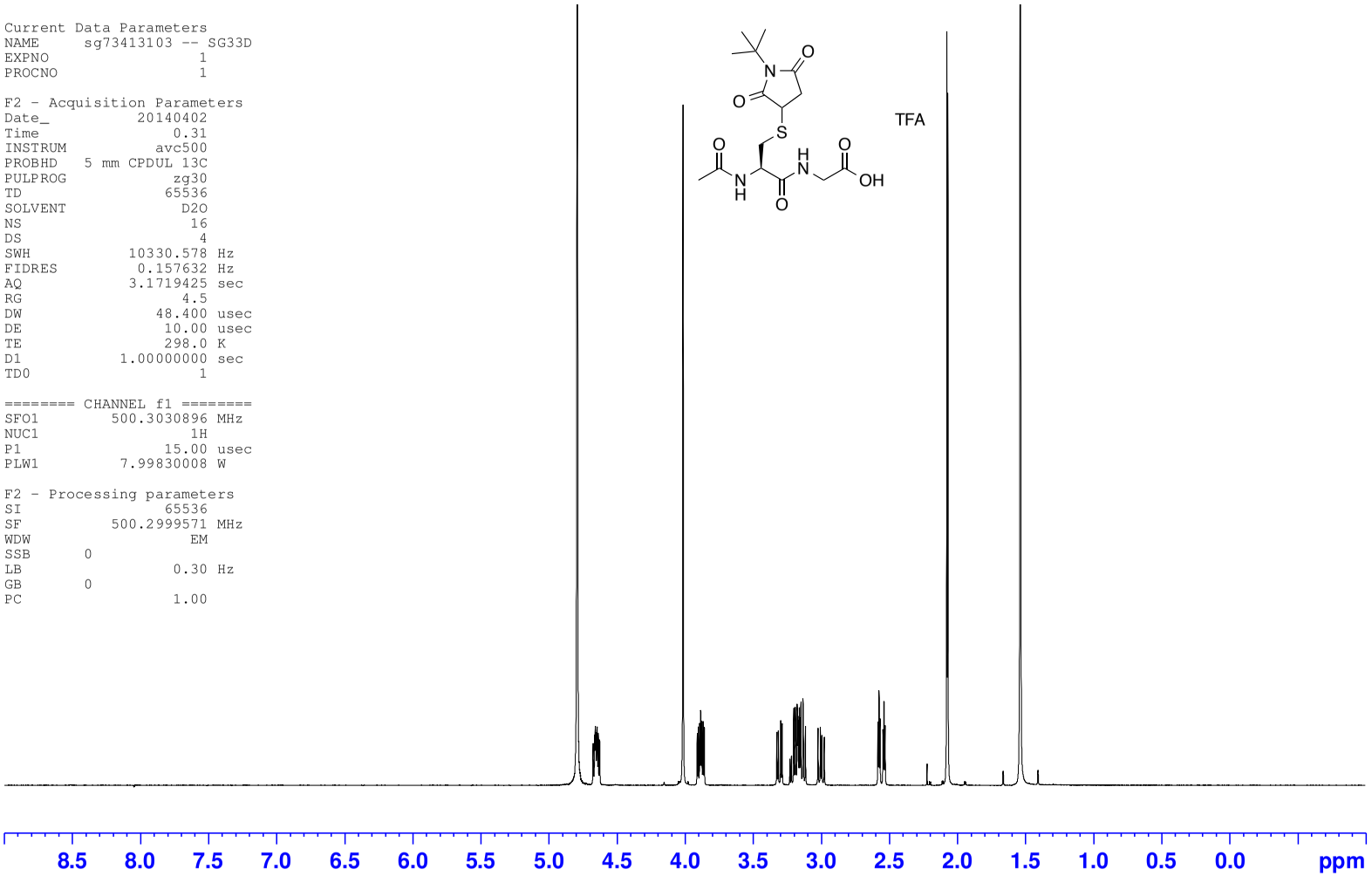


^{19}F NMR - *S*-(*N*-2,2,2-Trifluoroethylsuccinimido) L-glutathione 2,2,2-trifluoroacetate (99)

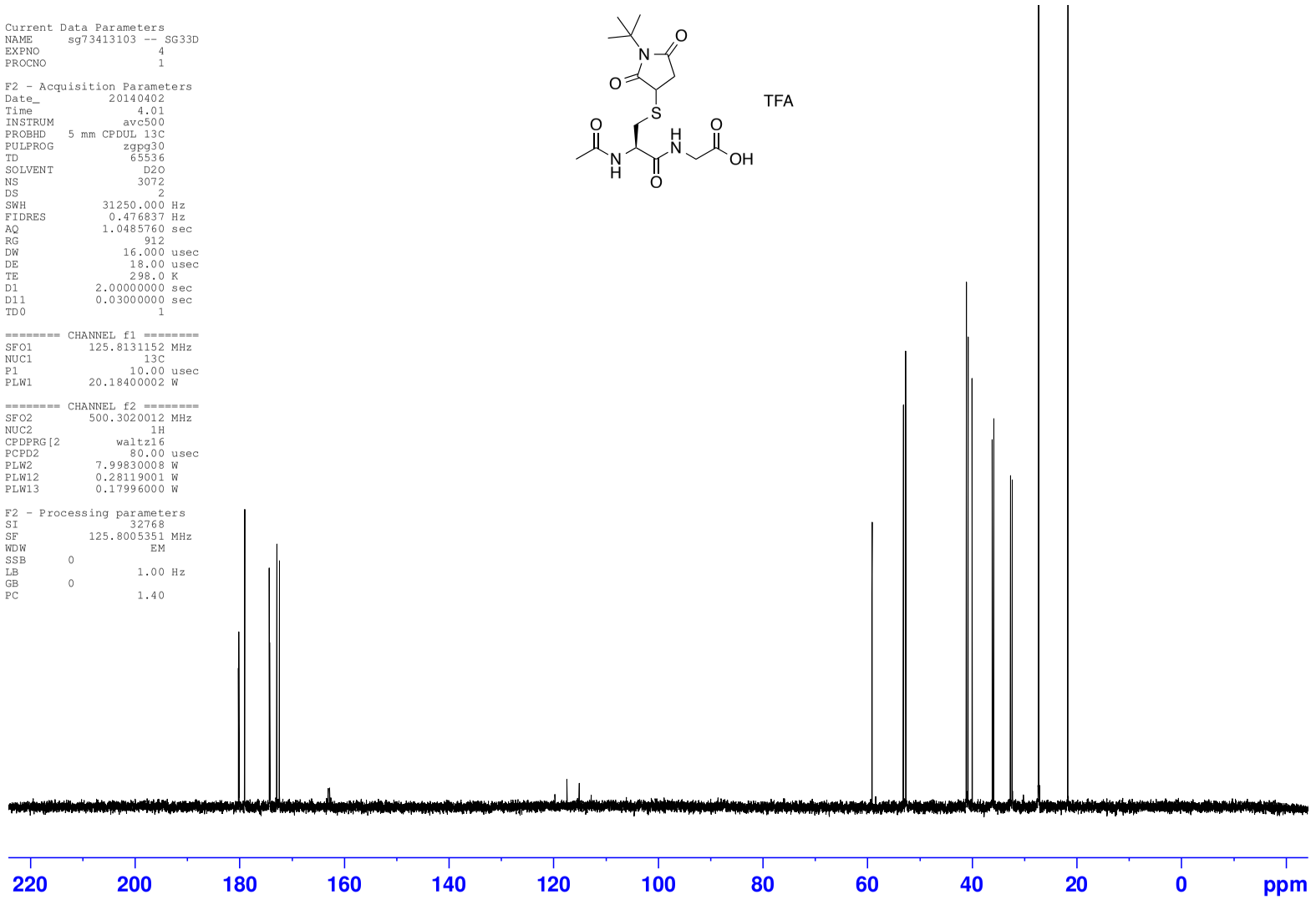


^{1}H NMR - *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (101)

```
Current Data Parameters
NAME      sg73413103 -- SG33D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140402
Time                0.31
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                   4.5
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                  15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999571 MHz
WDW                   EM
SSB      0
LB                  0.30 Hz
GB       0
PC                  1.00
```

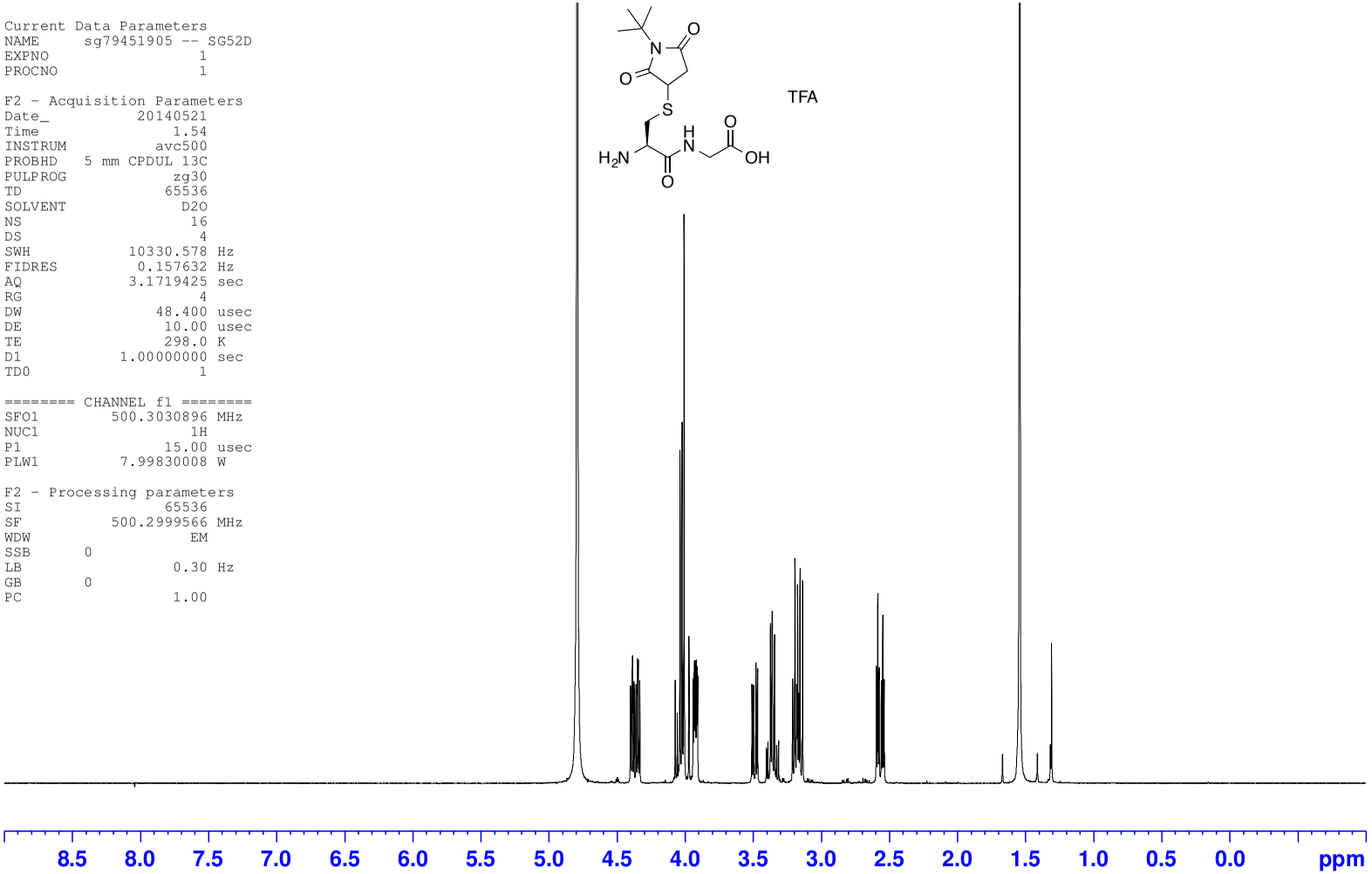


^{13}C NMR - *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (101)

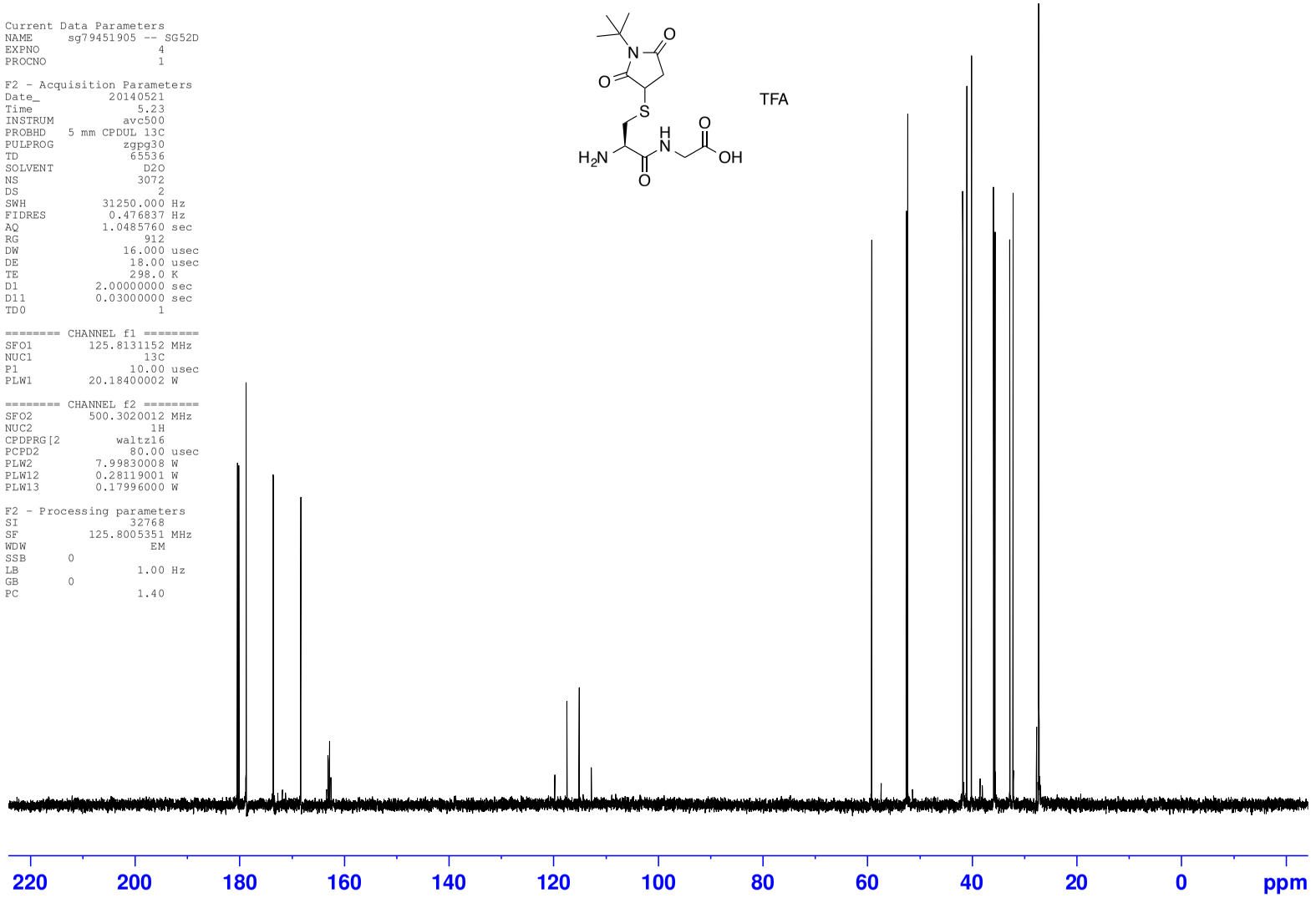


^{1}H NMR - *S*-(*N*-*tert*-Butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (102)

```
Current Data Parameters
NAME      sg79451905 -- SG52D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140521
Time                1.54
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                  1H
P1                 15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999566 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```

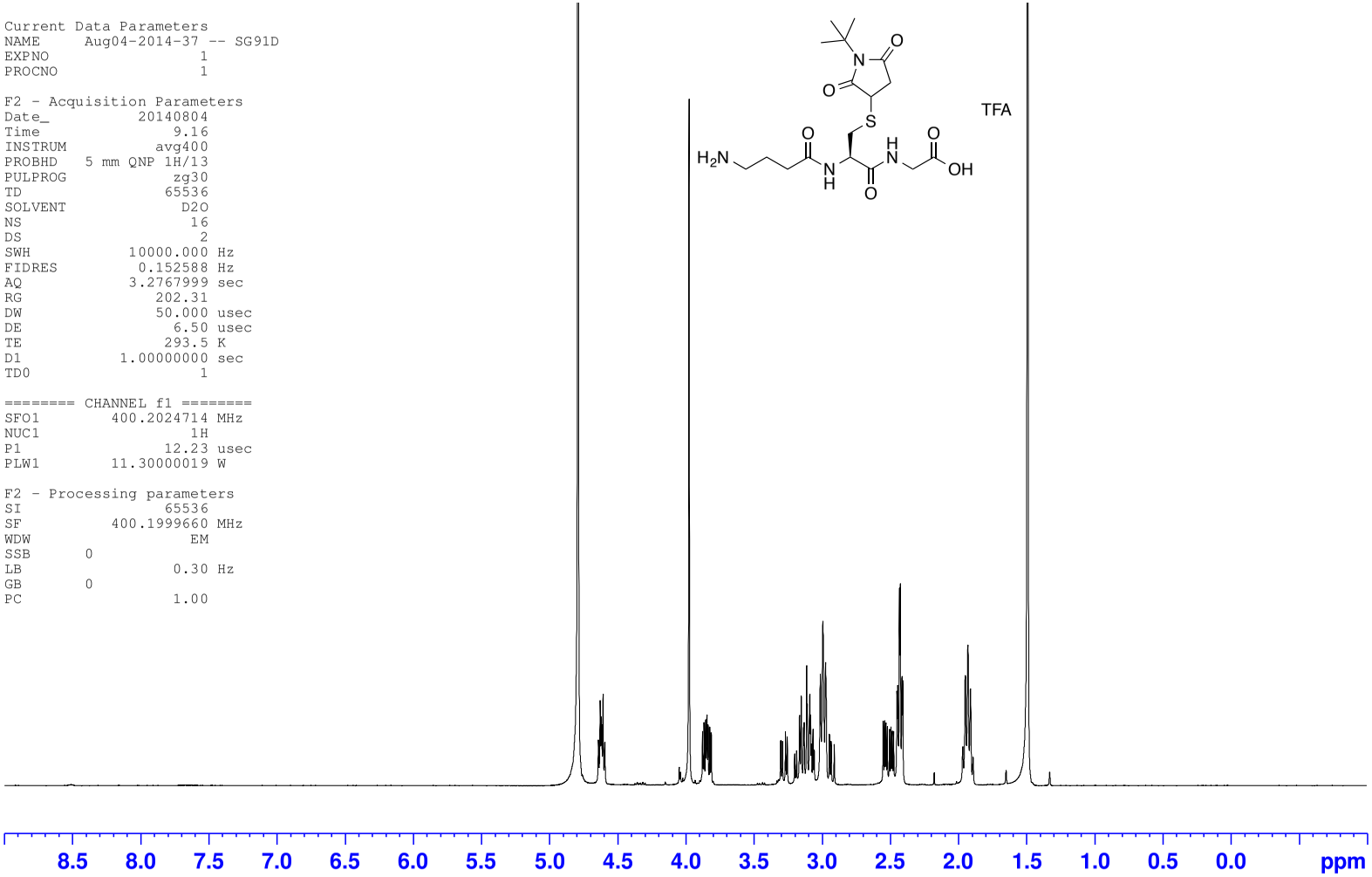


^{13}C NMR - *S*-(*N*-*tert*-Butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (102)

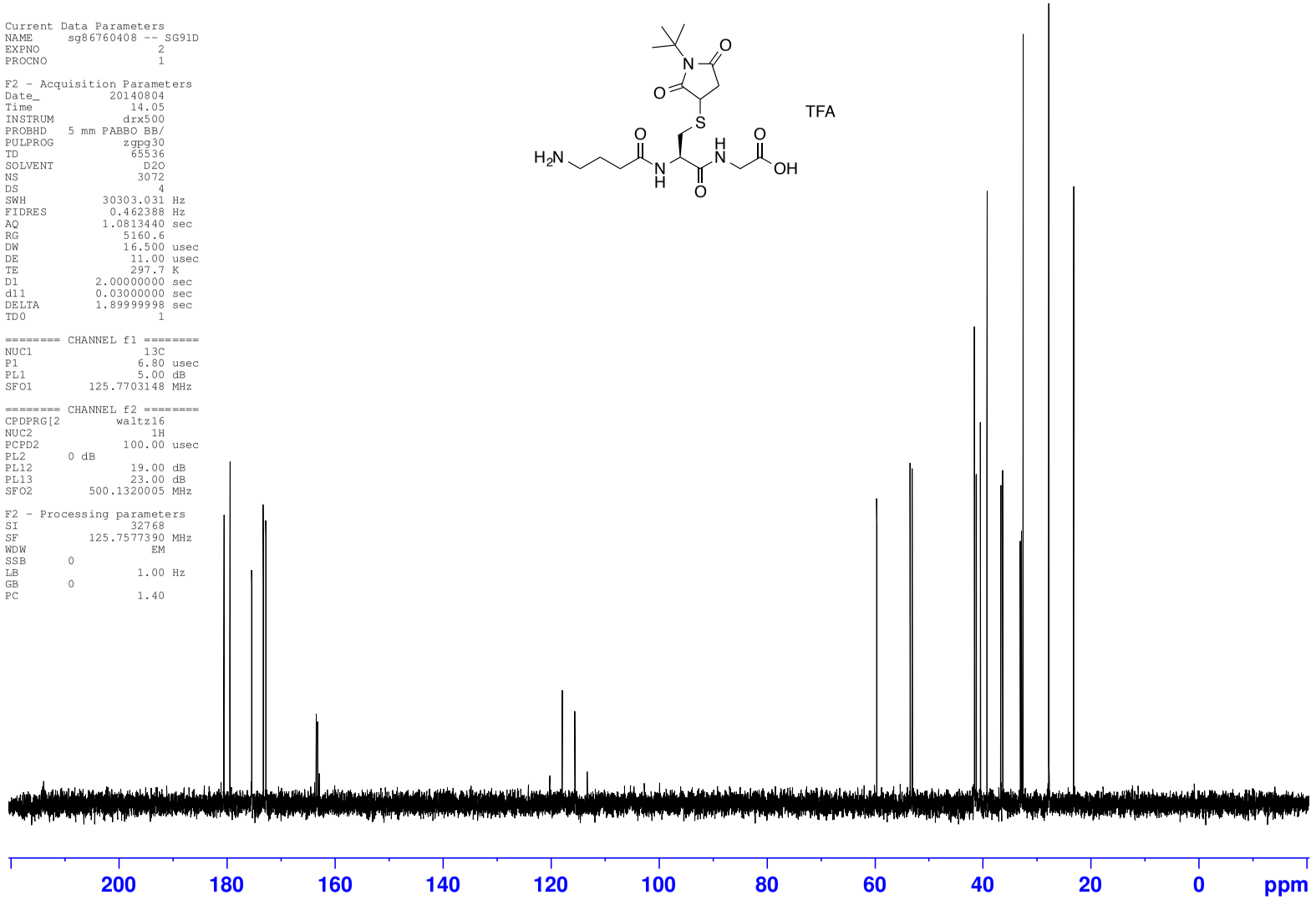


^{1}H NMR ɣ-Aminobutyric acid-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (109)

```
Current Data Parameters
NAME      Aug04-2014-37 -- SG91D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140804
Time               9.16
INSTRUM          avg400
PROBHD   5 mm QNP 1H/13
PULPROG            zg30
TD                65536
SOLVENT             D2O
NS                   16
DS                    2
SWH           10000.000 Hz
FIDRES         0.152588 Hz
AQ             3.2767999 sec
RG               202.31
DW               50.000 usec
DE                6.50 usec
TE                293.5 K
D1            1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         400.2024714 MHz
NUC1                 1H
P1                12.23 usec
PLW1         11.30000019 W

F2 - Processing parameters
SI                65536
SF           400.1999660 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```

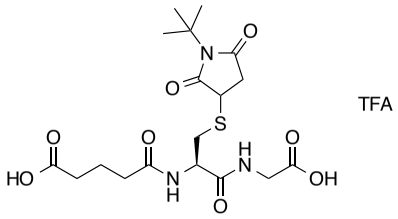


^{13}C NMR γ-Aminobutyric acid-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (109)

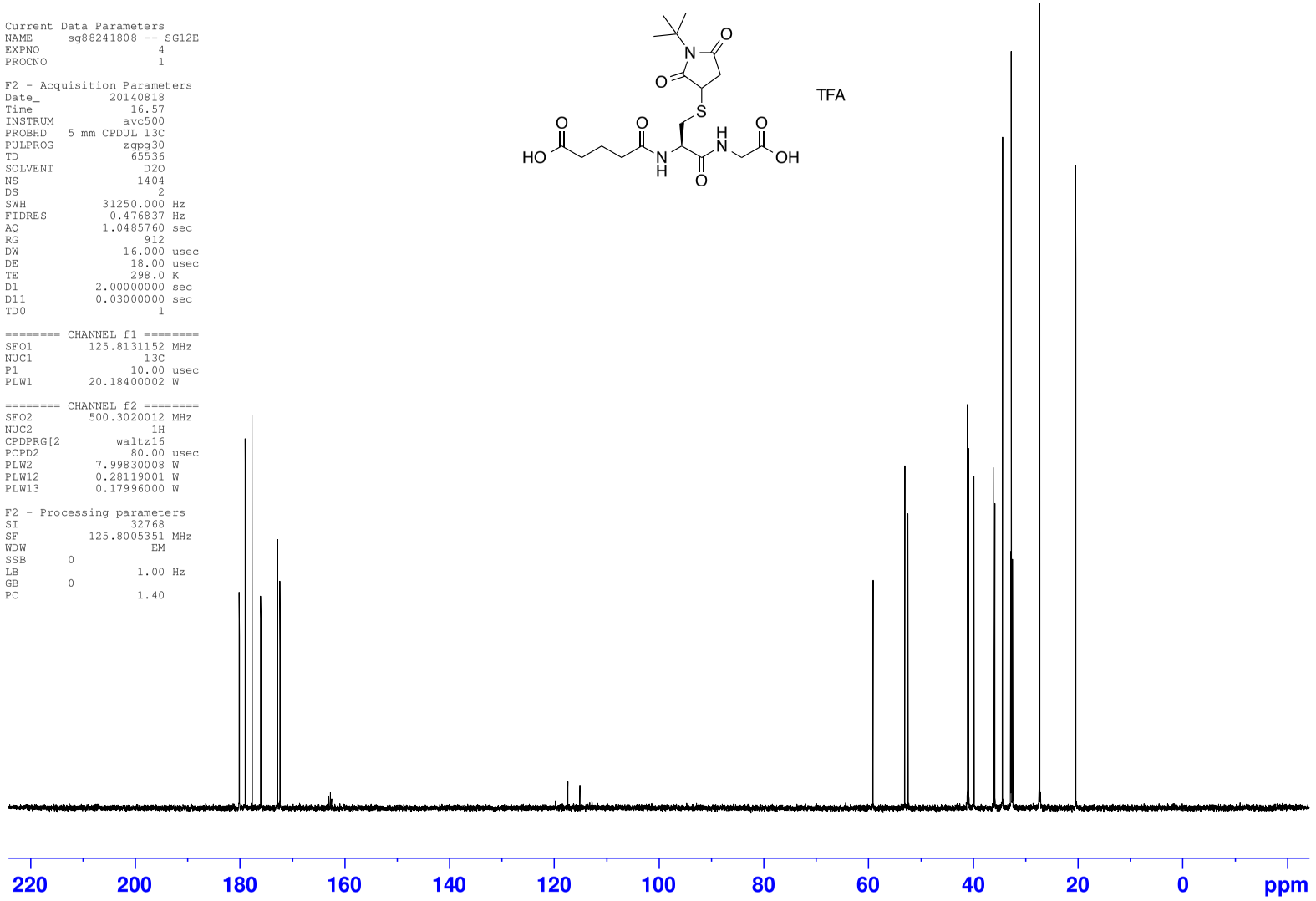


^{1}H NMR - Glutaric acid-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (112)

```
Current Data Parameters
NAME      sg88241808 -- SG12E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140818
Time              14.18
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                 15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999581 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```

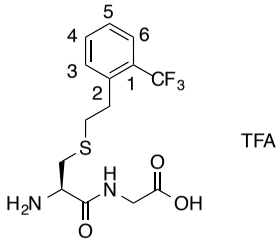


^{13}C NMR - Glutaric acid-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine 2,2,2-trifluoroacetate (112)



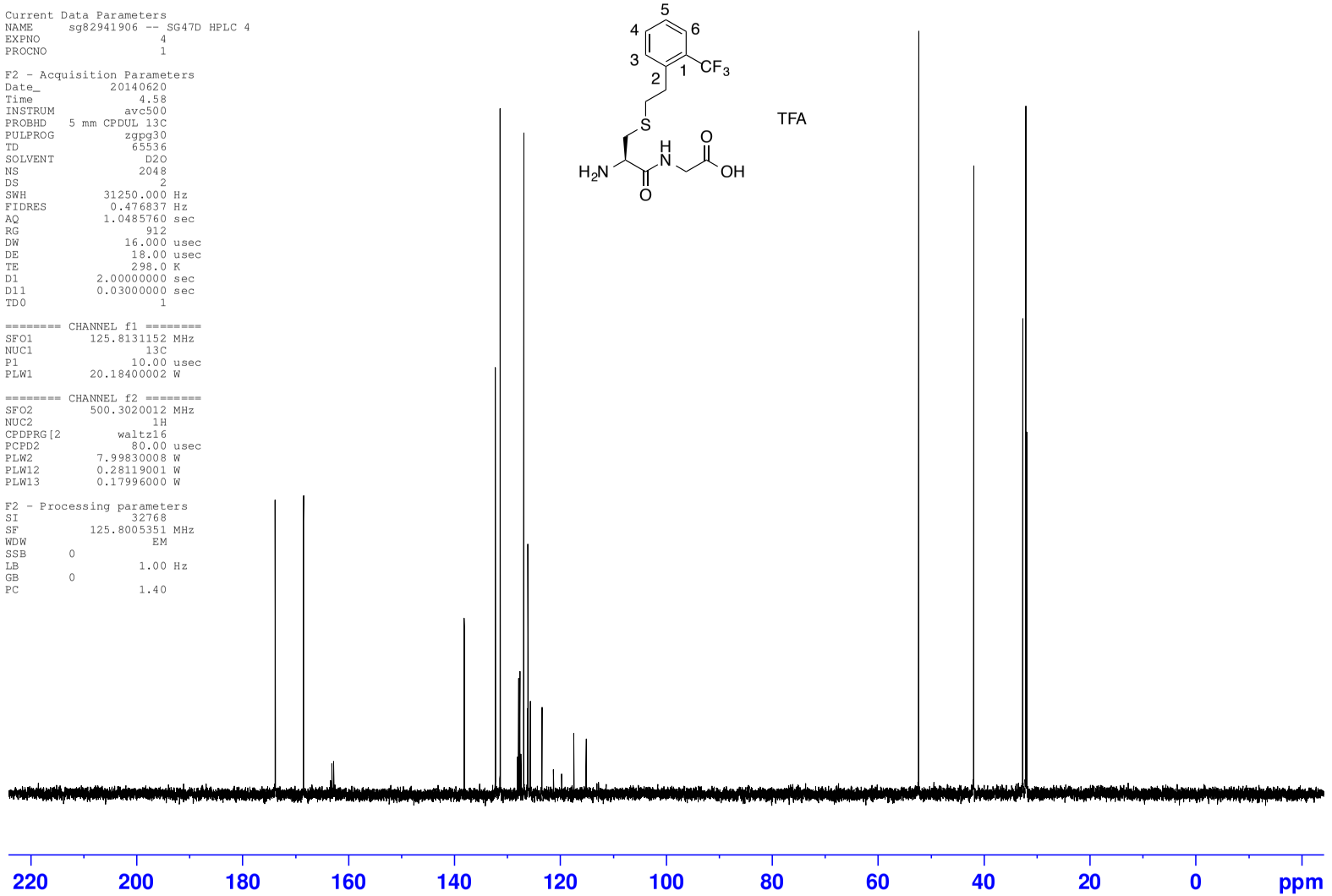
^{1}H NMR - *S*-2-(2-(Trifluoromethyl)phenyl)ethyl L-cys-gly 2,2,2-trifluoroacetate (113)

```
Current Data Parameters
NAME      sg82941906_2 -- SG47D HPLC 4
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_          20140620
Time               9.52
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES         0.157632 Hz
AQ             3.1719425 sec
RG                   4.5
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                 1H
P1                15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999569 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```



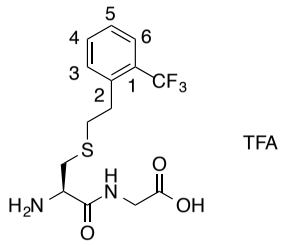
TFA

^{13}C NMR - *S*-2-(2-(Trifluoromethyl)phenyl)ethyl L-cys-gly 2,2,2-trifluoroacetate (113)



^{19}F NMR - *S*-2-(2-(Trifluoromethyl)phenyl)ethyl L-cys-gly 2,2,2-trifluoroacetate (113)

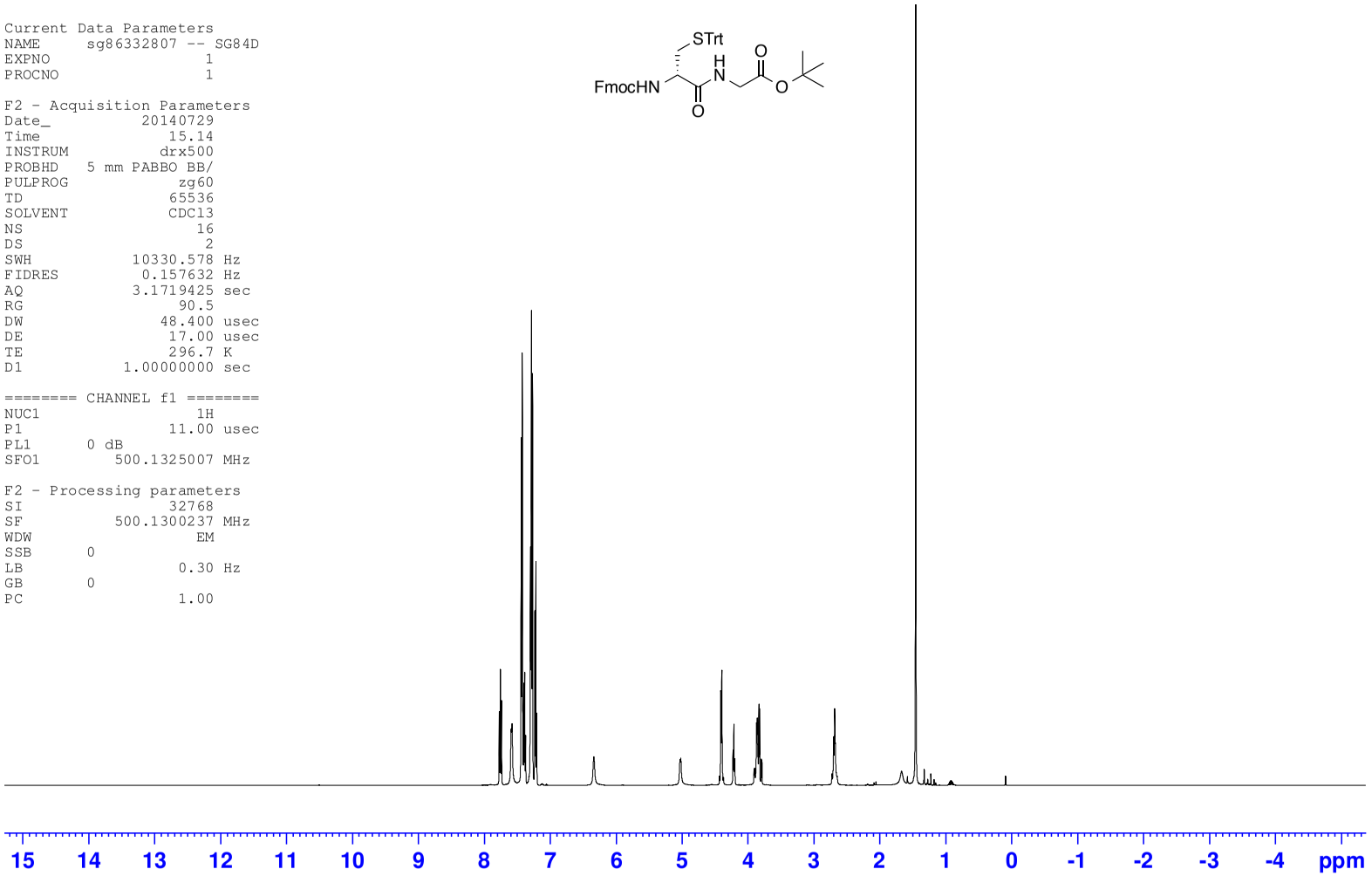
```
Current Data Parameters
NAME      Jun16-2014-20 -- SG58D
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140616
Time               9.14
INSTRUM          avf400
PROBHD   5 mm PABBO BB/
PULPROG          zgflqn
TD               131072
SOLVENT             D2O
NS                   16
DS                    4
SWH            75000.000 Hz
FIDRES          0.572205 Hz
AQ             0.8738133 sec
RG                205.43
DW                 6.667 usec
DE                 6.50 usec
TE                295.4 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         376.5547873 MHz
NUC1                19F
P1                13.50 usec
PLW1        19.00000000 W

F2 - Processing parameters
SI                 65536
SF           376.6112790 MHz
WDW                  EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```



TFA

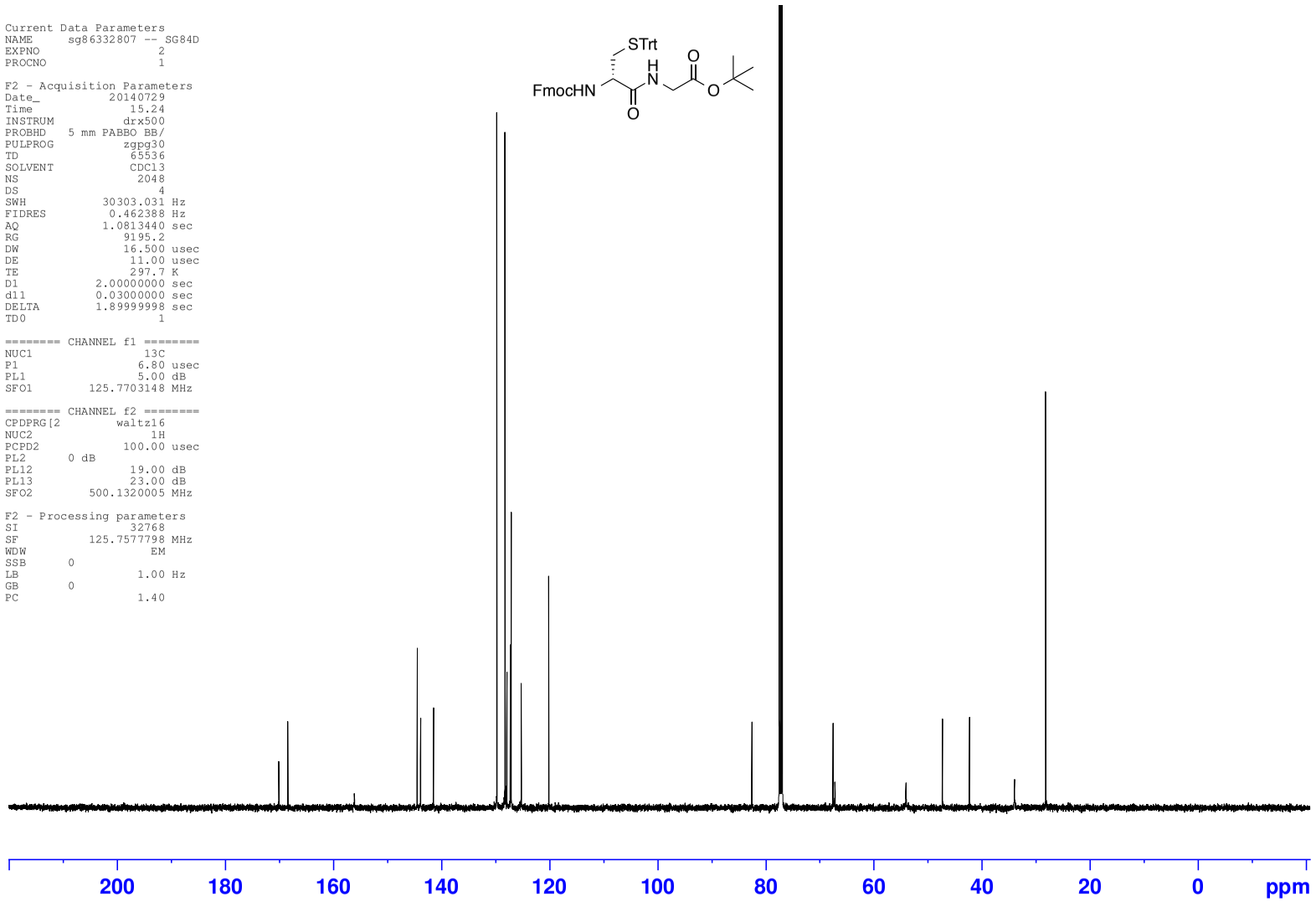
^{1}H NMR - *N*-Fmoc-*S*-trityl-D-cysteinylglycine *tert*-butyl ester (115)

```
Current Data Parameters
NAME      sg86332807 -- SG84D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140729
Time              15.14
INSTRUM          drx500
PROBHD   5 mm PABBO BB/
PULPROG            zg60
TD                65536
SOLVENT           CDCl3
NS                   16
DS                    2
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                 90.5
DW               48.400 usec
DE               17.00 usec
TE               296.7 K
D1           1.00000000 sec

======== CHANNEL f1 ========
NUC1                1H
P1                11.00 usec
PL1       0 dB
SFO1        500.1325007 MHz

F2 - Processing parameters
SI                 32768
SF          500.1300237 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```

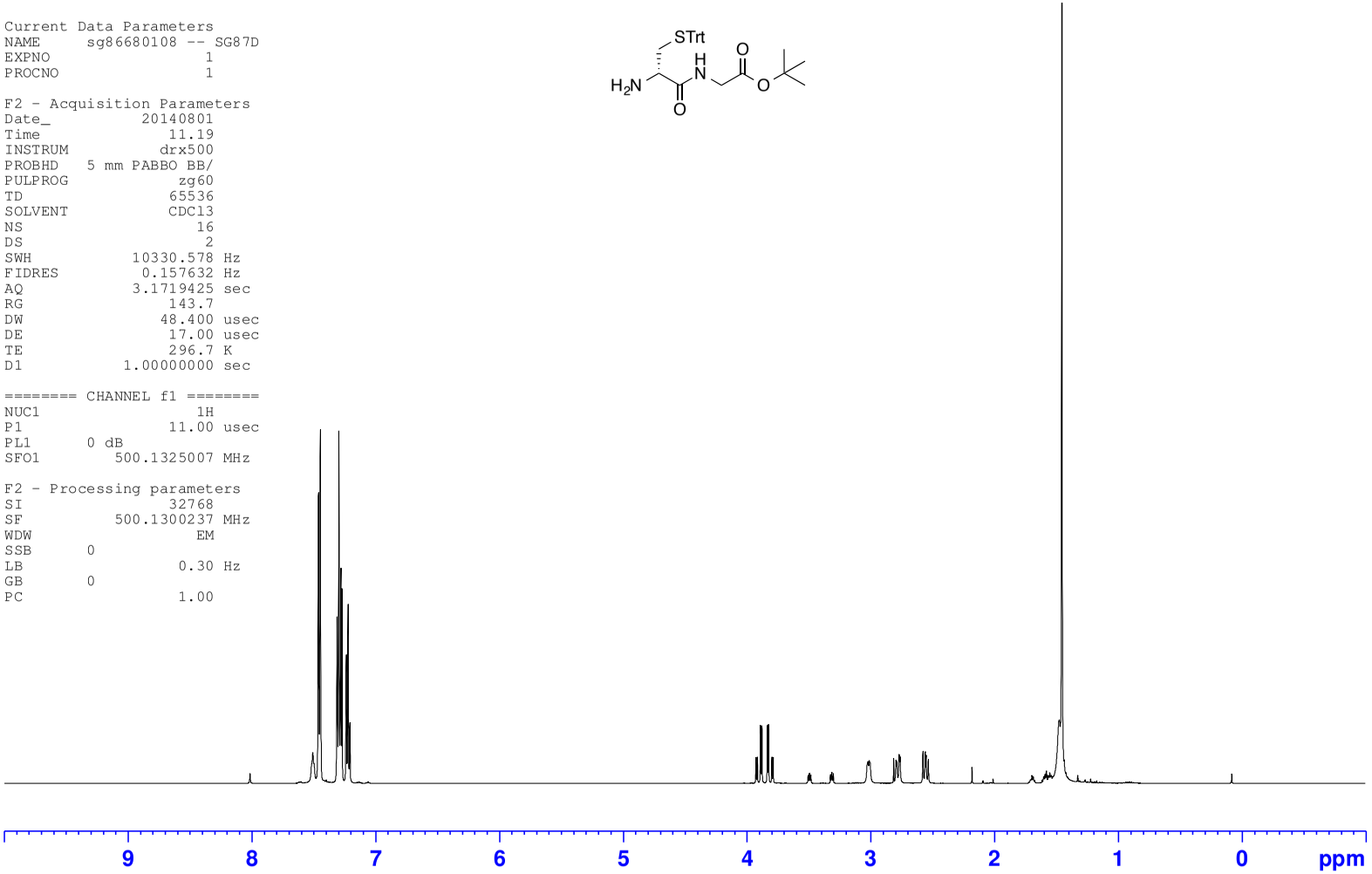


^{13}C NMR - *N*-Fmoc-*S*-trityl-D-cysteinylglycine *tert*-butyl ester (115)

```
Current Data Parameters
NAME      sg86332807 -- SG84D
EXPNO                2
PROCNO               1

F2 - Acquisition Parameters
Date_          20140729
Time             15.24
INSTRUM          drx500
PROBHD   5 mm PABBO BB/
PULPROG          zgpg30
TD                65536
SOLVENT           CDCl3
NS                 2048
DS                    4
SWH           30303.031 Hz
FIDRES         0.462388 Hz
AQ            1.0813440 sec
RG               9195.2
DW               16.500 usec
DE               11.00 usec
TE               297.7 K
D1           2.00000000 sec
d11          0.03000000 sec
DELTA        1.89999998 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                 13C
P1                  6.80 usec
PL1                 5.00 dB
SFO1       125.7703148 MHz

======== CHANNEL f2 ========
CPDPRG[2          waltz16
NUC2                 1H
PCPD2             100.00 usec
PL2      0 dB
PL12                19.00 dB
PL13                23.00 dB
SFO2       500.1320005 MHz

F2 - Processing parameters
SI                  32768
SF         125.7577798 MHz
WDW                    EM
SSB      0
LB                  1.00 Hz
GB       0
PC                  1.40
```

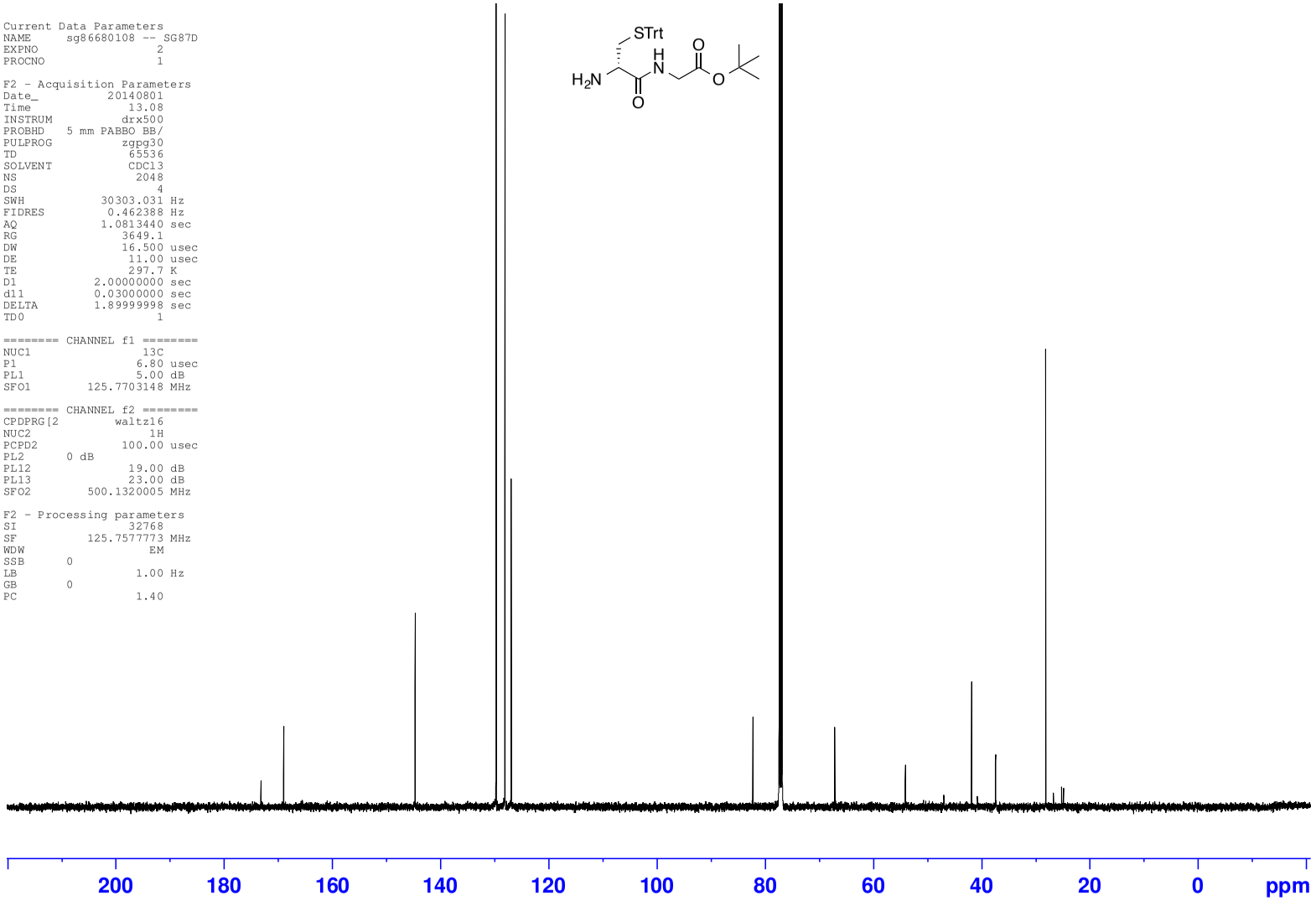


^{1}H NMR - *S*-Trityl-D-cysteinylglycine *tert*-butyl ester (116)



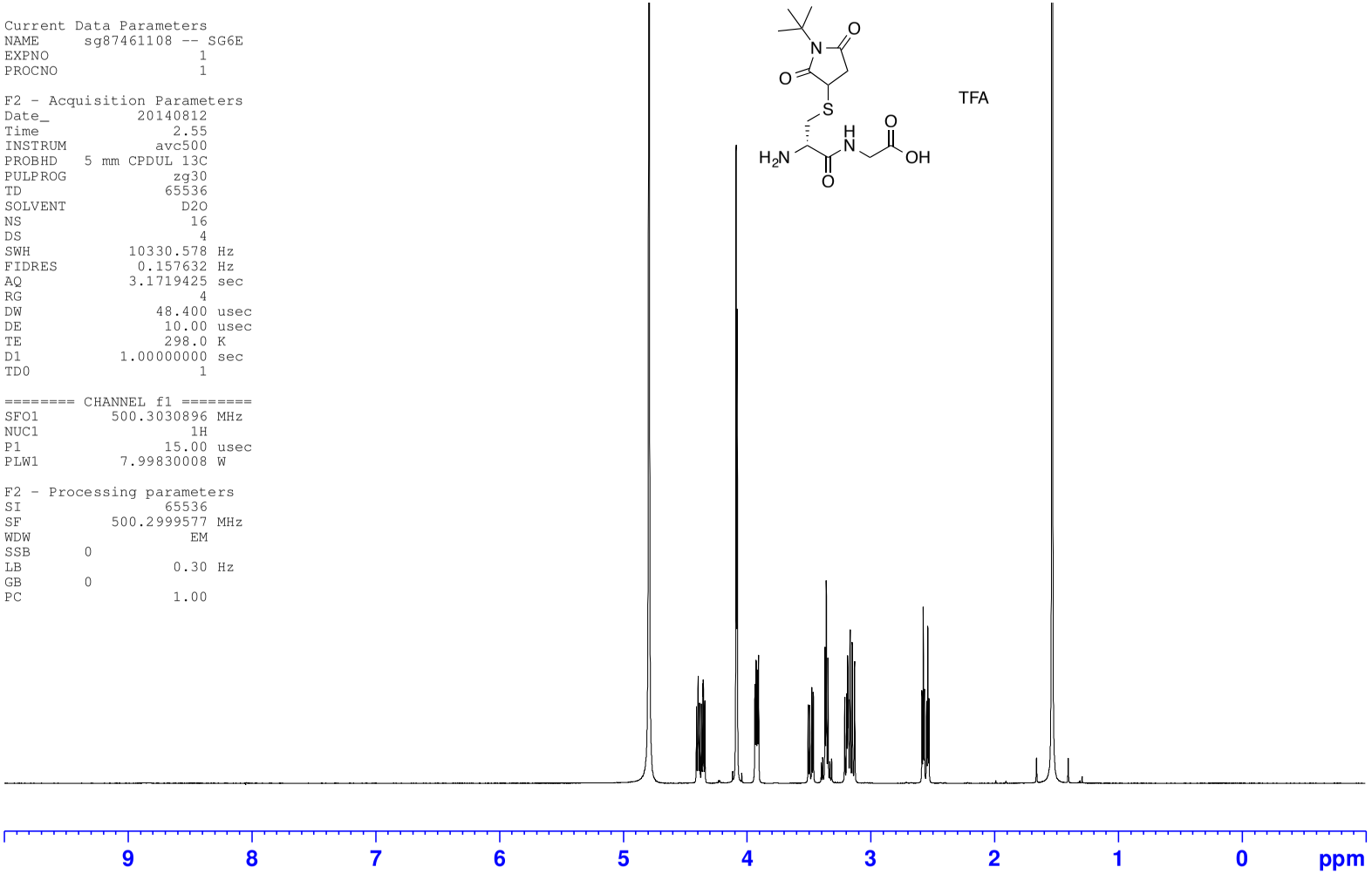
^{13}C NMR - *S*-Trityl-D-cysteinylglycine *tert*-butyl ester (116)

```
Current Data Parameters
NAME      sg86680108 -- SG87D
EXPNO                2
PROCNO               1

F2 - Acquisition Parameters
Date_          20140801
Time             13.08
INSTRUM          drx500
PROBHD   5 mm PABBO BB/
PULPROG          zgpg30
TD                65536
SOLVENT           CDCl3
NS                 2048
DS                    4
SWH           30303.031 Hz
FIDRES         0.462388 Hz
AQ             1.0813440 sec
RG                3649.1
DW                16.500 usec
DE                11.00 usec
TE                297.7 K
D1           2.00000000 sec
d11          0.03000000 sec
DELTA        1.89999998 sec
TD0                  1

======== CHANNEL f1 ========
NUC1                13C
P1                 6.80 usec
PL1                5.00 dB
SFO1       125.7703148 MHz

======== CHANNEL f2 ========
CPDPRG[2         waltz16
NUC2                1H
PCPD2           100.00 usec
PL2      0 dB
PL12               19.00 dB
PL13               23.00 dB
SFO2       500.1320005 MHz

F2 - Processing parameters
SI                 32768
SF         125.7577773 MHz
WDW                   EM
SSB      0
LB                 1.00 Hz
GB       0
PC                 1.40
```

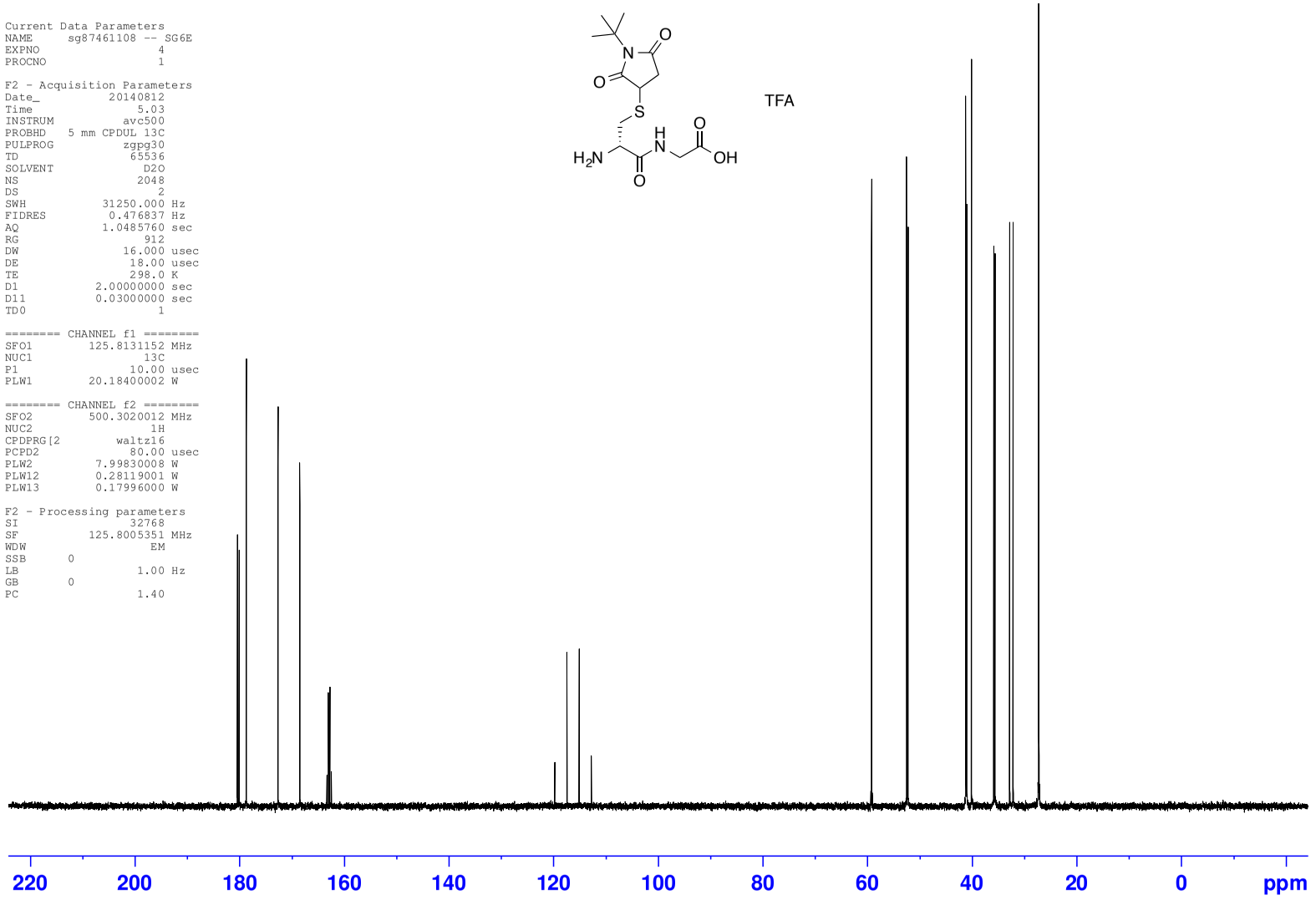


S-(*N*-*tert*-Butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate (118)

```
Current Data Parameters
NAME      sg87461108 -- SG6E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140812
Time                2.55
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                 15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999577 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```

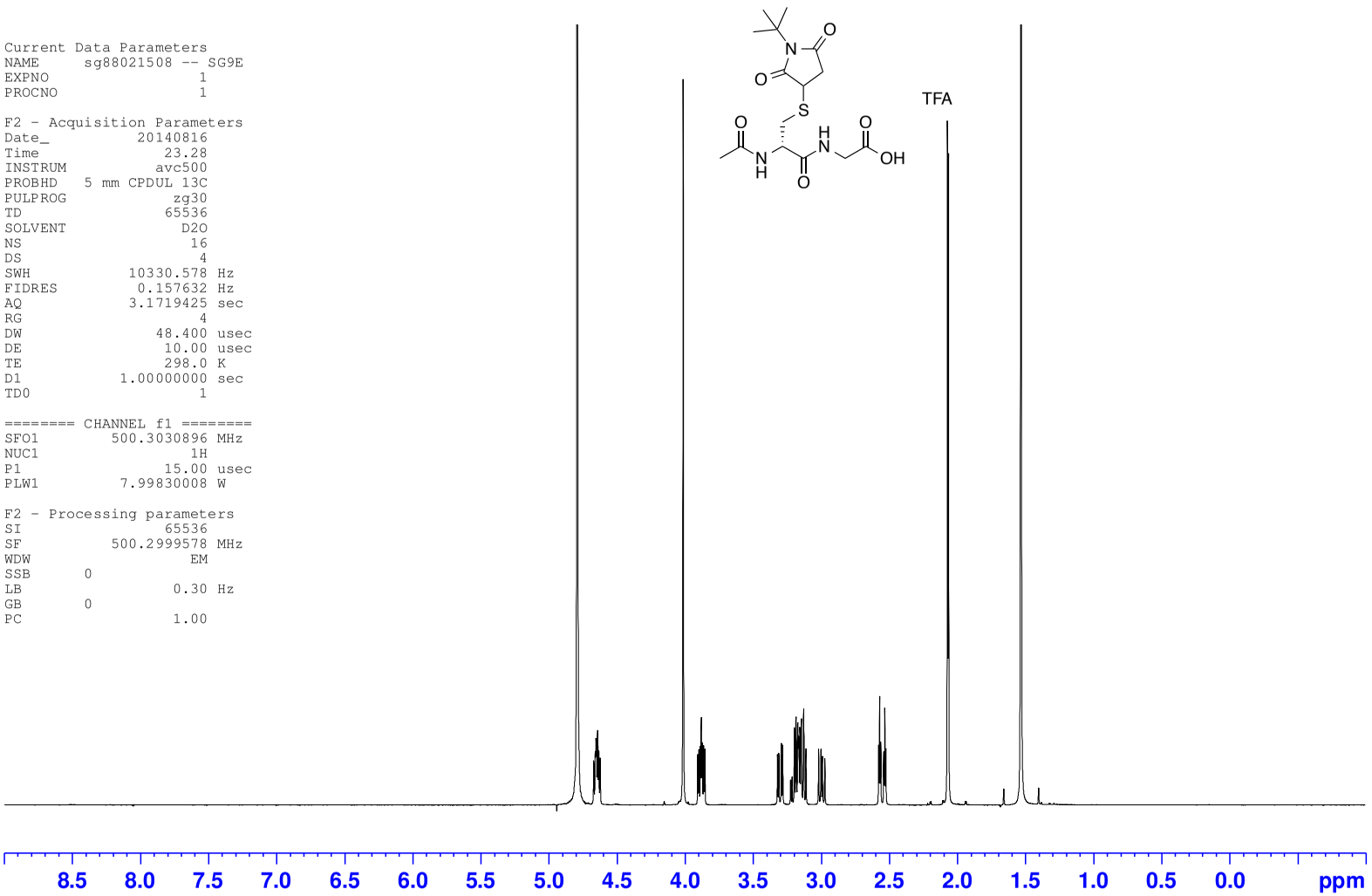


S-(*N*-*tert*-Butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate (118)

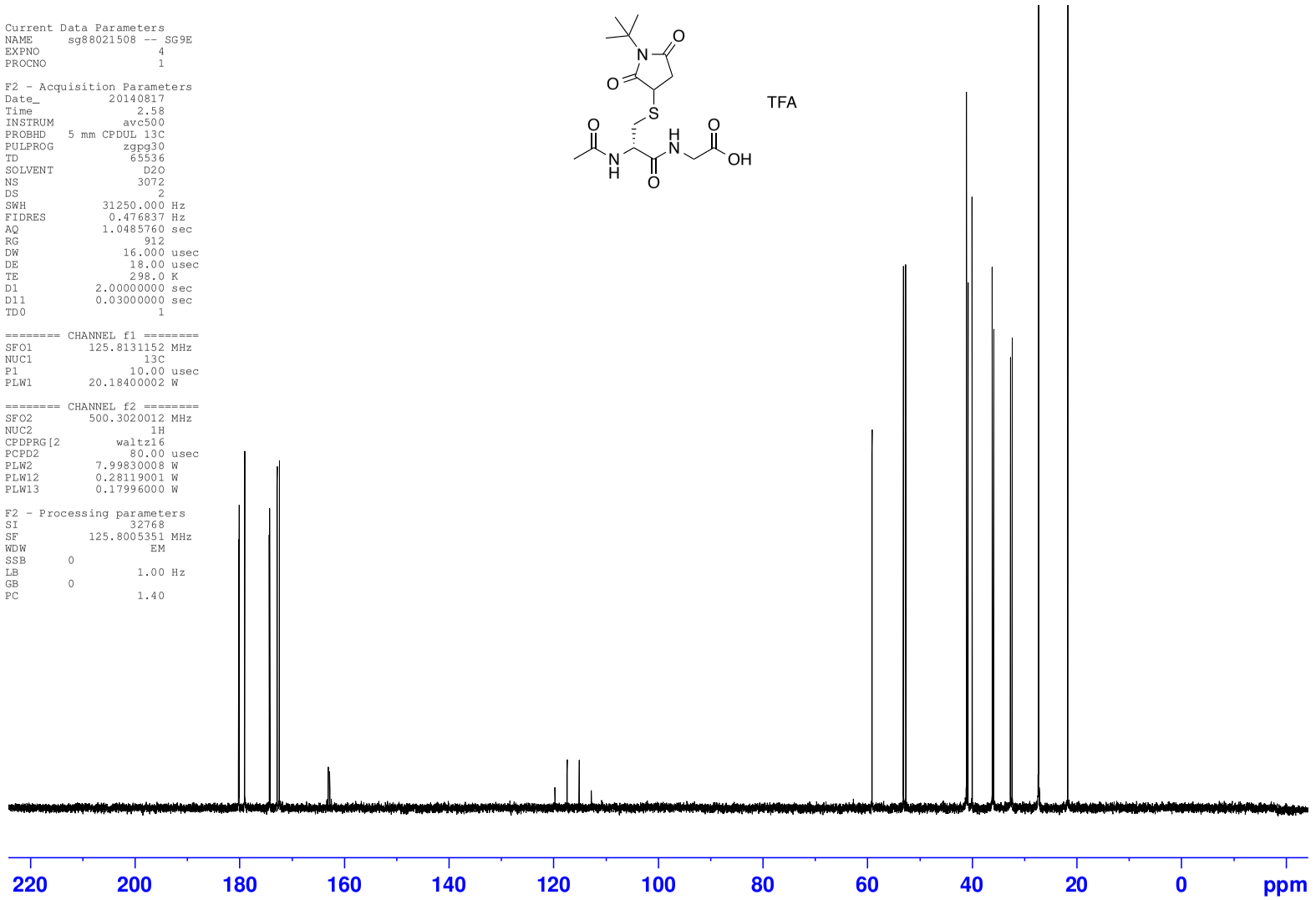


^{1}H NMR - *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate (119)

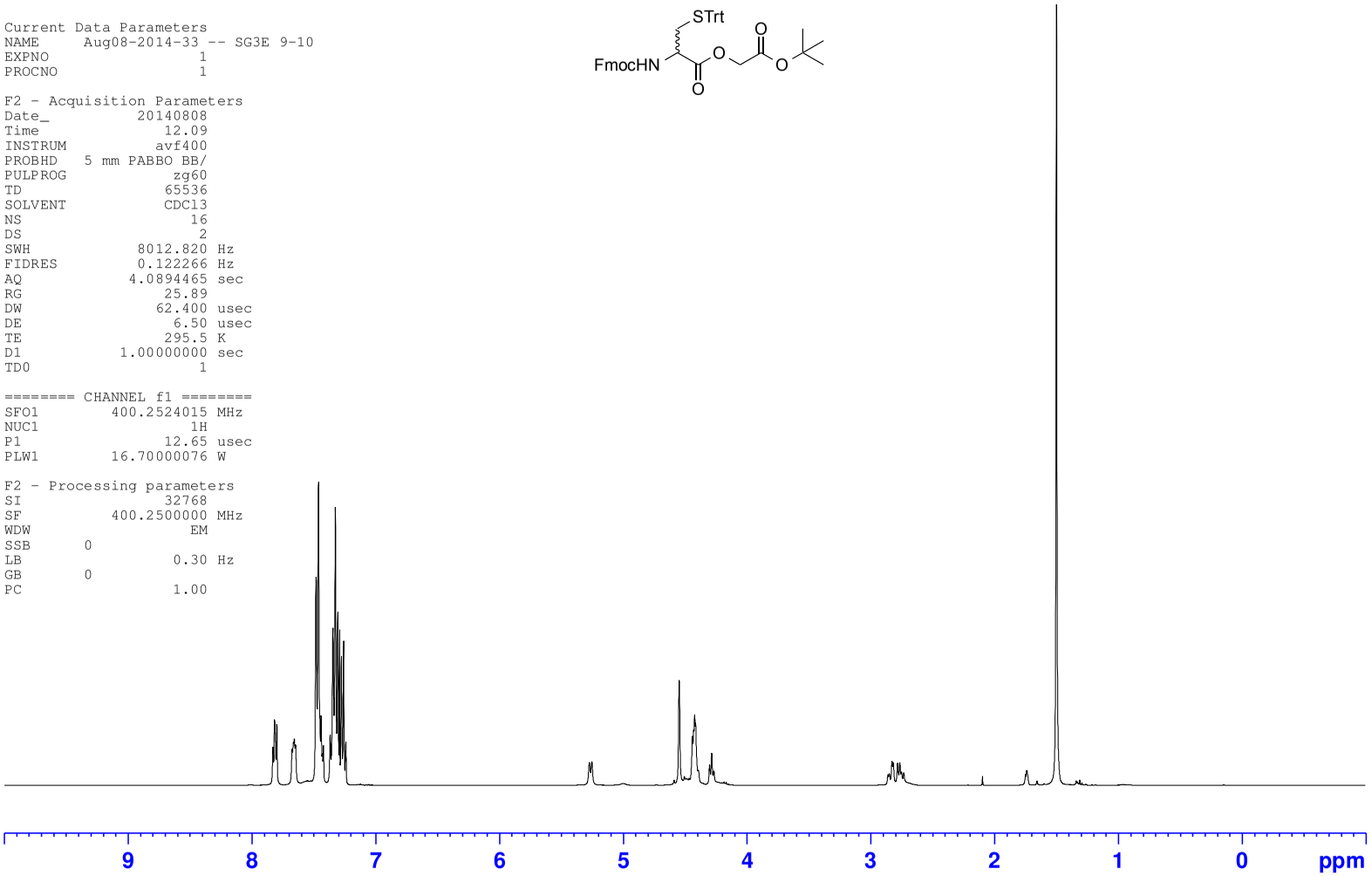
```
Current Data Parameters
NAME      sg88021508 -- SG9E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140816
Time              23.28
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG            zg30
TD                65536
SOLVENT             D2O
NS                   16
DS                    4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                    4
DW               48.400 usec
DE               10.00 usec
TE               298.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        500.3030896 MHz
NUC1                1H
P1                15.00 usec
PLW1         7.99830008 W

F2 - Processing parameters
SI                65536
SF           500.2999578 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```



^{13}C NMR - *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) D-cysteinylglycine 2,2,2-trifluoroacetate (119)



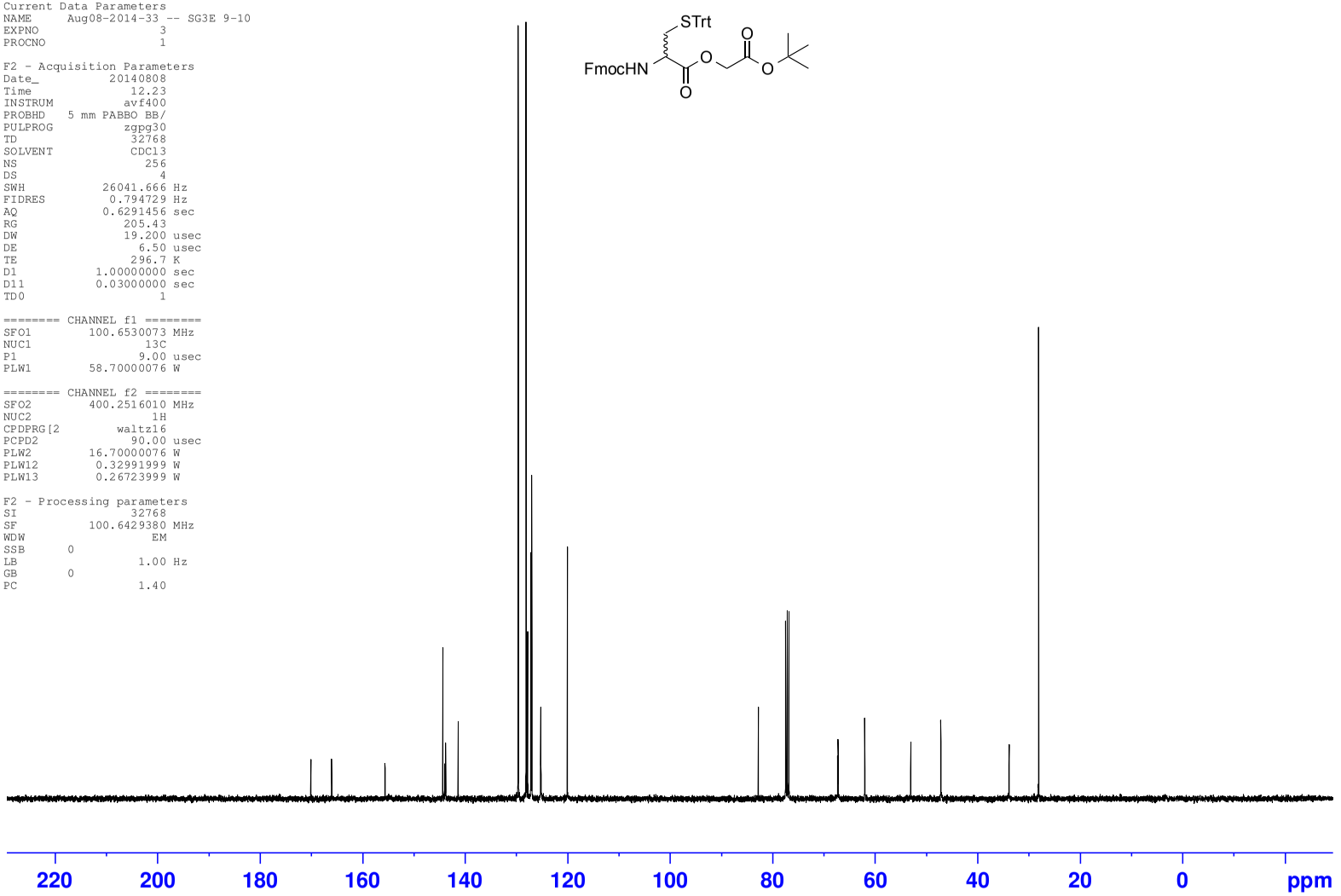
^{1}H NMR - 2-(*tert*-Butoxy)-2-oxoethyl *N*-Fmoc-*S*-trityl-L-cysteinate (120)




```
Current Data Parameters
NAME      Aug08-2014-33 -- SG3E 9-10
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140808
Time              12.09
INSTRUM           avf400
PROBHD   5 mm PABBO BB/
PULPROG             zg60
TD                 65536
SOLVENT            CDCl3
NS                    16
DS                     2
SWH             8012.820 Hz
FIDRES          0.122266 Hz
AQ              4.0894465 sec
RG                 25.89
DW                62.400 usec
DE                 6.50 usec
TE                295.5 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         400.2524015 MHz
NUC1                   1H
P1                 12.65 usec
PLW1         16.70000076 W

F2 - Processing parameters
SI                 32768
SF           400.2500000 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```

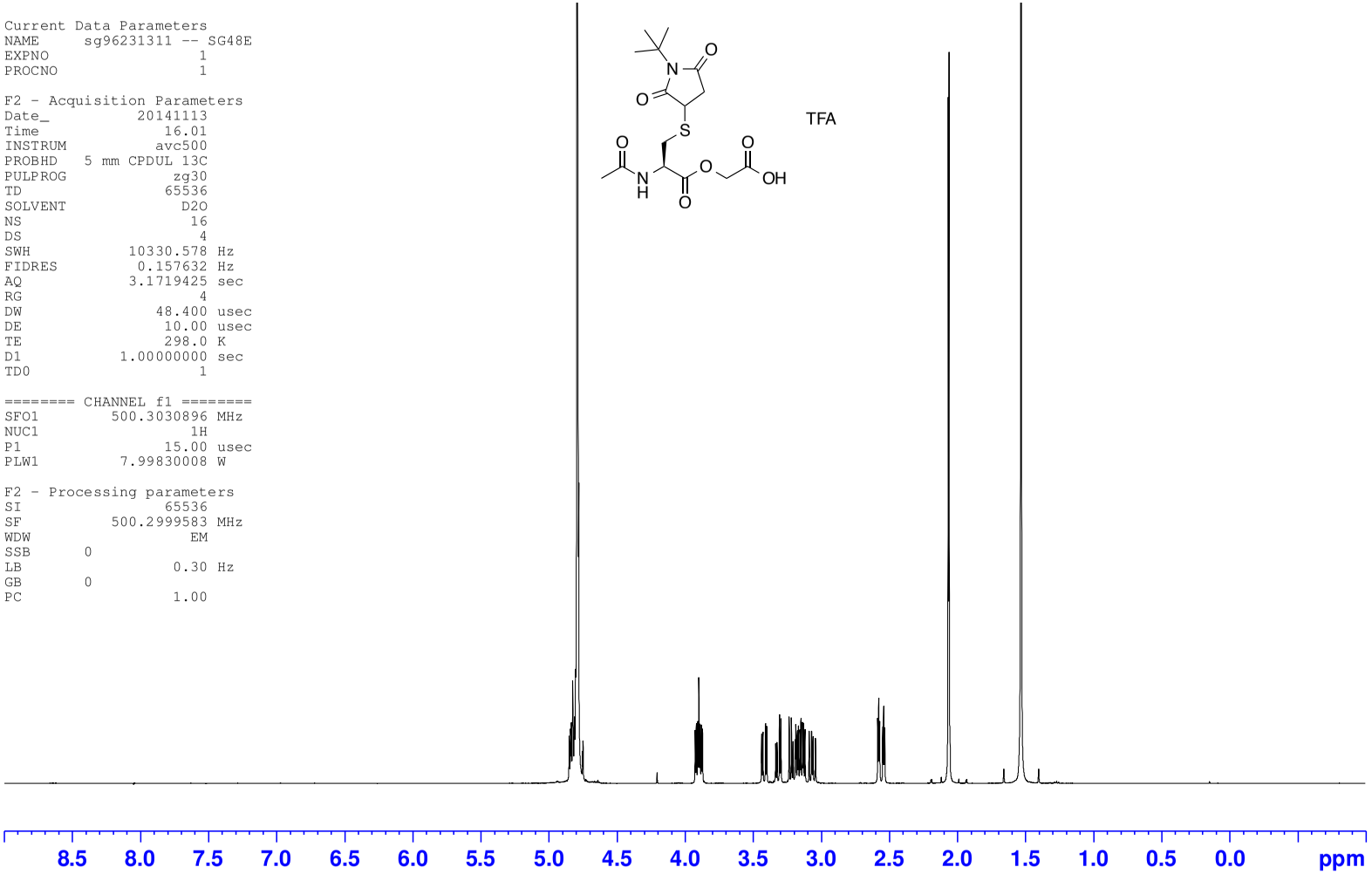
^{13}C NMR - 2-(*tert*-Butoxy)-2-oxoethyl *N*-Fmoc-*S*-trityl-L-cysteinate (120)

```
Current Data Parameters
NAME      Aug08-2014-33 -- SG3E 9-10
EXPNO                  3
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140808
Time              12.23
INSTRUM           avf400
PROBHD   5 mm PABBO BB/
PULPROG          zgpg30
TD                32768
SOLVENT           CDCl3
NS                  256
DS                    4
SWH           26041.666 Hz
FIDRES         0.794729 Hz
AQ            0.6291456 sec
RG               205.43
DW               19.200 usec
DE                6.50 usec
TE               296.7 K
D1           1.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        100.6530073 MHz
NUC1               13C
P1                9.00 usec
PLW1       58.70000076 W

======== CHANNEL f2 ========
SFO2        400.2516010 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2             90.00 usec
PLW2       16.70000076 W
PLW12       0.32991999 W
PLW13       0.26723999 W

F2 - Processing parameters
SI                32768
SF          100.6429380 MHz
WDW                EM
SSB      0
LB                1.00 Hz
GB       0
PC                1.40
```

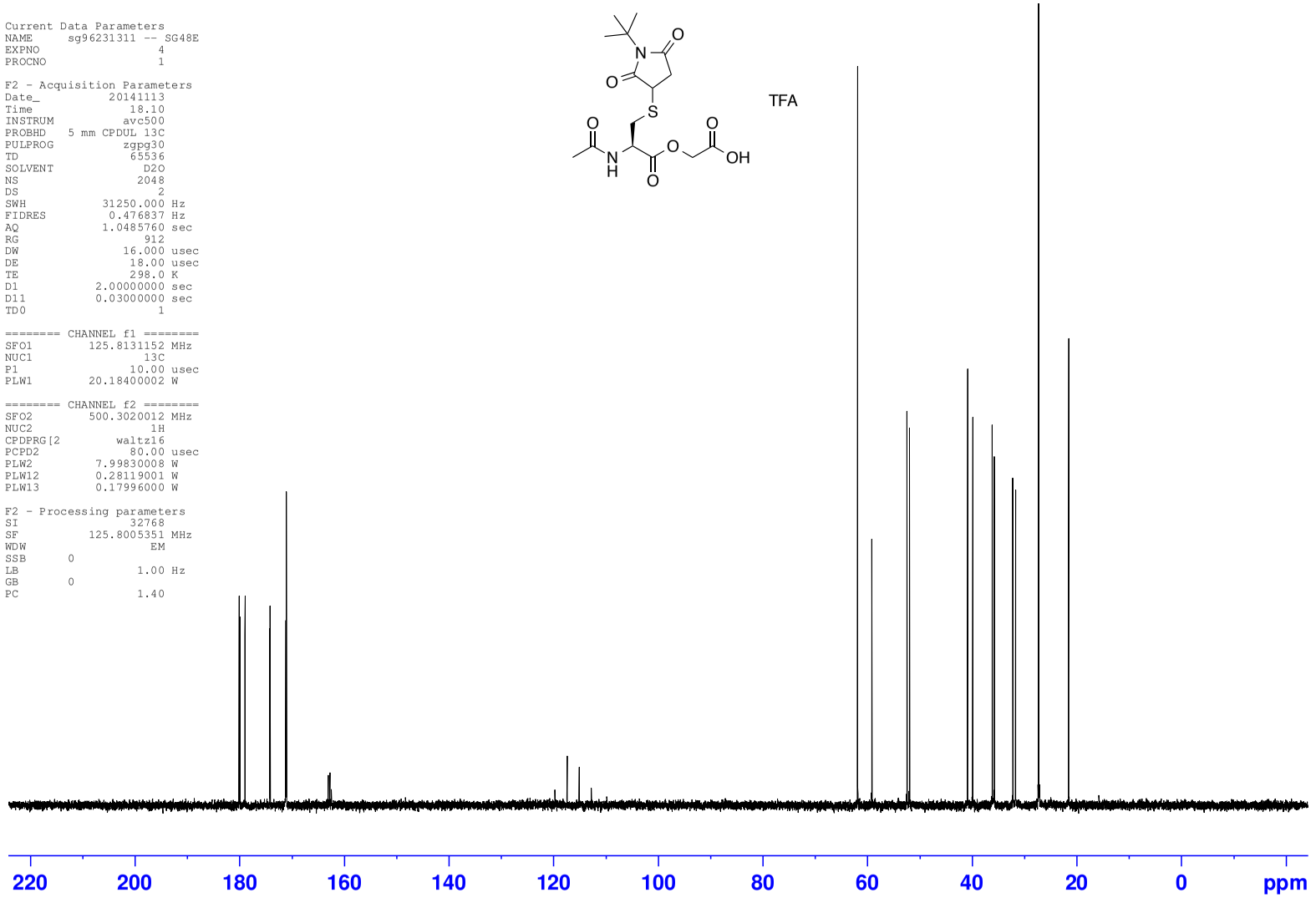


^{1}H NMR - 2-((*N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinyl)oxy)acetic acid 2,2,2-trifluoroacetate (125)

```
Current Data Parameters
NAME      sg96231311 -- SG48E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20141113
Time              16.01
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG            zg30
TD                65536
SOLVENT             D2O
NS                   16
DS                    4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                    4
DW               48.400 usec
DE               10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                 1H
P1                15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                65536
SF           500.2999583 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```

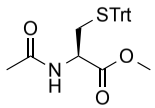


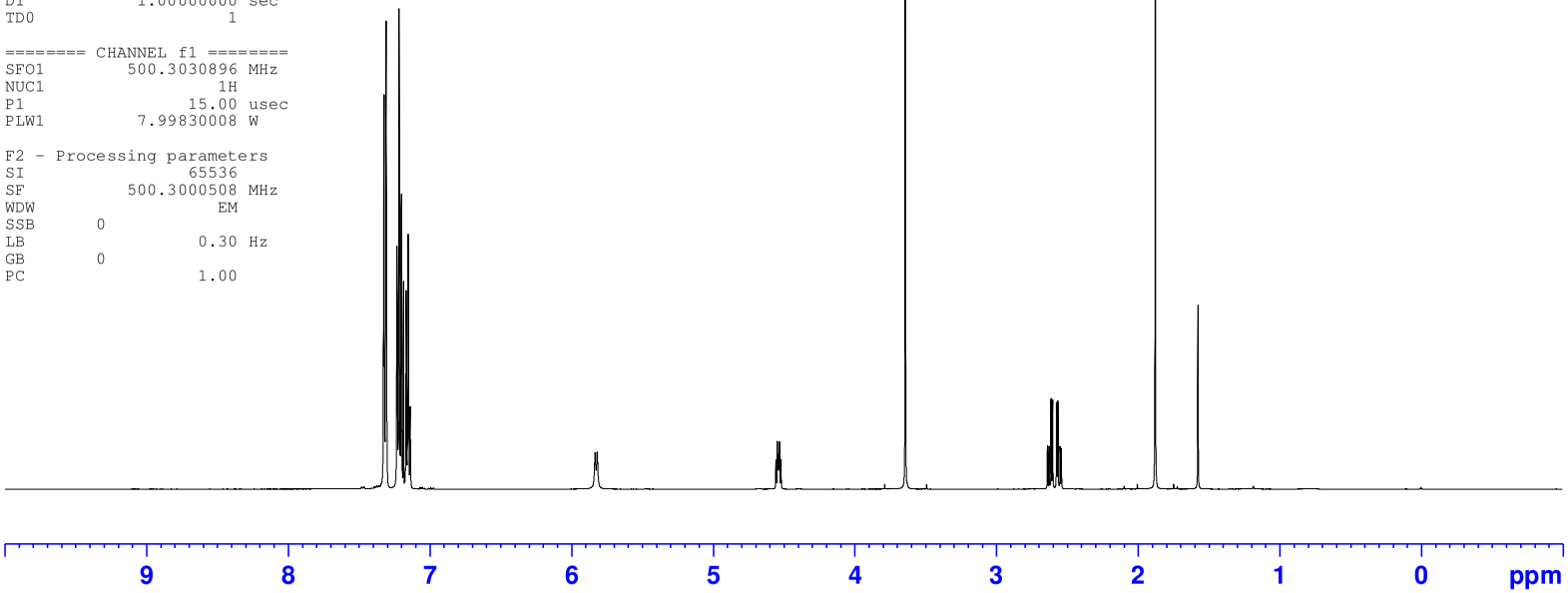
^{13}C NMR - 2-((*N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinyl)oxy)acetic acid 2,2,2-trifluoroacetate (125)



^{1}H NMR - Methyl *N*-acetyl-*S*-trityl-L-cysteinate (128)

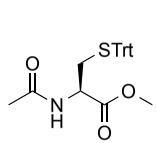
```
Current Data Parameters
NAME      sg97202111 -- SG51E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20141122
Time               18.36
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT            CDCl3
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                  3.56
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                  15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.3000508 MHz
WDW                   EM
SSB      0
LB                  0.30 Hz
GB       0
PC                  1.00
```



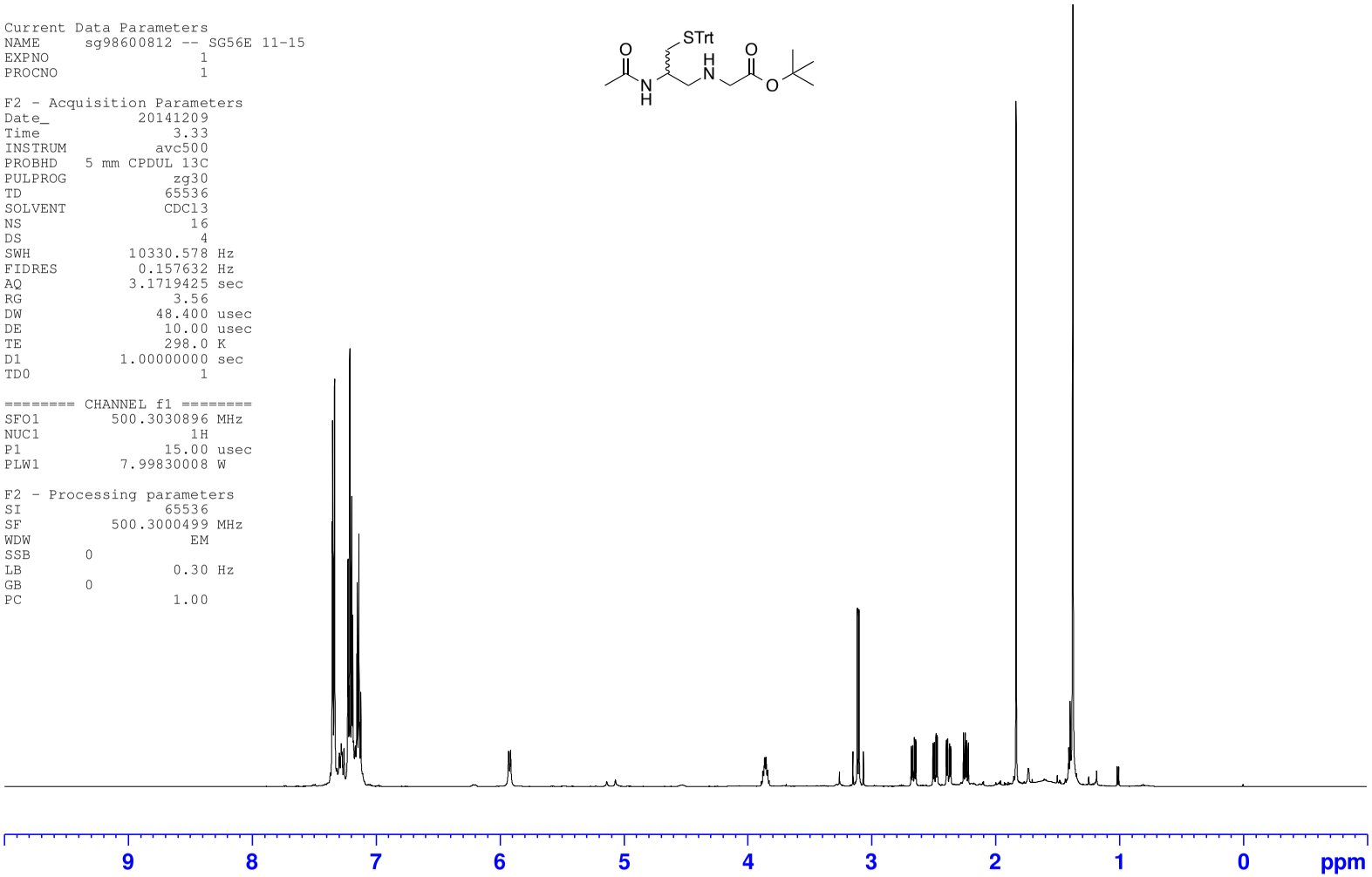


^{13}C NMR - Methyl *N*-acetyl-*S*-trityl-L-cysteinate (128)

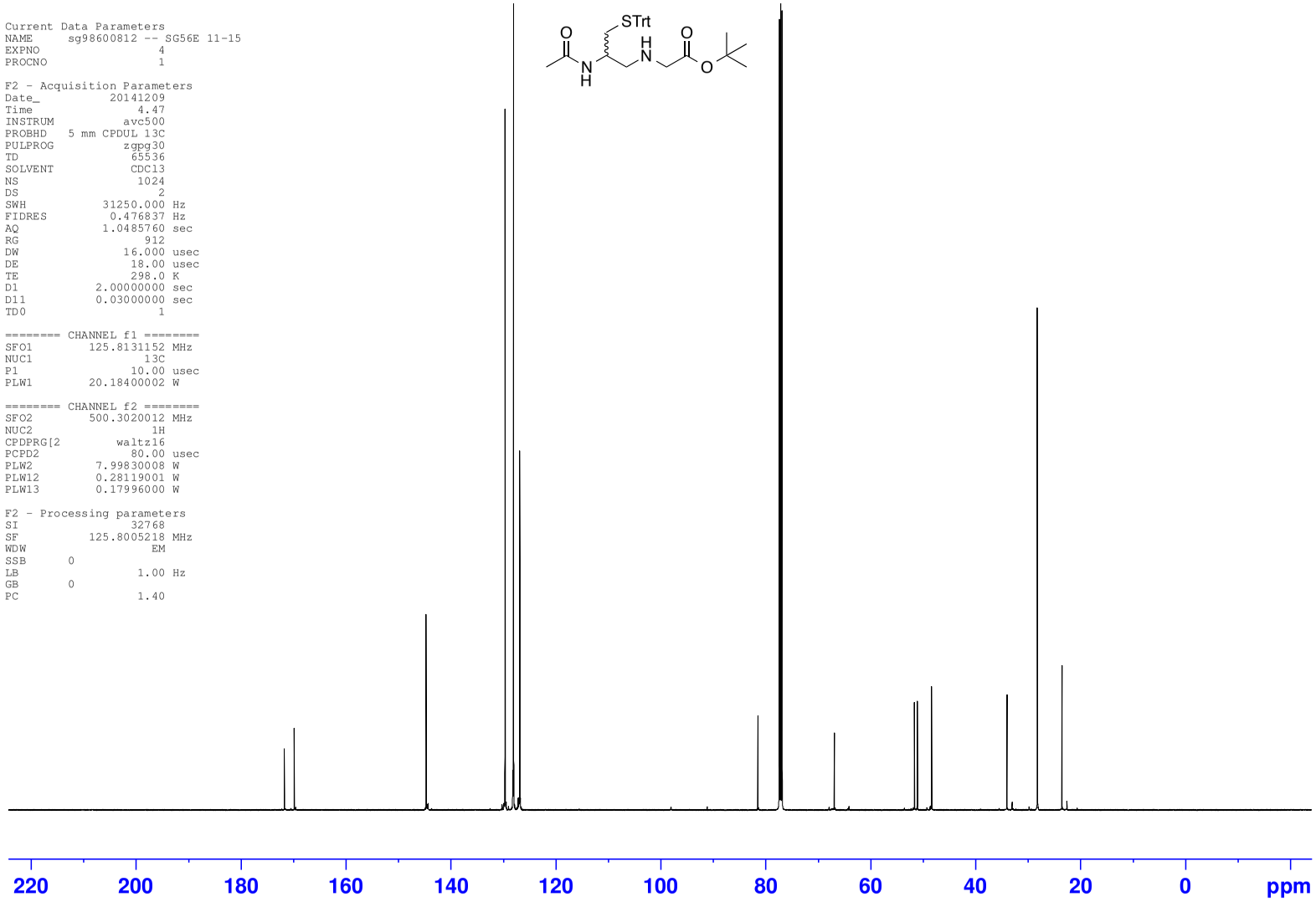
```
Current Data Parameters
NAME      sg97202111 -- SG51E
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_          20141122
Time             19.51
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD               65536
SOLVENT          CDCl3
NS                1024
DS                   2
SWH          31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                 912
DW              16.000 usec
DE              18.00 usec
TE              298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                  1

======== CHANNEL f1 ========
SFO1        125.8131152 MHz
NUC1               13C
P1              10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2            80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8005222 MHz
WDW                EM
SSB      0
LB                1.00 Hz
GB       0
PC                1.40
```



^{1}H NMR - *N*-Acetyl-*S*-trityl-*tert*-butyl (2-amino-3-mercaptopropyl)glycinate (130)

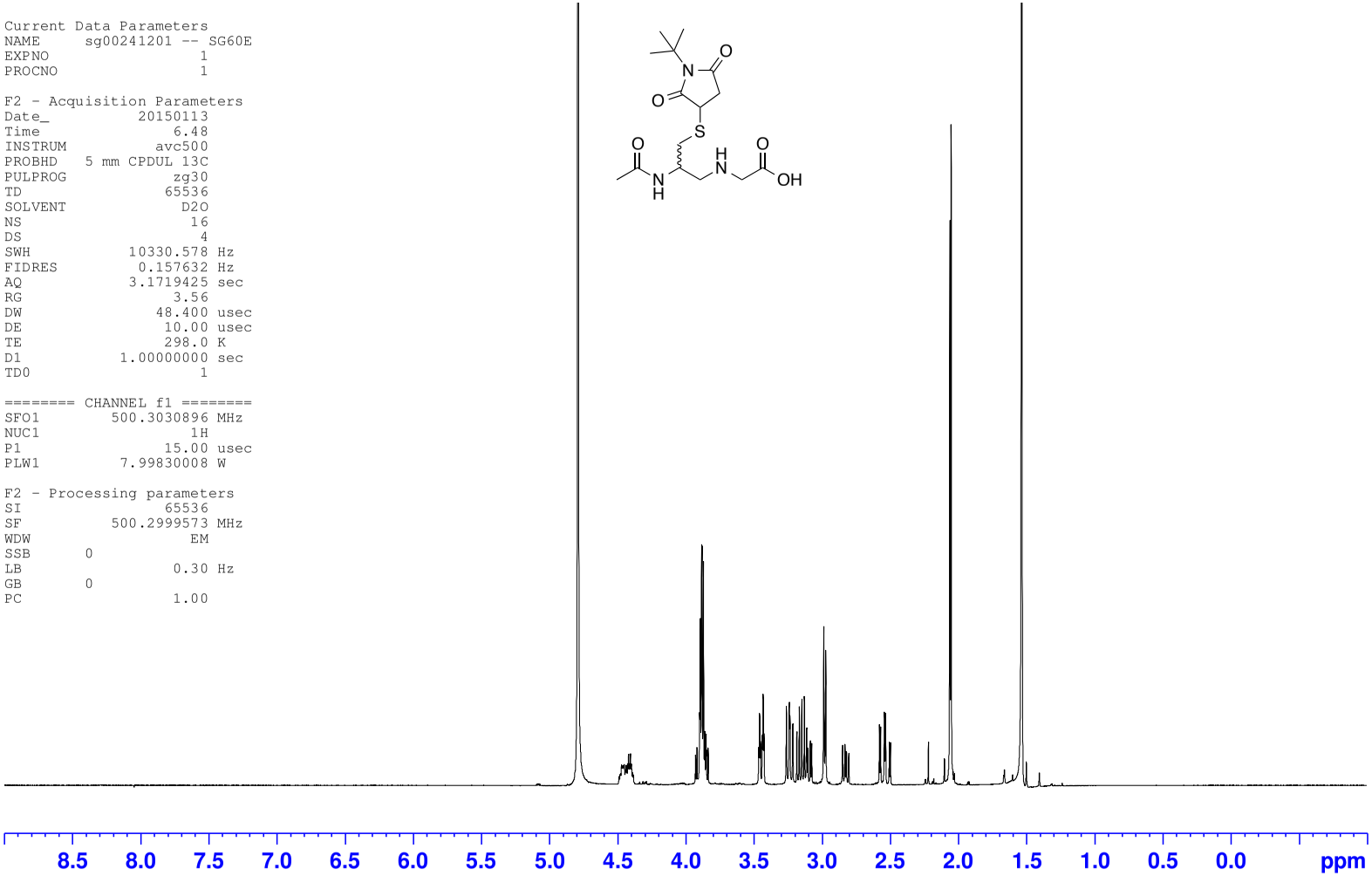


^{13}C NMR - *N*-Acetyl-*S*-trityl-*tert*-butyl (2-amino-3-mercaptopropyl)glycinate (130)

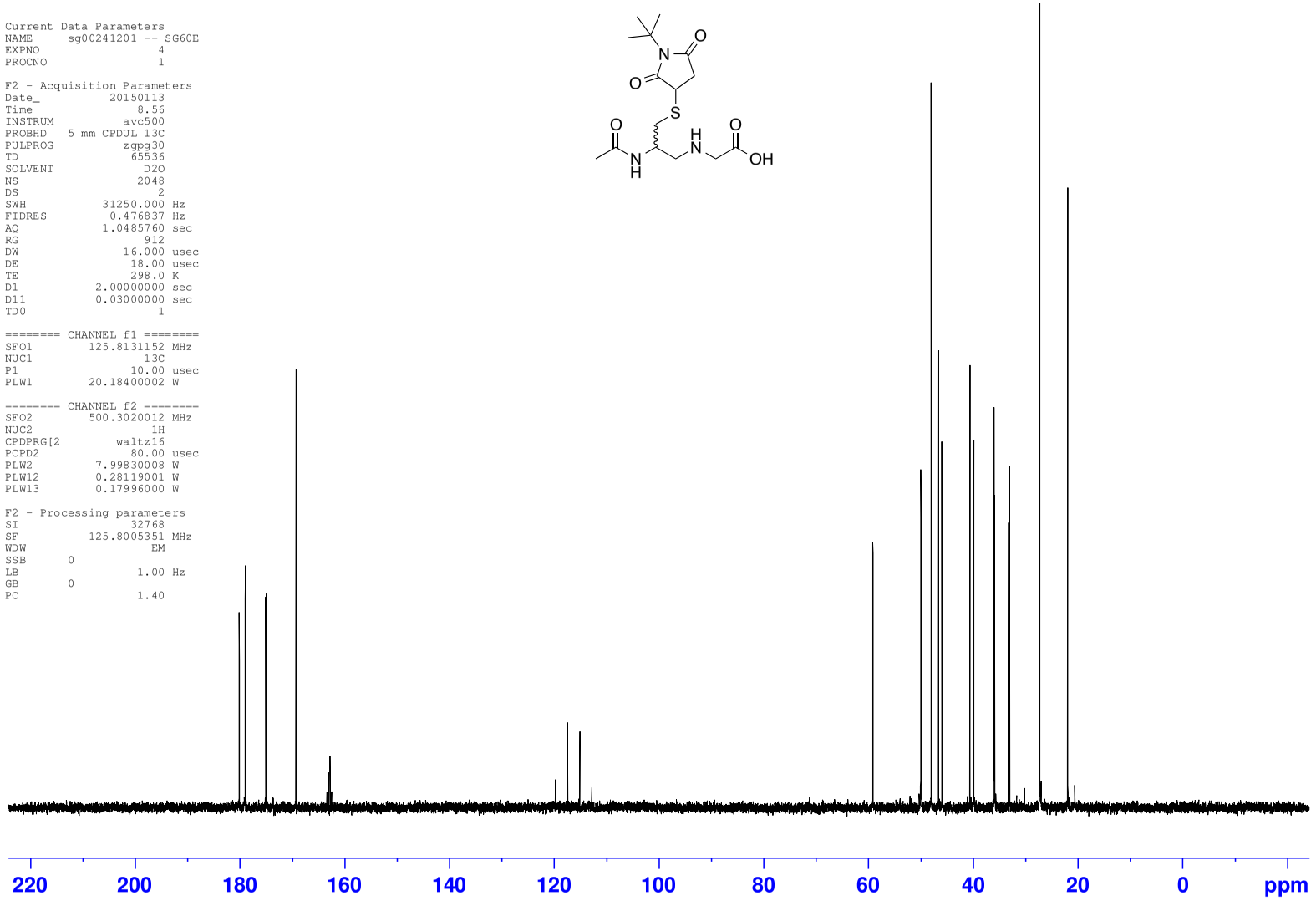


^{1}H NMR - *S*-(*N*-*tert*-Butylsuccinimido) (2-acetamido-3-mercaptopropyl)glycine 2,2,2-trifluoroacetate (132)

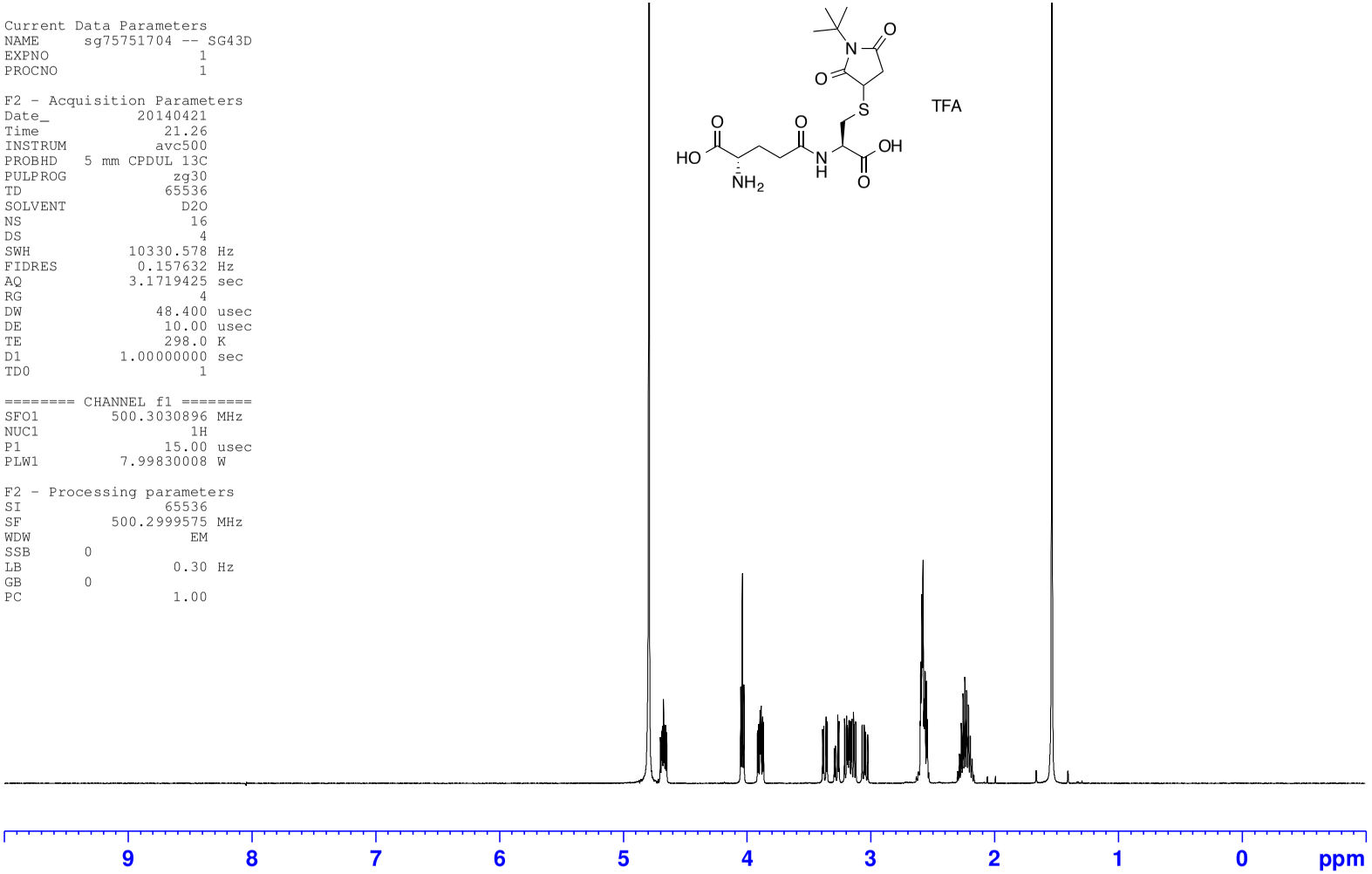
```
Current Data Parameters
NAME      sg00241201 -- SG60E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20150113
Time                6.48
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                  3.56
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                  15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999573 MHz
WDW                   EM
SSB       0
LB                  0.30 Hz
GB        0
PC                  1.00
```



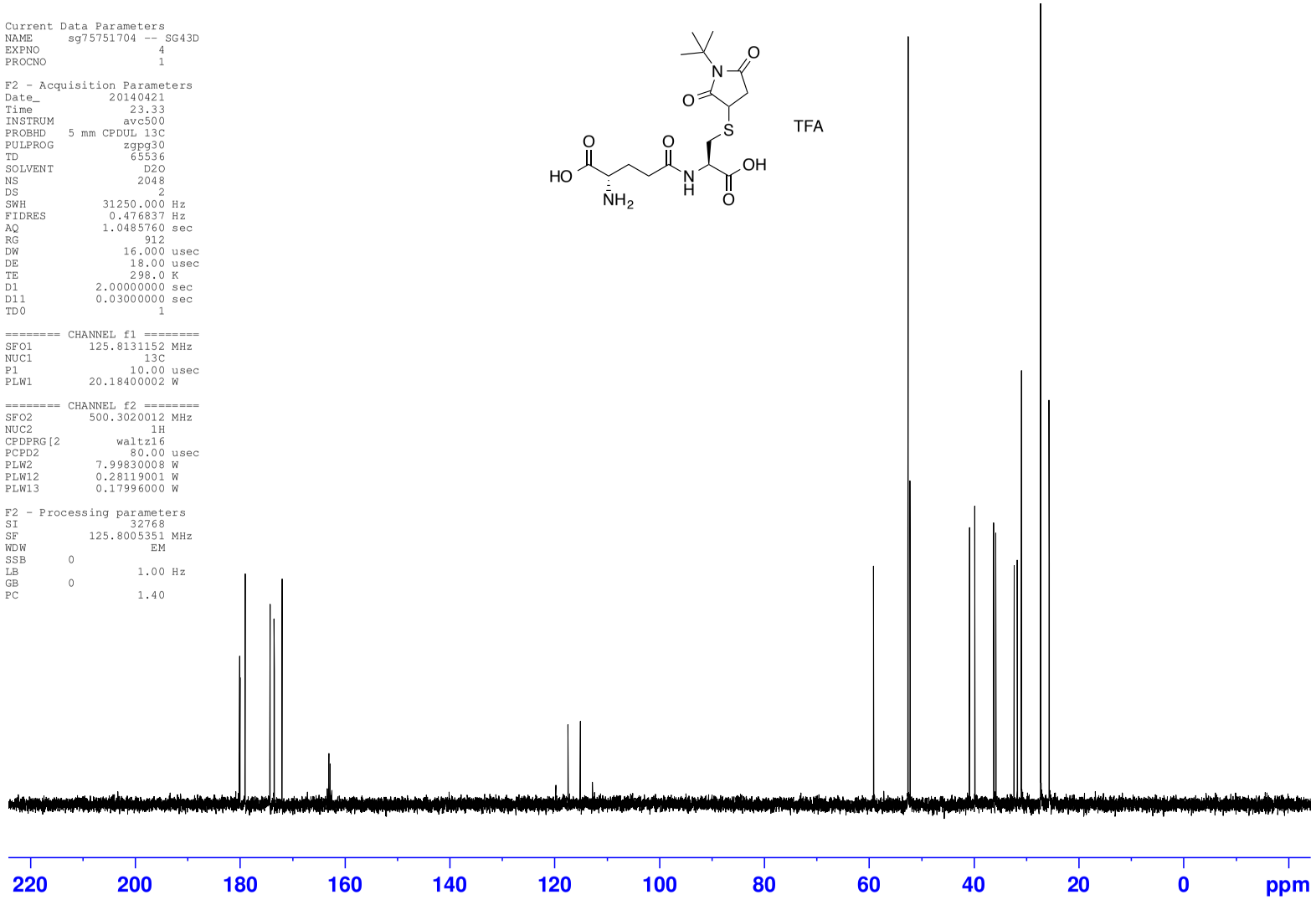
^{13}C NMR - *S*-(*N*-*tert*-Butylsuccinimido) (2-acetamido-3-mercaptopropyl)glycine 2,2,2-trifluoroacetate (132)



^{1}H NMR *S*-(*N*-*tert*-Butylsuccinimido) γ-L-glutamyl-L-cysteine 2,2,2-trifluoroacetate (134)

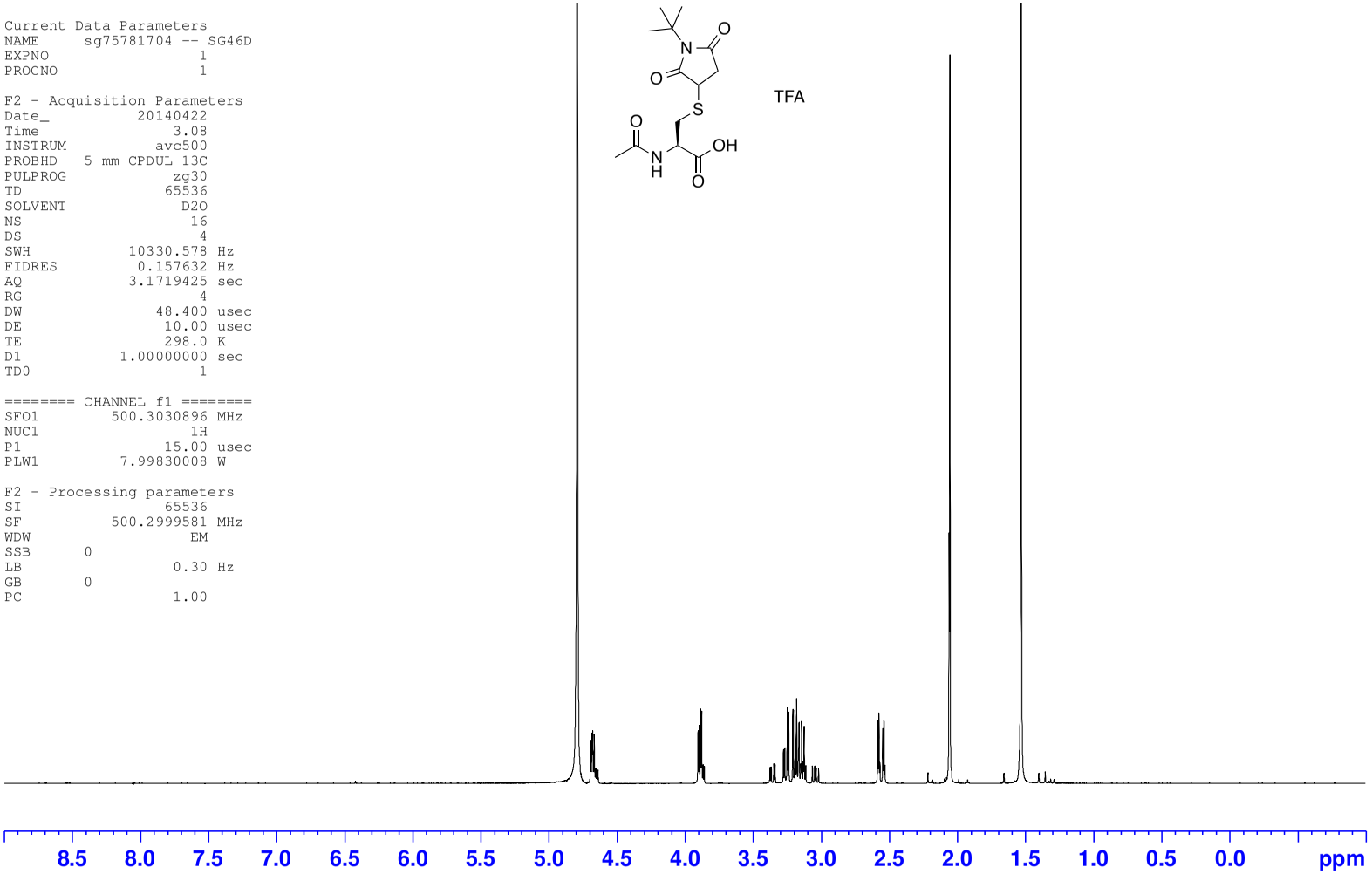


^{13}C NMR *S*-(*N*-*tert*-Butylsuccinimido) γ-L-glutamyl-L-cysteine 2,2,2-trifluoroacetate (134)

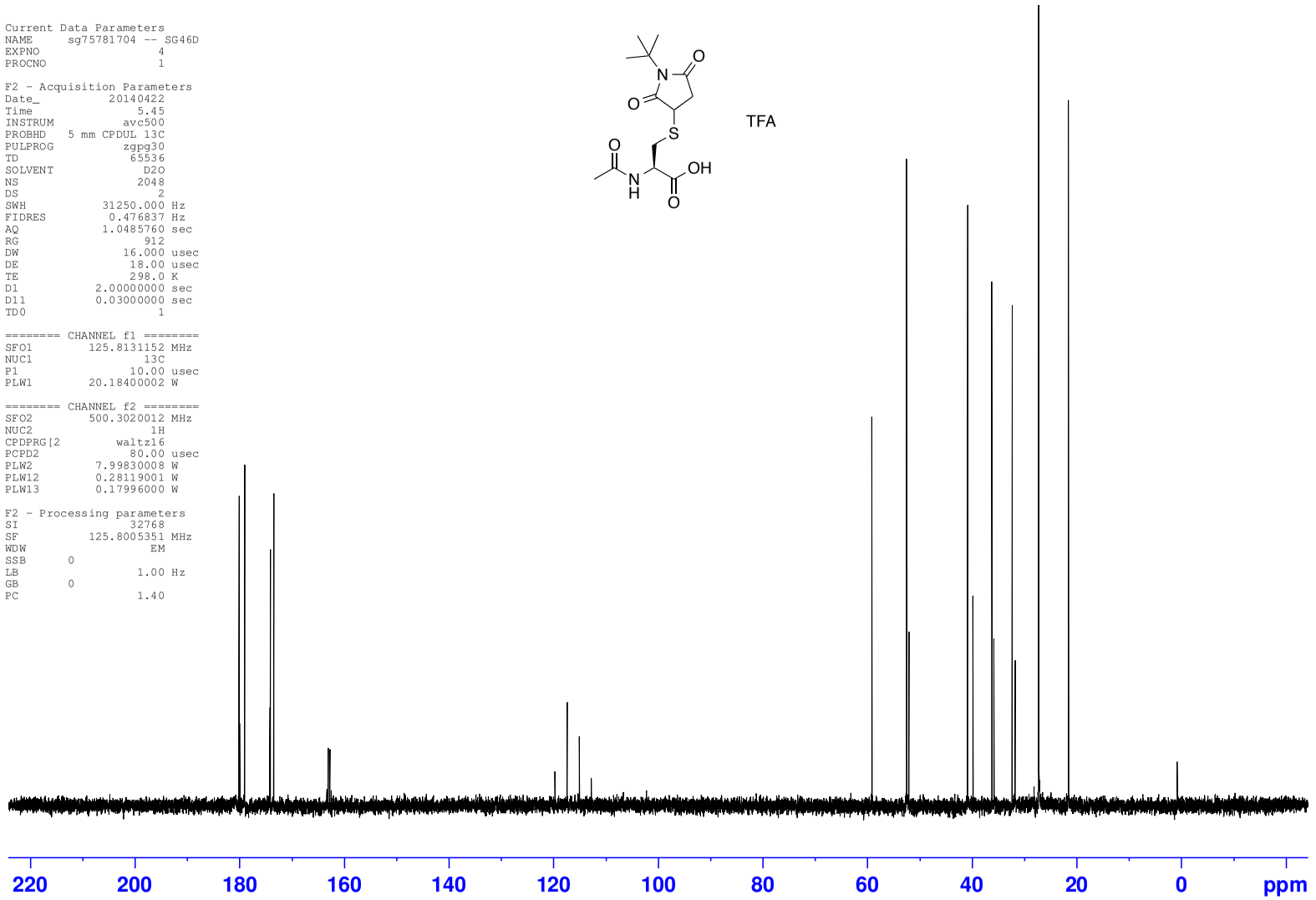


^{1}H NMR - *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteine 2,2,2-trifluoroacetate (135)

```
Current Data Parameters
NAME      sg75781704 -- SG46D
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20140422
Time                3.08
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                  1H
P1                15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999581 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```

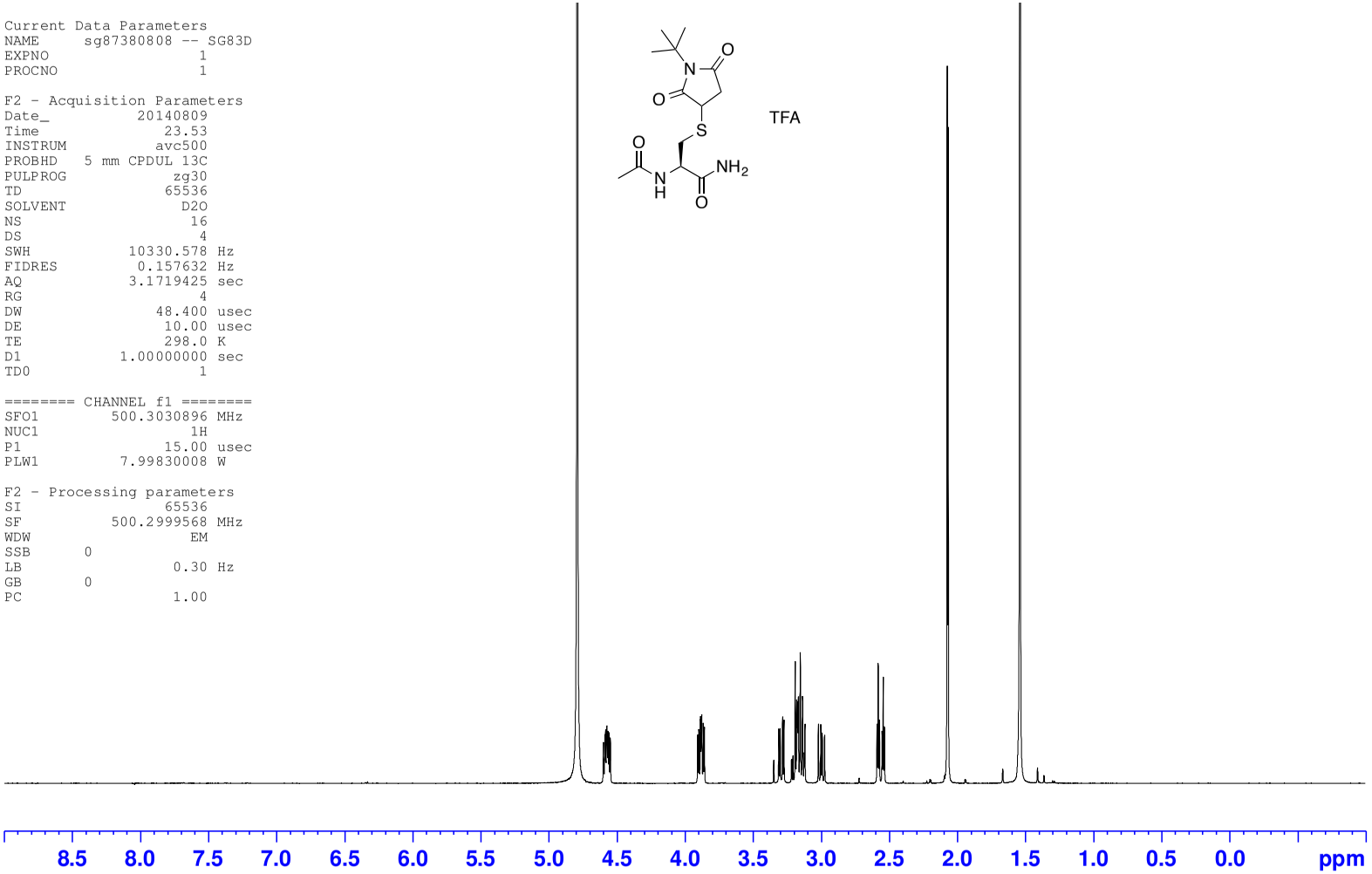


^{13}C NMR - *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteine 2,2,2-trifluoroacetate (135)



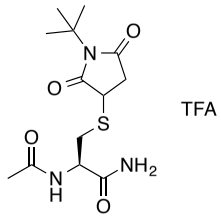
^{1}H NMR - *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinamide 2,2,2-trifluoroacetate (138)

```
Current Data Parameters
NAME      sg87380808 -- SG83D
EXPNO                1
PROCNO               1

F2 - Acquisition Parameters
Date_           20140809
Time              23.53
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG            zg30
TD                65536
SOLVENT             D2O
NS                   16
DS                    4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                    4
DW               48.400 usec
DE               10.00 usec
TE               298.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                1H
P1                15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                65536
SF           500.2999568 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```



^{13}C NMR - *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinamide 2,2,2-trifluoroacetate (138)

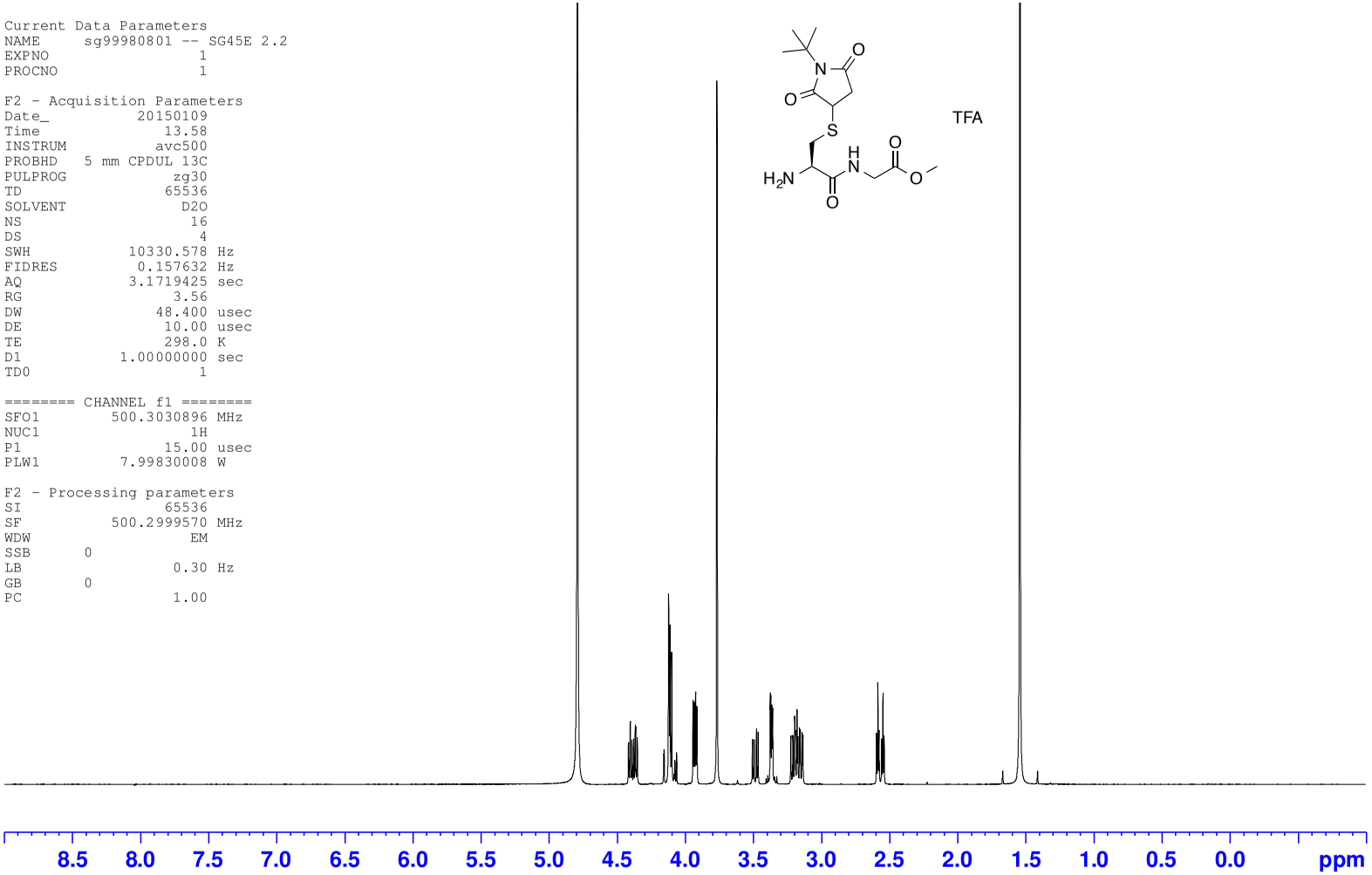
```
Current Data Parameters
NAME      sg87380808 -- SG83D
EXPNO                4
PROCNO               1

F2 - Acquisition Parameters
Date_          20140810
Time               1.08
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT             D2O
NS                 1024
DS                    2
SWH           31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               18.00 usec
TE               298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        125.8131152 MHz
NUC1                13C
P1                10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2                 1H
CPDPRG[2        waltz16
PCPD2              80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8005351 MHz
WDW                  EM
SSB      0
LB                1.00 Hz
GB       0
PC                1.40
```



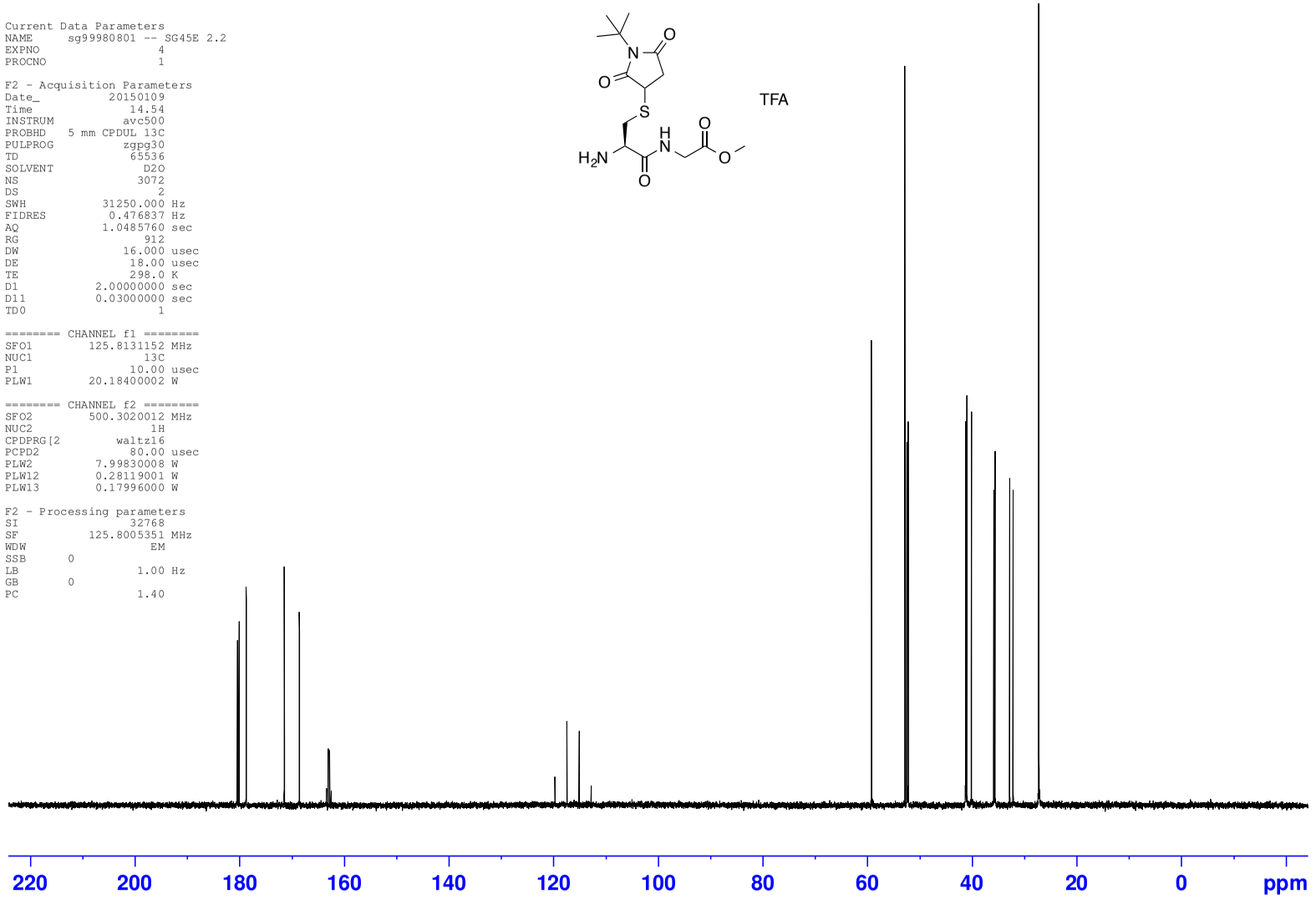
TFA

^{1}H NMR - *S*-(*N*-*tert*-Butylsuccinimido) L-cysteinylglycine methyl ester 2,2,2-trifluoroacetate (140)

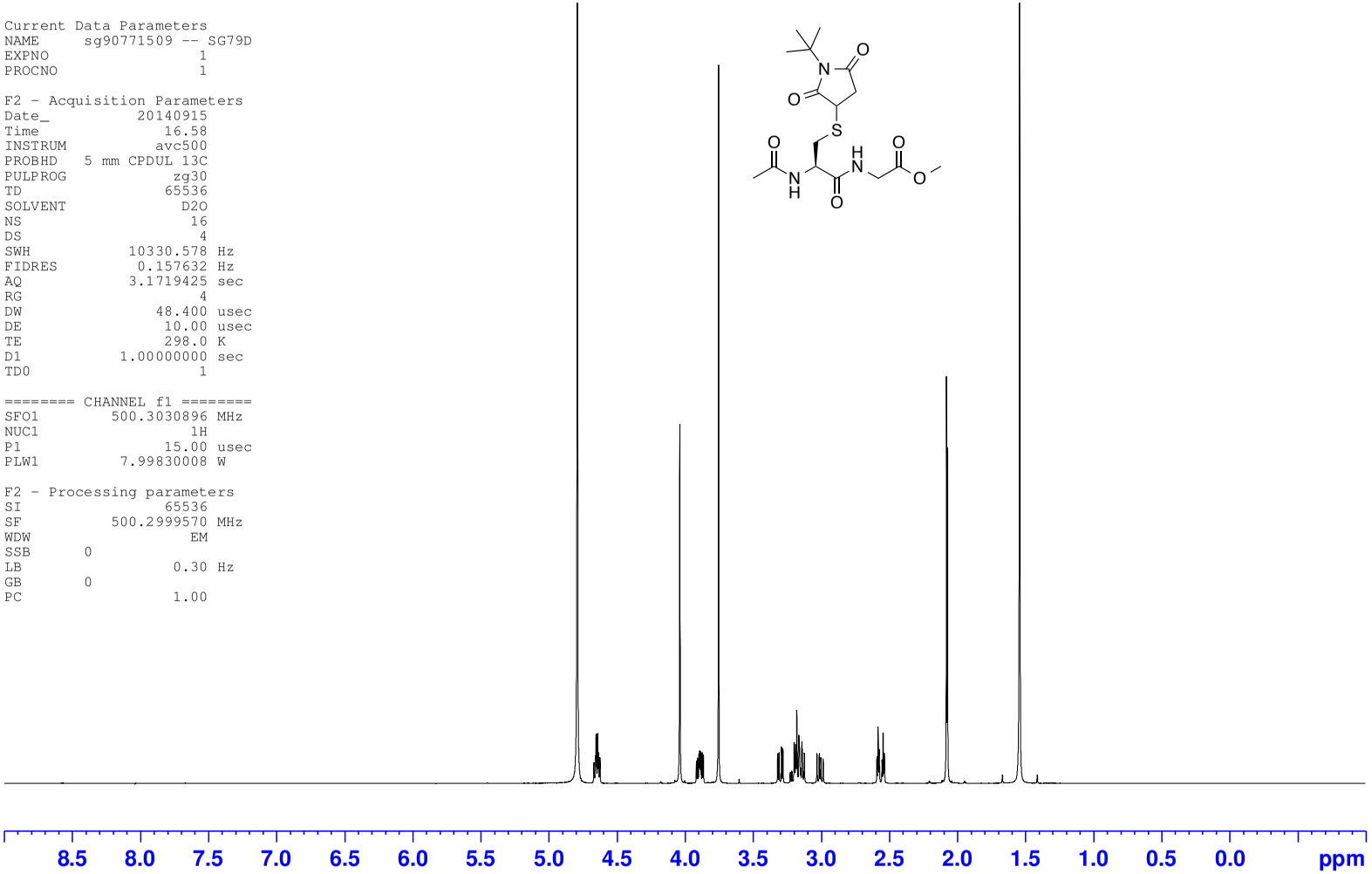
```
Current Data Parameters
NAME      sg99980801 -- SG45E 2.2
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_           20150109
Time              13.58
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                  3.56
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                1H
P1                15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.2999570 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```



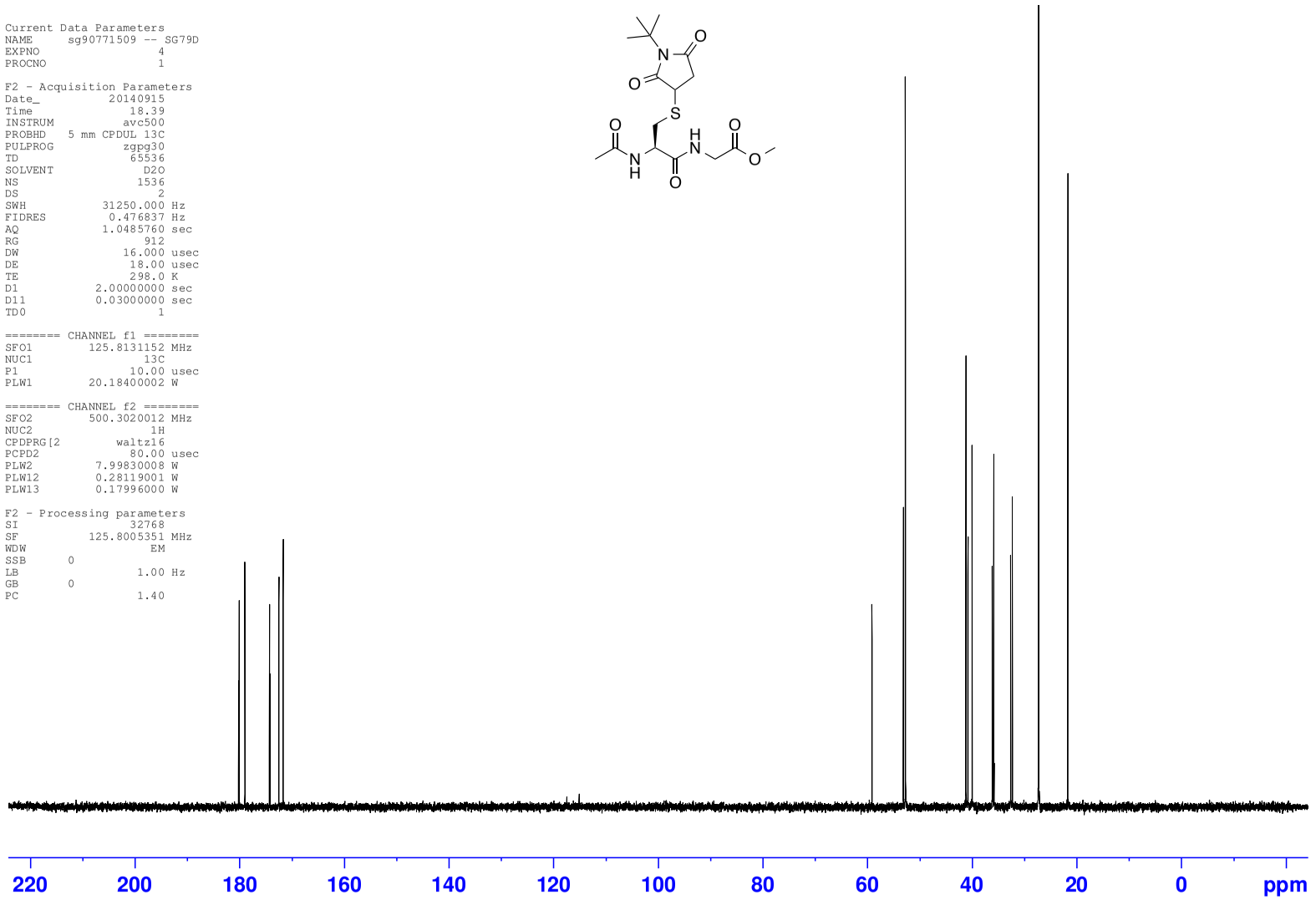
^{13}C NMR - *S*-(*N*-*tert*-Butylsuccinimido) L-cysteinylglycine methyl ester 2,2,2-trifluoroacetate (140)



^{1}H NMR *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine methyl ester (141)

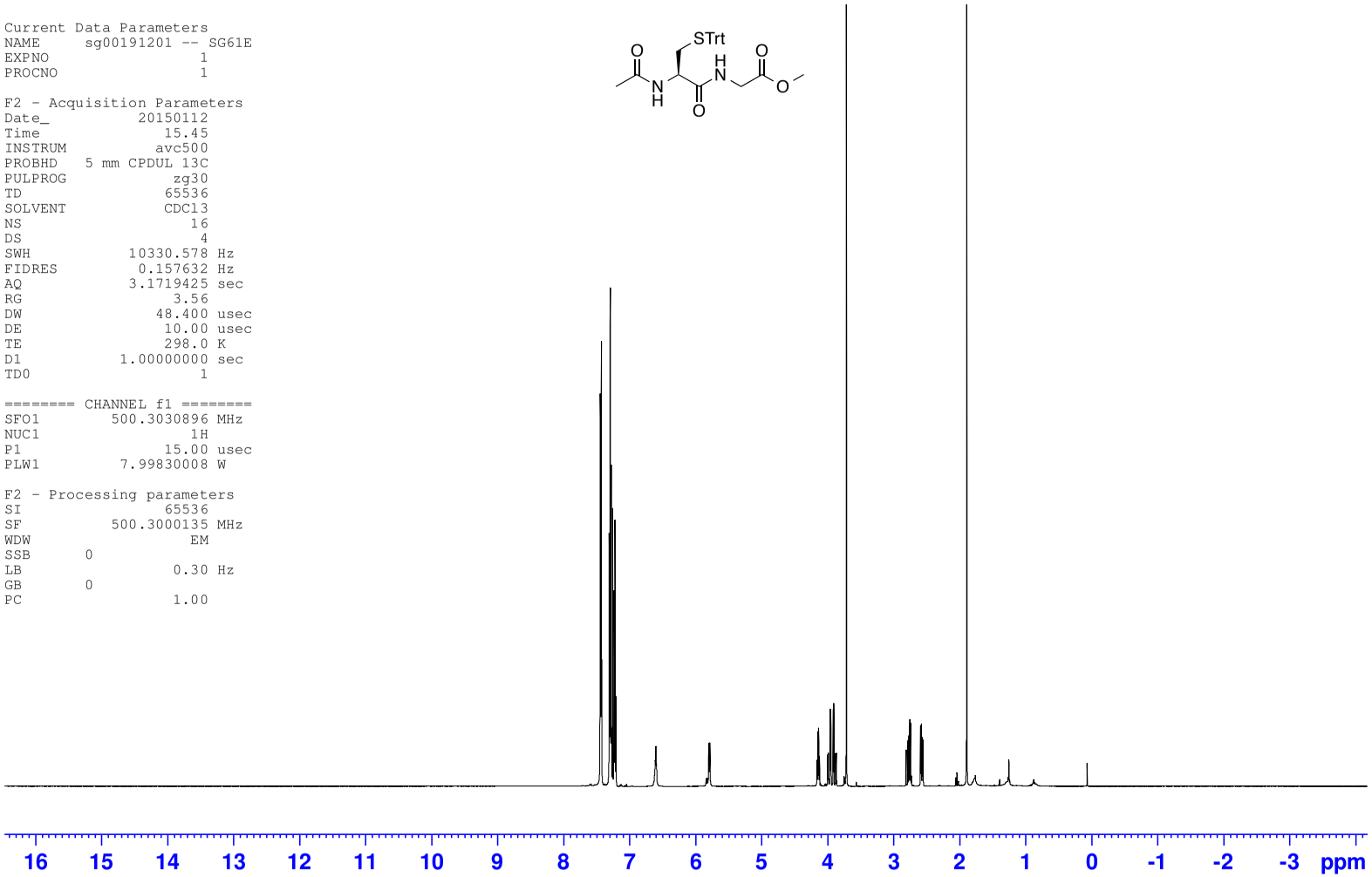


^{13}C NMR *N*-Acetyl-*S*-(*N*-*tert*-butylsuccinimido) L-cysteinylglycine methyl ester (141)



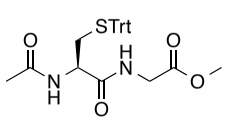
^{1}H NMR - *N*-Acetyl-*S*-trityl-L-cysteinylglycine methyl ester (142)

```
Current Data Parameters
NAME      sg00191201 -- SG61E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20150112
Time              15.45
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG            zg30
TD                65536
SOLVENT           CDCl3
NS                   16
DS                    4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                 3.56
DW               48.400 usec
DE                10.00 usec
TE                298.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        500.3030896 MHz
NUC1                   1H
P1                 15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF          500.3000135 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```



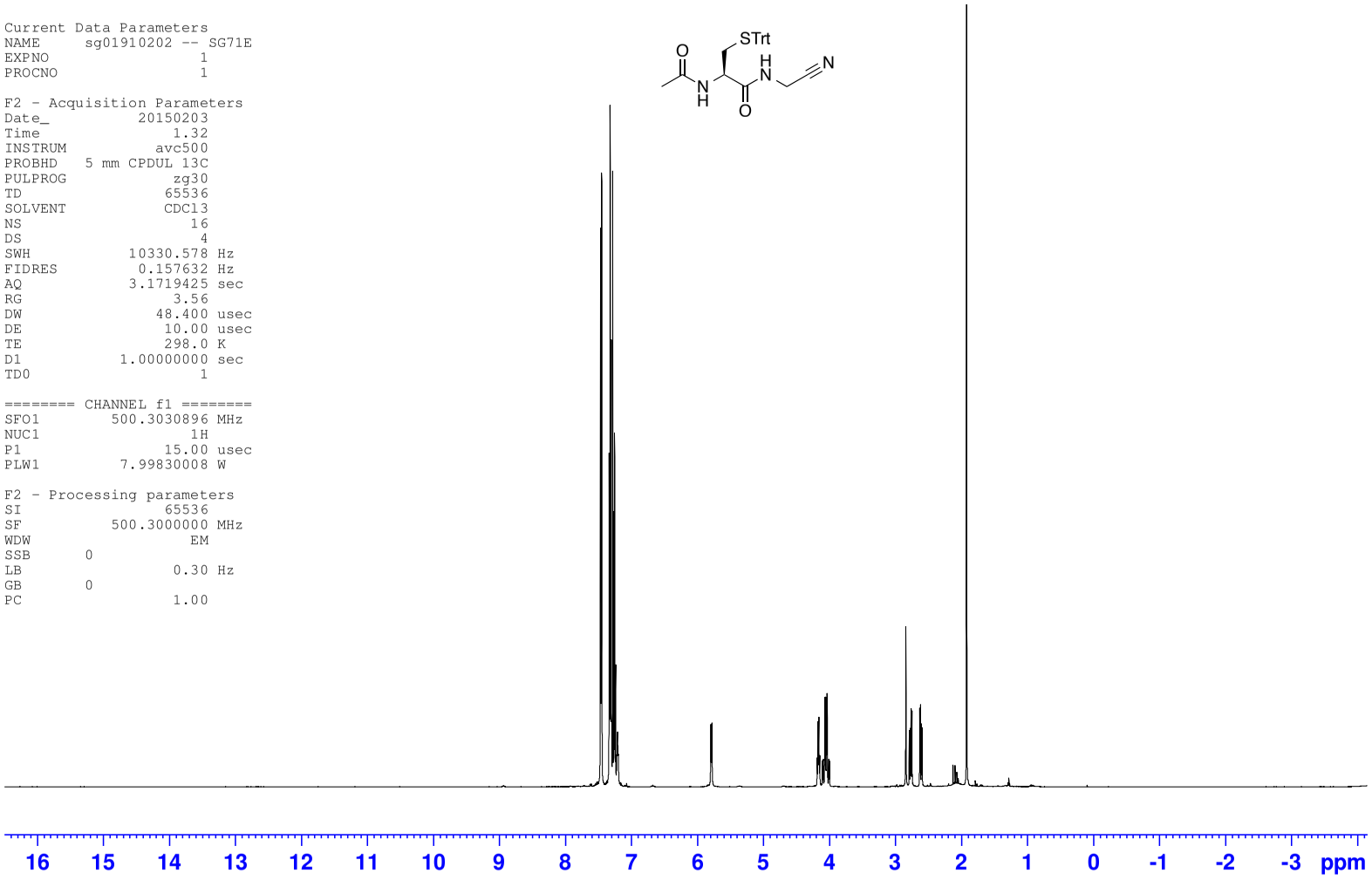
^{13}C NMR - *N*-Acetyl-*S*-trityl-L-cysteinylglycine methyl ester (142)

```
Current Data Parameters
NAME      sg00191201 -- SG61E
EXPNO                  4
PROCNO                 1

F2 - Acquisition Parameters
Date_           20150112
Time               17.54
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD                65536
SOLVENT           CDCl3
NS                 2048
DS                    2
SWH           31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.0485760 sec
RG                  912
DW               16.000 usec
DE               18.00 usec
TE               298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        125.8131152 MHz
NUC1               13C
P1               10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2            80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8005220 MHz
WDW                EM
SSB      0
LB                1.00 Hz
GB       0
PC                1.40
```

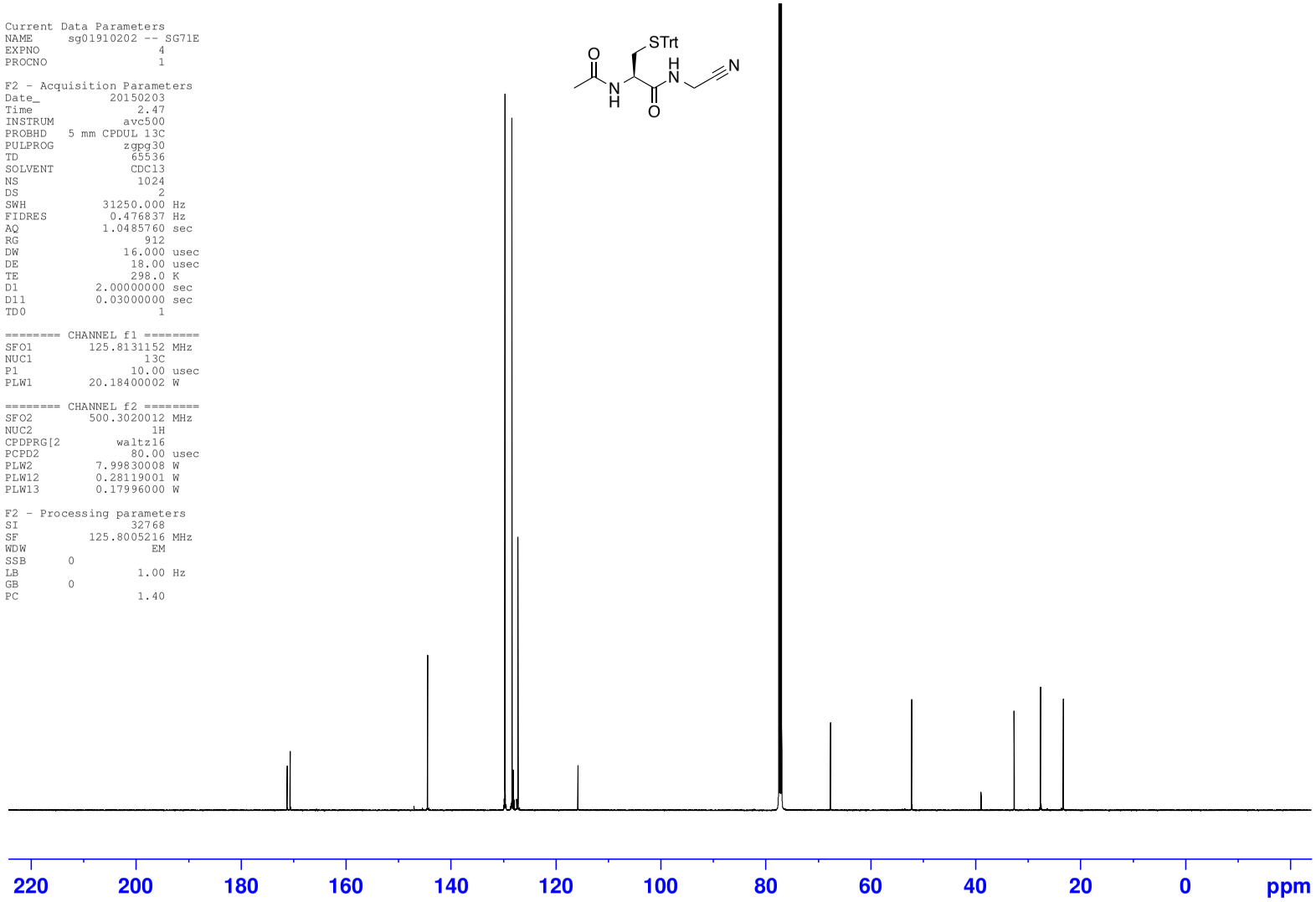


^{1}H NMR - *N*-Acetyl-*S*-trityl-(*R*)-2-amino-*N*-(cyanomethyl)-3-mercaptopropanamide (144)

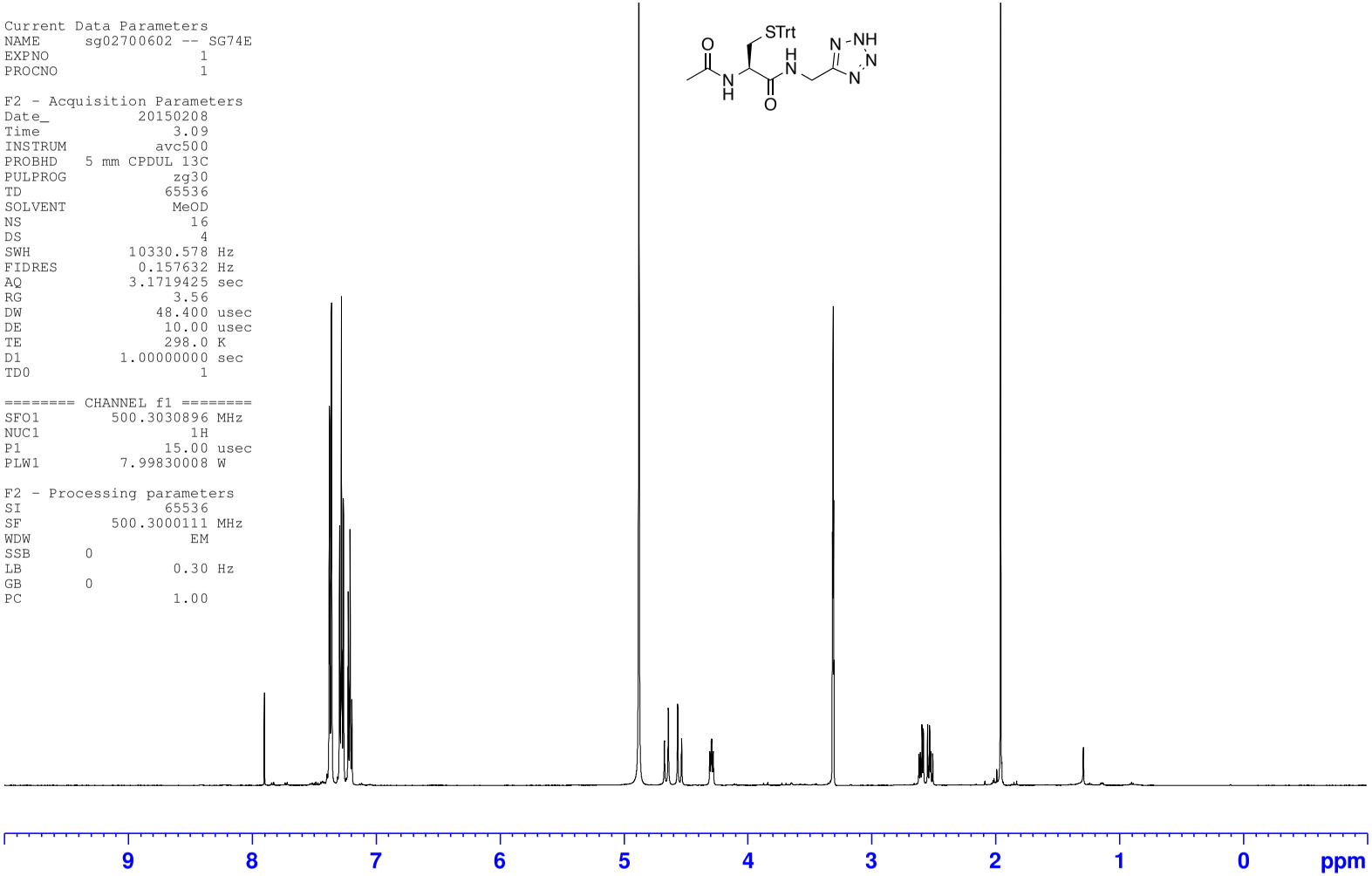
```
Current Data Parameters
NAME      sg01910202 -- SG71E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20150203
Time               1.32
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT            CDCl3
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                  3.56
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                  15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF            500.3000000 MHz
WDW                   EM
SSB       0
LB                  0.30 Hz
GB        0
PC                  1.00
```



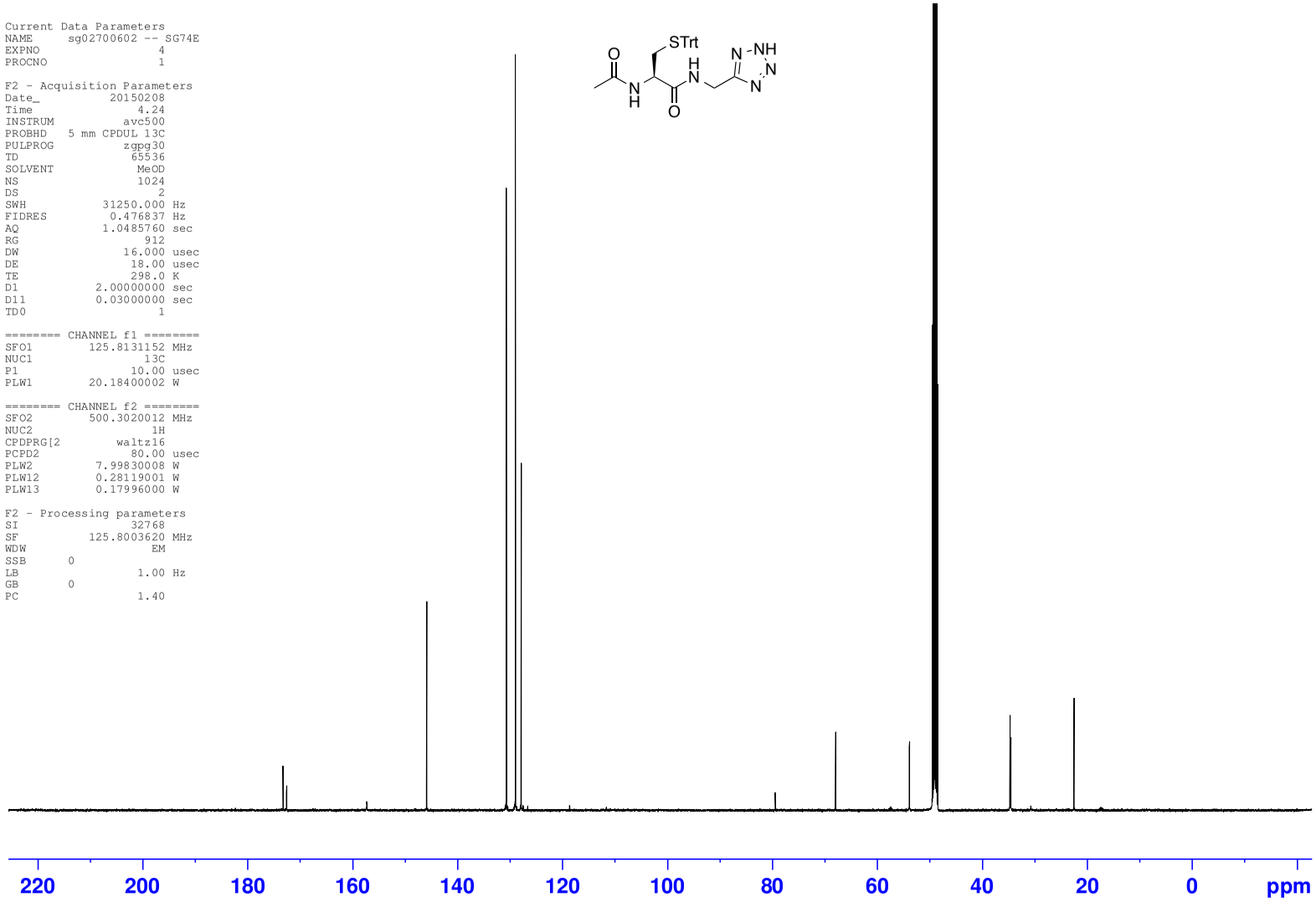
^{13}C NMR - *N*-Acetyl-*S*-trityl-(*R*)-2-amino-*N*-(cyanomethyl)-3-mercaptopropanamide (144)



^{1}H NMR - *N*-Acetyl-*S*-trityl-(*R*)-*N*-((2*H*-tetrazol-5-yl)methyl)-2-amino-3-mercaptopropanamide (145)

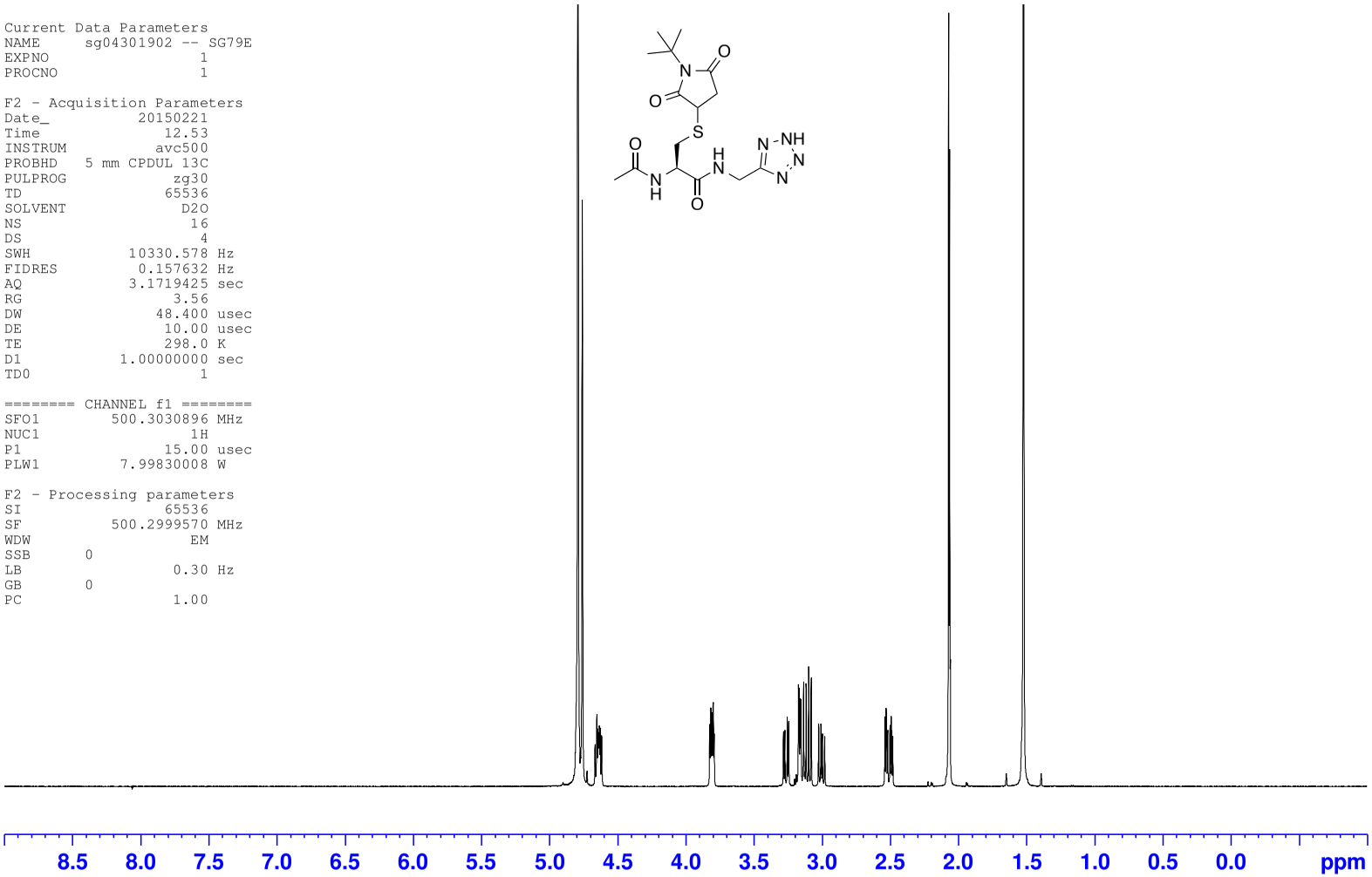


^{13}C NMR - *N*-Acetyl-*S*-trityl-(*R*)-*N*-((2*H*-tetrazol-5-yl)methyl)-2-amino-3-mercaptopropanamide (145)

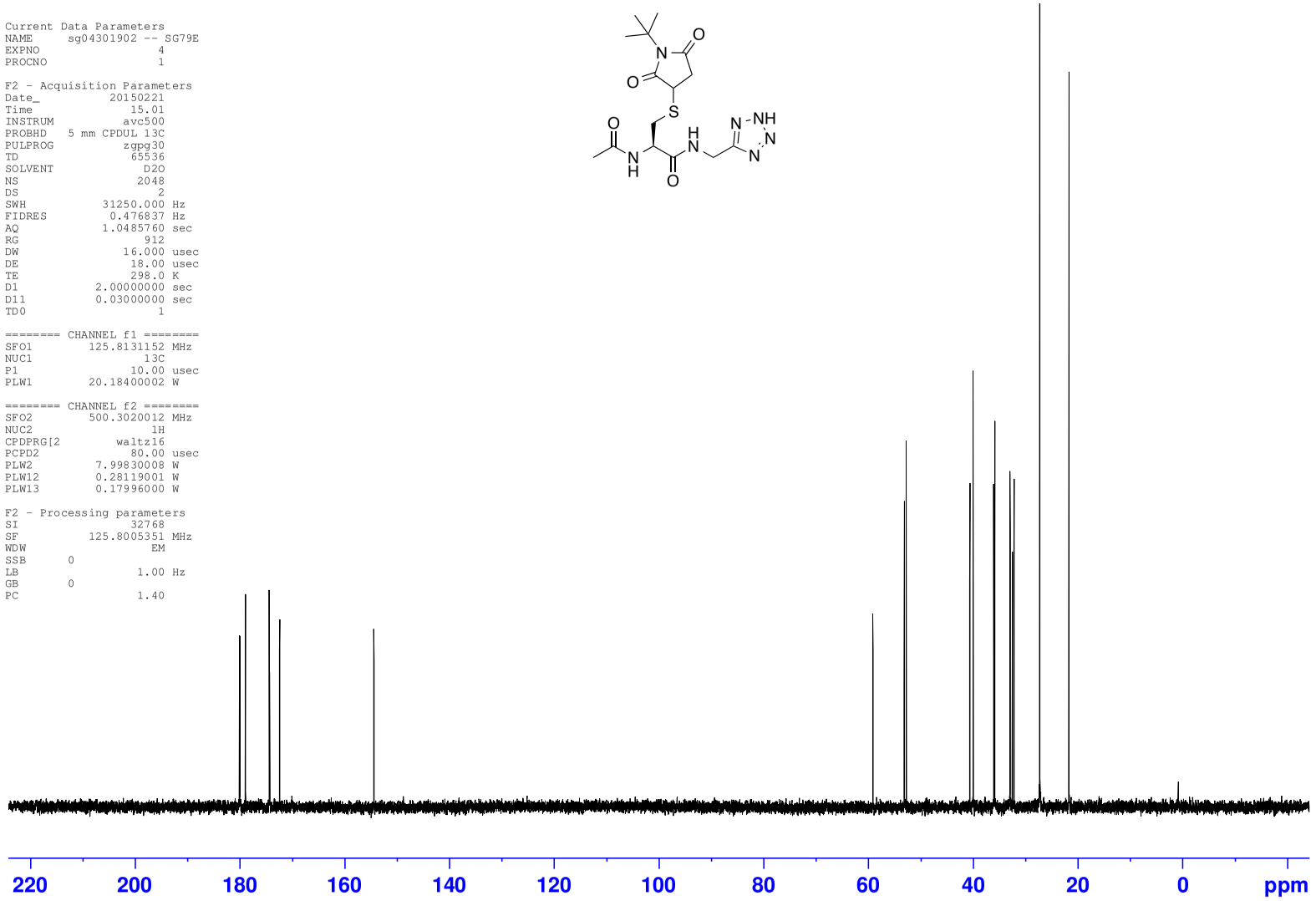


^{1}H NMR - *S*-(*N*-*tert*-Butylsuccinimido) (*R*)-*N*-((2*H*-tetrazol-5-yl)methyl)-2-acetamido-3-mercaptopropanamide (147)

```
Current Data Parameters
NAME      sg04301902 -- SG79E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20150221
Time              12.53
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG            zg30
TD                65536
SOLVENT             D2O
NS                   16
DS                    4
SWH           10330.578 Hz
FIDRES         0.157632 Hz
AQ            3.1719425 sec
RG                 3.56
DW               48.400 usec
DE               10.00 usec
TE               298.0 K
D1           1.00000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        500.3030896 MHz
NUC1                1H
P1                15.00 usec
PLW1         7.99830008 W

F2 - Processing parameters
SI                 65536
SF          500.2999570 MHz
WDW                  EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```

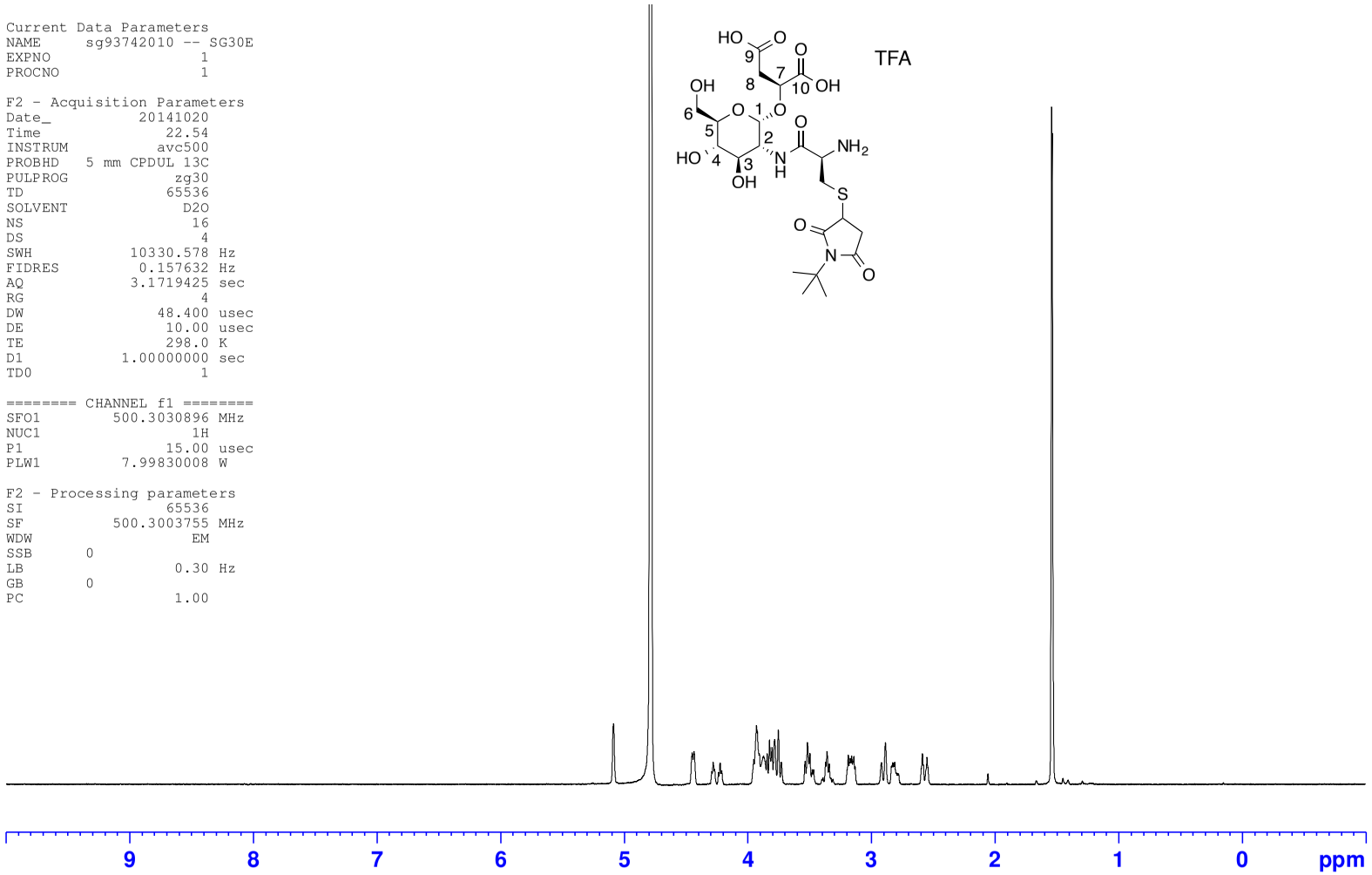


^{13}C NMR - *S*-(*N*-*tert*-Butylsuccinimido) (*R*)-*N*-((2*H*-tetrazol-5-yl)methyl)-2-acetamido-3-mercaptopropanamide (147)

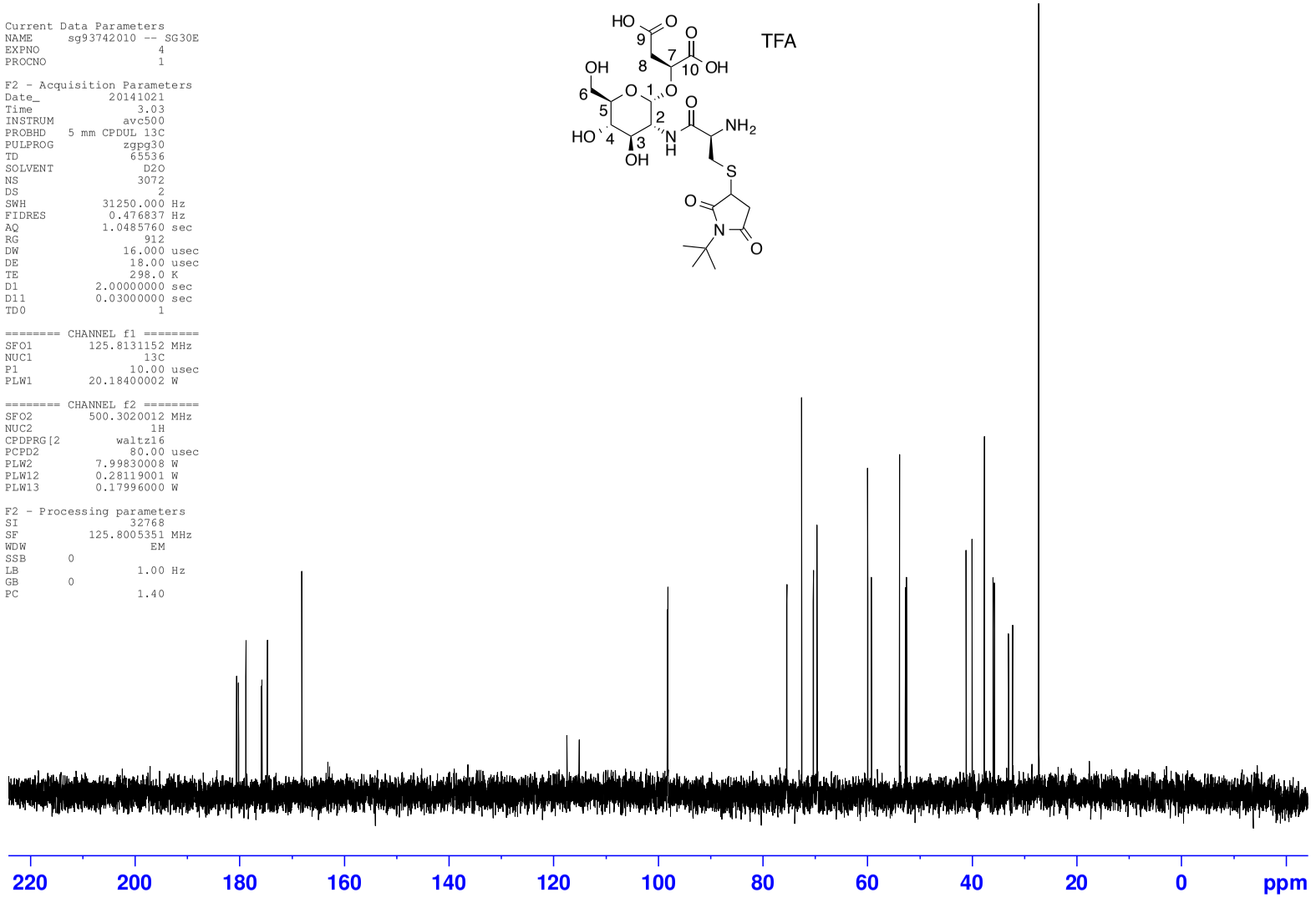


^{1}H NMR - *S*-(*N*-*tert*-Butylsuccinimido) bacillithiol 2,2,2-trifluoroacetate (155)

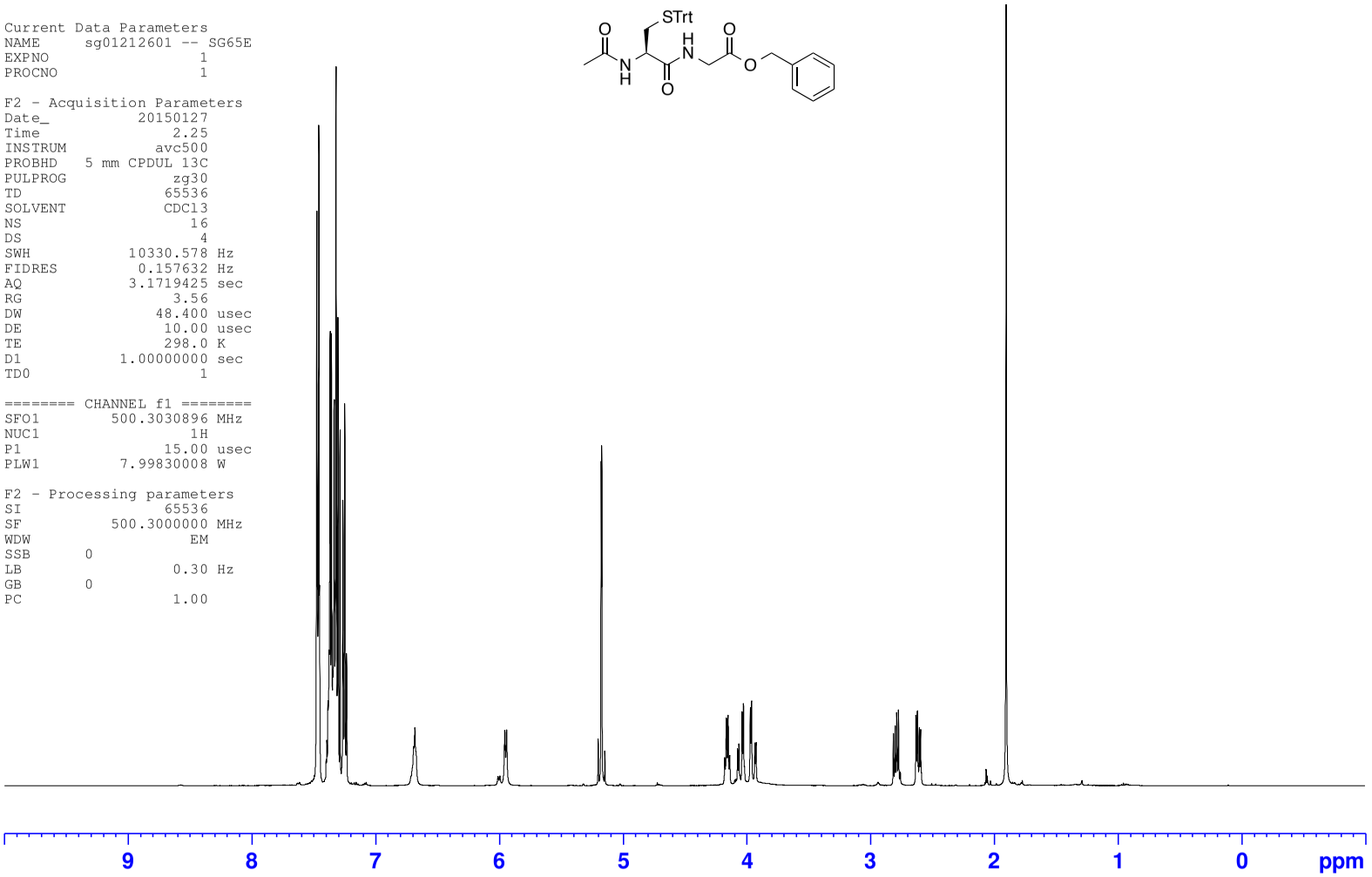
```
Current Data Parameters
NAME      sg93742010 -- SG30E
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20141020
Time               22.54
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG             zg30
TD                 65536
SOLVENT              D2O
NS                    16
DS                     4
SWH            10330.578 Hz
FIDRES          0.157632 Hz
AQ             3.1719425 sec
RG                     4
DW                48.400 usec
DE                10.00 usec
TE                298.0 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1         500.3030896 MHz
NUC1                   1H
P1                 15.00 usec
PLW1          7.99830008 W

F2 - Processing parameters
SI                 65536
SF           500.3003755 MHz
WDW                   EM
SSB      0
LB                 0.30 Hz
GB       0
PC                 1.00
```



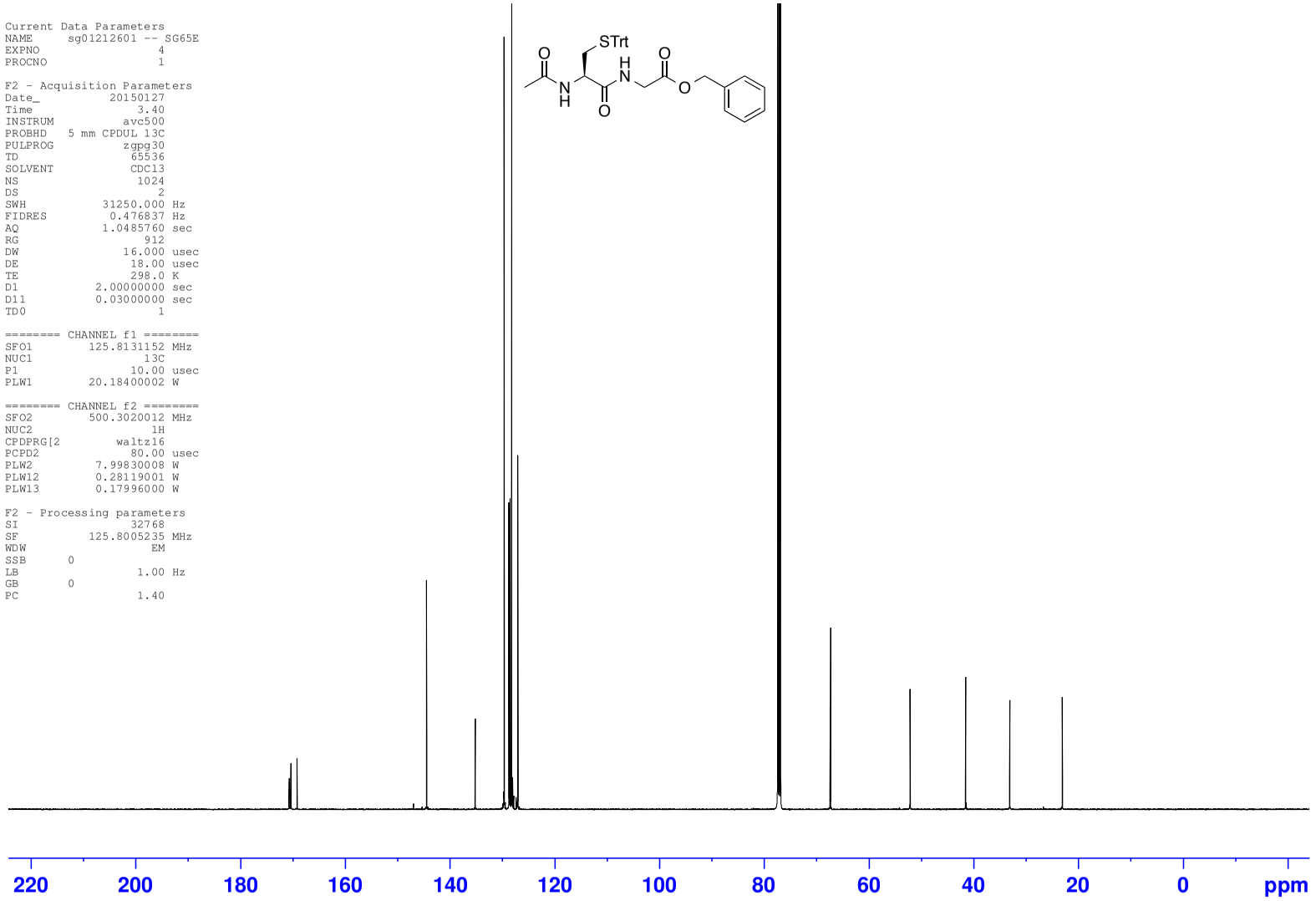
^{13}C NMR - *S*-(*N*-*tert*-Butylsuccinimido) bacillithiol 2,2,2-trifluoroacetate (155)



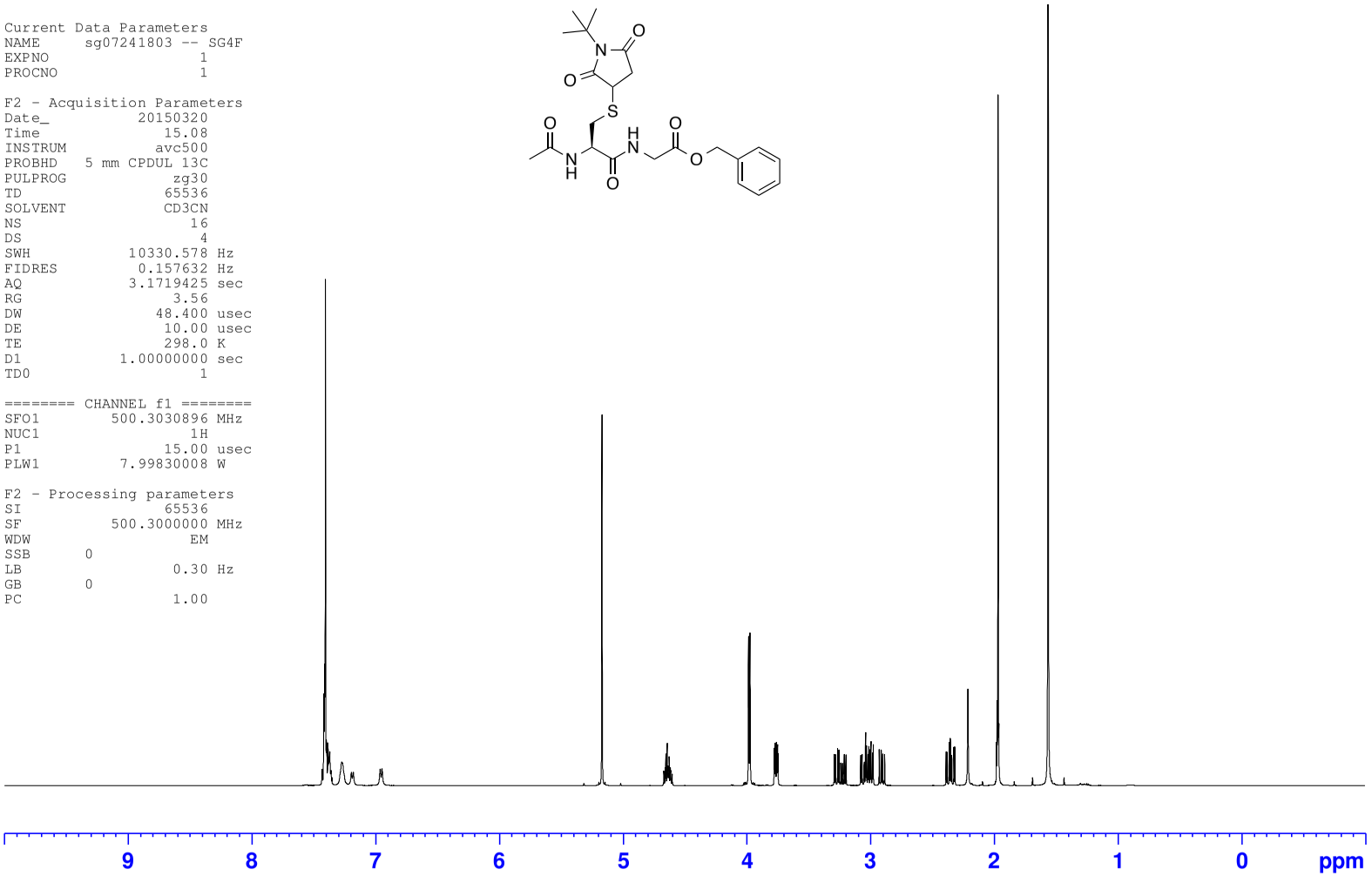
^{1}H NMR - *N*-Acetyl-*S*-trityl-L-cysteinylglycine benzyl ester (156)



^{13}C NMR - *N*-Acetyl-*S*-trityl-L-cysteinylglycine benzyl ester (156)



^{1}H NMR - *N*-Acetyl-*S*-*N*-*tert*-butylsuccinimido L-cysteinylglycine benzyl ester (158)



^{13}C NMR - *N*-Acetyl-*S*-*N*-*tert*-butylsuccinimido L-cysteinylglycine benzyl ester (158)