



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

A target-guided synthesis approach to the discovery of novel bivalent inhibitors of Glutathione Transferases



Alexandra J. Clipson

University of Edinburgh

Submitted for the degree of Doctor of Philosophy

6th June, 2011

Declaration

The work described in this thesis is entirely my own, except where I have either acknowledged help from a named person or given reference to a published source. Text taken from another source will be enclosed in quotation marks and a reference given. This thesis has not been submitted, in whole or in part, for any other degree.

Signature:

Date:

Acknowledgements

I would firstly like to thank my supervisor Dr. Michael Greaney for giving me the opportunity to work on this challenging project and for his guidance throughout my PhD.

I would also like to thank Dr. Dominic Campopiano, our collaborator on this project, for providing the resources and expertise for the expression and purification of the glutathione transferases. A special mention to Scott Baxter and Sam Capewell for their time and patience teaching a chemist the art of protein purification.

I thank cancer research UK for providing the funding for this research.

I owe a lot to the excellent groundwork carried out by Venugopal Bhat and Dr. Anne Caniard, my predecessors in the dynamic combinatorial work. Thanks to Venu for his unrivalled passion for HPLC and to Anne for the purification of several GST enzymes.

I would like to acknowledge the biophysical characterization facility in the centre for translational and chemical biology at the University of Edinburgh for the use of their facilities and in particular Dr. Liz Blackburn for her guidance in performing the ITC experiments and Dr. Janice Bramham for her assistance in setting up enzyme kinetics experiments on the multimode platereader.

I thank Dr. Juraj Bella and Dr. Marika DeCremoux for providing an excellent NMR service in the department. I also thank the EPSRC National Mass Spectrometry Service Centre at Swansea for their HRMS service.

This research could not have been completed without the help and support from all of my colleagues in the Greaney group, both past and present. I would like to say a special thank you to Naim Nazef for his LC/MS technical know-how.

To my friends and family, I cannot thank you enough for your unwavering support and for the belief you have shown in me.

Last, but certainly not least I would like to thank James. Quite simply I couldn't have done this without you. You kept a smile on my face even through the toughest times. Thank you.

Abbreviations

$[\alpha]_D$	specific rotation at 589 nm
A	Absorbance
Ac	Acetyl
AChE	Acetylcholinesterase
ADP	Adenosine diphosphate
ALD	Aldehyde
ALK	Alkyne
ASK1	Apoptosis signal-regulating kinase
AZ	Azide
BITC	Benzyl isothiocyanate
BSA	Bovine serum albumin
c	concentration
(b/h)CAII	(bovine/human) carbonic anhydrase II
CDNB	1-chloro-2,4-dinitrobenzene
DCC	Dynamic combinatorial chemistry
DCL	Dynamic combinatorial library
DCM	Dichloromethane
DIOS-MS	Desorption/ionisation on silicon mass spectrometry
DMAP	4-dimethylaminopyridine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNP GSH	<i>S</i> -(2,4-dinitrophenyl)glutathione
EA	Ethacrynic acid
EDC.HCl	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
eq	Equivalents
ESI-MS	Electrospray ionisation mass spectrometry
FT/IR	Fourier transform infrared spectroscopy
<i>g</i>	Relative centrifugal force

ΔG°	Change in free Gibbs energy
GSH	Glutathione
G-site	Glutathione binding pocket
GST	Glutathione transferase
ΔH	Enthalpy
H-bond	Hydrogen bond
HIV	Human immunodeficiency virus
HPLC	High performance liquid chromatography
H-site	Hydrophobic binding pocket
HZD	Hydrazide
HZN	Acyldiazide
IC ₅₀	half maximal inhibitory concentration
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IR	Infrared
ISCC	<i>In situ</i> click chemistry
ITC	Isothermal titration calorimetry
JNK	c-Jun NH ₂ -terminal kinase
K	Binding constant
K_d	Dissociation constant
K_i	Inhibition constant
K_M	Michaelis constant (concentration of substrate that leads to half-maximal velocity)
LB	Luria Bertani
LC	Liquid chromatography
LC/MS	liquid chromatography mass spectrometry
LC-MS-MS	liquid chromatography tandem mass spectrometry
LC/MS-SIM	liquid chromatography mass spectrometry with selected ion monitoring
MALDI-MS/MS	Matrix-assisted laser desorption/ionisation tandem mass spectrometry
MALDI-TOF MS	Matrix-assisted laser desorption/ionisation time of flight mass spectrometry

MAP	Mitogen-activated protein
MP	Melting point
n	reaction stoichiometry
NAD ⁺	Nicotinamide adenine dinucleotide
NMR	Nuclear magnetic resonance
NT	No triazoles
OBOC	One bead one compound
OD ₆₀₀	Optical density at wavelength = 600 nm
PEG	Polyethylene glycol
PNA	Peptide nucleic acid
R	universal gas constant
RNA	Ribonucleic acid
rpm	revolutions per minute
rt	Room temperature
ΔS	Entropy
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
S _N 2	Bimolecular nucleophilic substitution
S _N Ar	Nucleophilic aromatic substitution
SNP	Single-nucleotide polymorphism
SPR	Surface Plasmon resonance
T	Absolute temperature
TEMED	Tetramethylethylenediamine
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
TFA	Trifluoroacetic acid
TGS	Target guided synthesis
TGS buffer	Tris-glycine-SDS buffer
TLC	Thin layer chromatography
TZ	Triazole
UV	Ultraviolet
vol	Volume
wt %	percentage by weight
% w/v	weight/volume percentage

Nuclear Magnetic Resonance spectroscopic data

δ	chemical shift in ppm	m	multiplet
$^{13}\text{C}\{^1\text{H}\}$	proton decoupled ^{13}C NMR spectroscopy	MHz	Megahertz
d	doublet	NMR	Nuclear Magnetic Resonance
dd	doublet of doublets	ppm	parts per million
ddd	doublet of doublet of doublets	q	quartet
dq	doublet of quartets	s	singlet
^1H	Proton NMR spectroscopy	t	triplet
J	coupling constant	td	triplet of doublets
Hz	Hertz		

Mass spectroscopic data

amu	atomic mass units	HRMS	High resolution mass spectrometry
EI	electron impact	m/z	mass to charge ratio
ESI	electrospray ionisation	M^+	molecular ion

Units

$^{\circ}\text{C}$	degrees Celsius	m	metres
cal	calories	M	moles per decimetre cubed
cm^{-1}	wavenumbers	min	minute(s)
d	day(s)	mol	mole(s)
Da	Daltons	psi	pounds per square inch
g	gram(s)	sec	second(s)
h	hour(s)	V	volts
L	litre(s)		

Abstract

Target-guided synthesis is an approach to drug discovery that uses the biological target as a template to direct synthesis of its own best inhibitors from small molecule fragments. The process bridges the gap between chemical synthesis of drug candidates and their biological binding assay, merging the two operations into a single process whereby the active site or a binding pocket within the structure of the biological target directly controls the assembly of the best inhibitor *in situ*. Two different approaches to target-guided synthesis, the thermodynamic approach, making use of reversible reactions, and the kinetic approach, which uses an irreversible reaction, have been employed to discover novel, isoform selective inhibitors of the glutathione transferase (GST) enzyme family – possible drug targets in cancer and parasitic disease treatments.

The thermodynamic approach described in this thesis uses the aniline-catalysed reversible acyl hydrazone formation reaction to create a dynamic covalent library of bivalent ligands designed to bind the dimeric structure of GST. In the presence of GST one of the bivalent ligands was selectively amplified at the expense of the other library members. This ligand was shown, *via* biological assays, to be a specific inhibitor for one isoform of GST, the mu isoform mGSTM1-1.

A kinetic approach has also been investigated as a way to identify novel bivalent GST inhibitors utilising the Huisgen 1, 3 dipolar cycloaddition reaction. An azide and alkyne fragment library was designed to bind across the dimeric GST structure. The inhibitor structures are therefore bivalent, containing two anchoring fragments known to bind to the GST active site, linked by a triazolepeptide spacer. The triazole provides the click chemistry disconnection, enabling rapid *in situ* screening of candidate alkyne and azide fragments for inhibitor discovery. Whilst the *in situ* reaction with GST yielded inconclusive results, a number of the triazole products were found to have low nanomolar inhibitory activity towards GST.

Table of Contents

Declaration.....	i
Acknowledgements.....	ii
Abbreviations.....	iii
Abstract.....	vii
Table of Contents.....	viii
Chapter 1 - Introduction.....	1
1.1 Target-guided synthesis.....	1
1.2 The first example of target-guided synthesis.....	3
1.3 Dynamic combinatorial chemistry.....	4
1.3.1 Recent developments in DCC –	
(i) DCC for challenging targets.....	7
1.3.2 Recent developments in DCC –	
(ii) Compound detection in large DCLs.....	9
1.3.3 Recent developments in DCC – (iii) Novel DCC reactions.....	11
1.4 Kinetic target-guided synthesis.....	13
1.5 <i>In situ</i> click chemistry.....	17
1.5.1 Cucurbituril.....	19
1.5.2 Acetylcholinesterase.....	19
1.5.3 Carbonic anhydrase II.....	22
1.5.4 HIV-1 protease.....	23
1.5.5 DNA binding groove.....	24
1.5.6 Chitinase.....	25
1.5.7 Protein-capture agents for carbonic anhydrase II.....	26
1.5.8 Myelin-associated glycoprotein –	
combinatorial approach with NMR spectroscopy.....	27

1.5.9 Mycobacterial monooxygenase EthR.....	28
1.6 Summary – DCC versus kinetic TGS.....	29
1.7 References.....	32

Chapter 2: A dynamic combinatorial chemistry approach to identify novel bivalent

inhibitors of glutathione transferases	39
2.1 Glutathione transferases	39
2.2 Design and synthesis of initial dynamic combinatorial library	49
2.2.1 Optimisation and identification of the initial dynamic combinatorial library.....	51
2.2.2 Isoform specificity in the initial dynamic combinatorial library	58
2.2.3 Synthesis of DCL01 products	62
2.2.4 Binding data for DCL01 products.....	62
2.2.5 Controls for protein-templated DCL01	64
2.2.6 Further binding studies for DCL01 products - competition assay and isothermal titration calorimetry.....	69
2.3 Problems with HZN01-02-01	73
2.4 Increasing complexity within the dynamic combinatorial library	77
2.4.1 Exploring isoform specificity with DCL02	81
2.4.2 Control experiments for DCL02	84
2.4.3 Synthesis and binding data for homodimers from DCL02.....	87
2.5 Conclusion and future work	88
2.6 References.....	90

Chapter 3: An <i>in situ</i> click chemistry approach to identify novel bivalent inhibitors of glutathione transferases	93
3.1 Introduction	93
3.2 Designing the <i>in situ</i> click chemistry library – a triazole as a peptide mimic	94
3.2.1 Synthesis of the azide linker library	98
3.2.2 Design and synthesis of a small alkyne library	99
3.3 The initial <i>in situ</i> click chemistry reaction	100
3.3.1 Optimising the ISCC reaction - the volume of DMSO.	105
3.3.2 Optimising the ISCC reaction – the time of the reaction	106
3.3.3 Optimising the ISCC reaction – fragment concentration	109
3.3.4 Optimising the ISCC reaction – enzyme concentration.....	112
3.4 The optimised ISCC reaction for AZ01 with ALK03 and ALK04	113
3.4.1 Further controls for the ISCC reaction between AZ01 and ALK03 or ALK04	115
3.4.2 Synthesis and binding studies of the triazole products observed in the ISCC reaction	118
3.5 Increasing the alkyne library	120
3.5.1 ISCC reaction with AZ01 and the 20-member alkyne library	122
3.6 Investigating the sensitivity of LC/MS-SIM.....	125
3.7 Binding data of triazole TZ18a-01, a triazole not formed in the ISCC reaction	127
3.8 <i>In situ</i> click chemistry screen of TZ03a-01 with alkyne library	128
3.9 ISCC screen of TZ04a-01 with alkyne library using mGSTM1-1	133

3.9.1 ISCC screen of AZ04a-01 with alkyne ‘hits’ using mGSTM1-1 – further controls	136
3.10 Summary of the ISCC reaction between TZ04a-01 and alkyne ‘hits’ with mGSTM1-1 and comparison with the binding data for the triazole products	140
3.11 Summary of the ISCC reaction between TZ04a-01 and alkyne ‘hits’ with hGSTP1-1 and comparison with the binding data for the product triazoles	144
3.12 Implications for ISCC as an approach for TGS	146
3.13 Inconsistencies in the ISCC reaction	149
3.14 Implications for multivalent drug design	150
3.15 Conclusions and future work	151
3.16 References	153
 Chapter 4 Conclusions.....	 157
 Chapter 5 Experimental Data	 161
5.1 Product synthesis and characterisation	161
5.1.1 General methods	161
5.1.2 Synthesis of compounds described in Chapter 2.....	161
5.1.3 Synthesis of compounds described in Chapter 3.....	171
5.2 Target-guided synthesis reactions	200
5.2.1 Dynamic combinatorial chemistry reactions.....	200
5.2.2 <i>In situ</i> click chemistry reactions	205
5.3 GST Binding assays	211
5.3.1 CDNB binding assay – to determine K_M and V_{MAX}	211
5.3.2 CDNB binding assay – to determine IC_{50} values	212
5.3.3 CDNB competition assay – to determine K_i values	214

5.3.4 Isothermal titration calorimetry	214
5.4 Expression and purification of GST	215
5.4.1 Expression and purification of hGSTP1-1	215
5.4.2 Polyacrylamide Gel Electrophoresis (PAGE)	216
5.4.3 Expression and purification of mGSTM1-1, SjGST and mGSTA4-4	217
5.5 References	217
Appendix	A1
6.1 Identification of peaks observed in DCL01 and DCL02.....	A1
6.1.1 Identification of peaks observed in DCL01.....	A1
6.1.2 Identification of peaks observed in DCL02.....	A8
6.2 LC/MS-SIM traces from <i>in situ</i> click chemistry reactions..	A14
6.2.1 <i>In situ</i> click reactions between AZ01 and ALK04	A14
6.2.2 <i>In situ</i> click reactions between TZ04a-01 and ALK04	A16
6.2.3 Reference samples for <i>in situ</i> click reactions between azide AZ01 and alkyne library.....	A19
6.2.4 Reference samples for <i>in situ</i> click reactions between azide TZ04a- 01 and alkyne library.....	A20
6.3 Results from <i>in situ</i> click reactions between TZ04a-01 and alkyne hits using hGSTP1-1	A23

Chapter 1 - Introduction

1.1 Target-guided synthesis

Despite a trade surplus of some £7 billion in the UK in 2009,^[1] the pharmaceutical industry has come under much criticism over the last fifteen years with the number of new drugs, defined as New Molecular Entities (NME), coming to market levelling off whilst R&D costs continue to rise dramatically.^[2] The development of a new drug will typically take between 12 and 15 years. Of 25000 compounds tested in the laboratory, only five will make it to market and only one will recoup its investment.^[1] The discrepancy between cost and productivity has resulted in a shift towards the development of new, more cost effective approaches to drug discovery to both complement and advance current methods. These techniques include the development of combinatorial chemistry and high-throughput screening,^[3, 4] improved computer-based virtual screening,^[5] use of solid-phase and microarray technology,^[6] and the development of novel strategies such as fragment-based drug design.^[7]

The traditional method of structure based design relies on a step-wise procedure whereby compound libraries are screened against the biological target to identify lead compounds. These leads will then be further functionalised to increase binding affinity. This approach can be very labour-intensive, with each new compound having to be synthesised, purified and then assayed for activity against the protein target, see Figure 1.

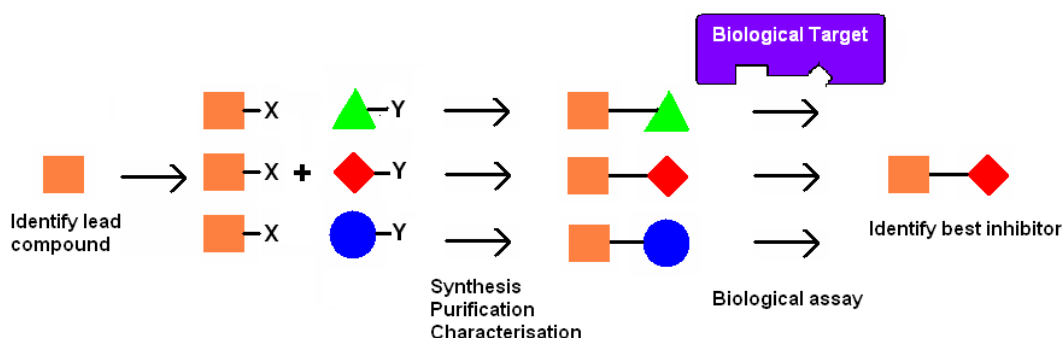


Figure 1: The traditional step-wise approach to drug discovery.

Clearly one solution to reduce time and cost would be the combination of synthesis with biological screening; a result achieved utilising target-guided synthesis (TGS). This methodology uses the biological target as a template to synthesise its own inhibitor from a mixture of small molecules, thereby reducing the number of synthetic steps required as well as removing the need for biological assaying of each possible product from the mixture.

Using the biological target as a template to synthesise its own inhibitor can be approached with two different methods; using dynamic combinatorial chemistry (DCC) or kinetically controlled TGS. DCC uses reversible reactions between small molecule fragments to make use of Le Chatelier's principle. "*Le Chatelier's principle states that when an external change is made to a system in equilibrium, the system will respond so as to oppose the change.*"^[8] In DCC a dynamic equilibrium is set up and on addition of the biological target; the most strongly binding product will be removed from the equilibrium mixture thereby shifting the equilibrium towards the synthesis of more of this product, resulting in its amplification, as shown in Figure 2.

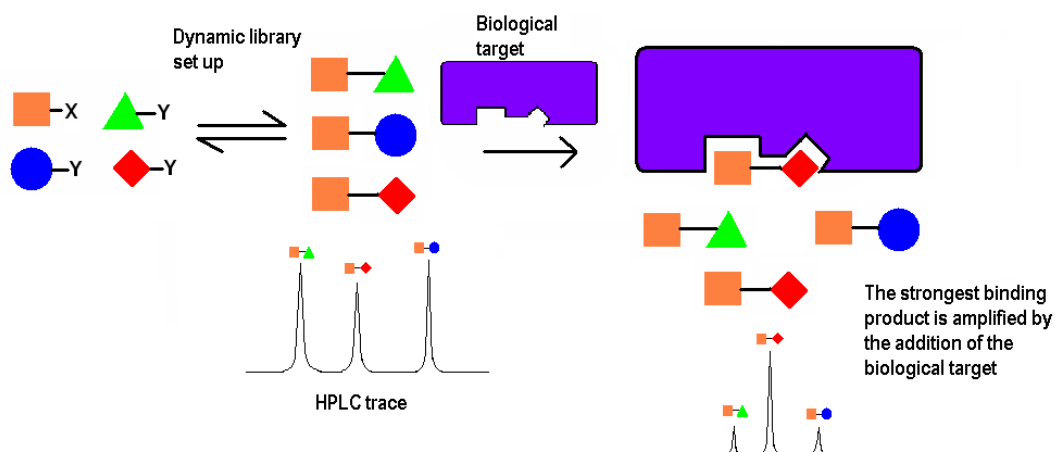


Figure 2: Cartoon outlining the principles of DCC.

Kinetic TGS on the other hand makes use of irreversible reactions with fragment libraries synthesised with functional handles attached. On addition of the biological template, those fragments with the best binding affinity will be held in close proximity by the biological target. This stabilises the transition state and promotes formation of the product from the two best-binding fragments. The resultant amplification of this product can then be detected from the reaction mixture, Figure 3.

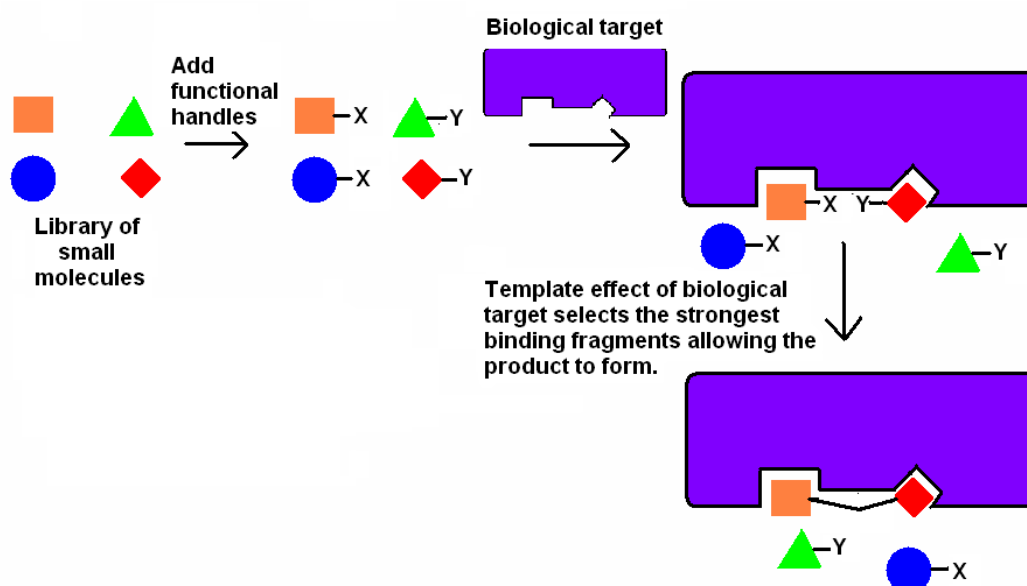


Figure 3: Cartoon outlining the principles of kinetic TGS.

1.2 The first example of target-guided synthesis

The first reported use of TGS was an example of kinetic TGS discovered by chance in 1969 by Chase and Tubbs.^[9] Whilst designing irreversible inhibitors for the enzyme carnitine acetyltransferase by synthesising bromoacetyl derivatives of the substrates CoA and carnitine to alkylate amino acid residues on or near the enzyme active site, it was noticed that in the presence of either bromoacetyl CoA and (-)-carnitine or CoA and bromoacetyl-(±)-carnitine a highly potent reversible inhibition was occurring. The authors determined that instead of alkylating nearby amino acids, the close proximity in which the two substrates were being held promoted the alkylation of the terminal thiol group of CoA to produce a biligand adduct of CoA-carnitine **1**, which displayed reversible inhibition, as demonstrated in Figure 4.

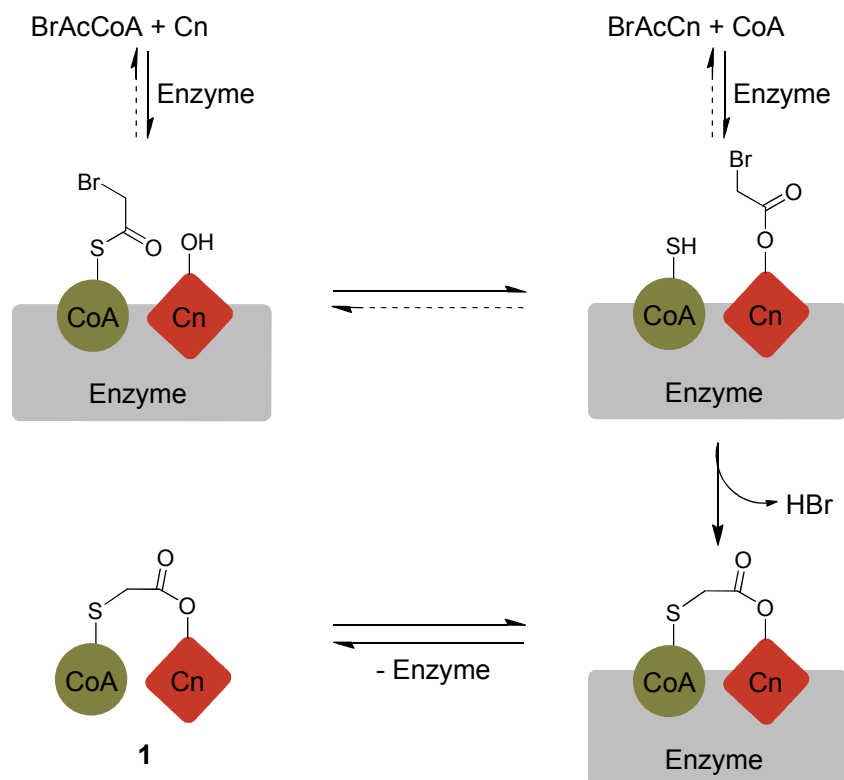


Figure 4: Reaction scheme outlining the proposed reversible inhibition of carnitine acetyltransferase by the templated reaction of CoA and bromoacetyl-(±)-carnitine.

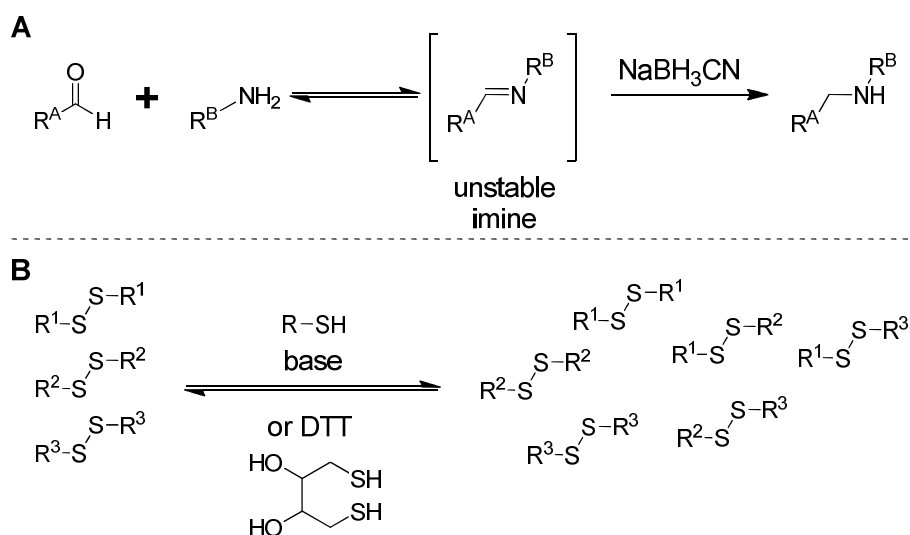
However, it was not until the 1980s that the idea of TGS as a method of drug discovery was explored.^[10] In 1986 Rideout described the *in situ* formation of cytotoxic hydrazones.^[11] A mixture of decanal and *N*-amino-*N'*-1-octylguanidine (AOG) was added to human erythrocytes, which were lysed within 80 minutes. Adding decanal or AOG alone resulted in no lysis after 420 minutes. The decanal and AOG self assembled to cytotoxic hydrazones *in situ*. Since then, the use of DCC as a drug discovery tool has really expanded, with kinetic TGS as a drug discovery strategy lagging somewhat behind. Both approaches to TGS are reviewed in this chapter.

1.3 Dynamic combinatorial chemistry

The concept of DCC is used in a wide variety of applications, not only the drug discovery role described above but also in the design of synthetic receptors, self replication systems and materials science.^[12] The use of DCC with biological

systems poses a particular set of requirements as the reversible reaction must be functional under the physiological conditions essential for the stability of the biological template. In 1997 Huc and Lehn proposed the reversible formation of imines from the reaction of an aldehyde and an amine, Scheme 1A.^[13] The reaction between two reactive species that are readily available proceeded at pH 6.0 and was successfully used to identify novel carbonic anhydrase II inhibitors. A drawback for this approach is that the imine products are unstable and therefore cannot be isolated or analysed directly. This necessitates a further *in situ* step of reduction of the imine to an amine. This additional kinetic step can complicate the results of the initial thermodynamic selectivity of the biological template leading to false positives and false negatives.

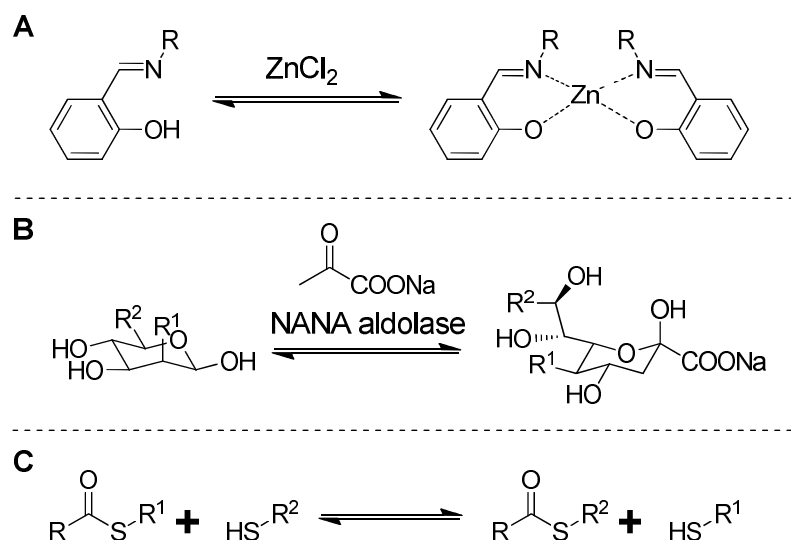
A second reversible reaction, introduced to the biologically-templated DCC arsenal by Ramström and Lehn, is disulfide bond formation, Scheme 1B.^[14] The suitability of the reaction was demonstrated in the discovery of new inhibitors for the lectin concanavalin A. The reaction between disulfides and thiols undergoes rapid interchange at pH ≥ 7 under aqueous conditions. The disulfide bond becomes stable at low pH (< 5) allowing the reversible library to be “locked” by a reduction in a pH. Both the imine- and disulfide-forming reactions have been successfully employed in the design and identification of many novel inhibitors.^[12, 15]



Scheme 1: A) reversible imine formation, followed by *in situ* reduction to an amine, B) reversible disulfide bond formation.

Several other reversible reactions that can be applied to

biologically-templated DCC have been reported in the literature. The Miller group took advantage of the reversible nature of transition metal complexation to create a dynamic combinatorial library (DCL) of salicylaldehydes forming coordination complexes with divalent zinc, as outlined in Scheme 2A.^[16] The reaction was shown to be reversible under mild aqueous conditions and the DCL was used to find the complex with the best binding affinity for a select double-stranded DNA sequence. The first example of a carbon-carbon bond forming reaction used in DCC was described by Flitsch *et al.*, who made use of the reversible nature of the enzyme-catalysed aldol reaction, Scheme 2B.^[17] A small DCL was generated using *N*-acetylneuraminic acid aldolase (NANA aldolase) as the aldol catalyst with one library member amplified in the presence of the biological template wheat germ agglutinin. Ramström and co-workers have demonstrated transthioesterification as a viable reversible reaction for use in biologically-templated DCC, Scheme 2C.^[18] A proof of concept study with acetylcholinesterase as the biological target showed the reaction proceeded at pH 7, at room temperature and with rapid equilibration of the DCL.



Scheme 2: **A)** reversible metal complexation reaction, **B)** reversible enzyme-catalysed aldol reaction, **C)** reversible transthioesterification reaction.

Recent developments in DCC can be divided into three areas; (i) application to challenging targets, (ii) development of techniques for monitoring large libraries and (iii) identification of novel reversible reactions suitable for biologically-templated DCC.

1.3.1 Recent developments in DCC – (i) DCC for challenging targets

The power of DCC as a tool for identifying novel enzyme inhibitors has clearly been demonstrated over the last twenty years.^[12, 15] The ability of DCC to be used for more challenging targets is now being explored. DNA and RNA targets are particularly problematic to medicinal chemists attempting to discover selective and highly specific ligands.^[19, 20] Balasubramanian and co-workers described the use of a disulfide-based DCL to identify ligands for specific DNA G-quadruplex targets.^[21] Miller *et al.* used a solid phase DCC approach, described in more detail below, to discover ligands for RNA structures involved in myotonic dystrophy type 1.^[22] In 2011 the Marchán group reported a DCC approach to the identification of novel ligands for the Tau exon 10 splicing regulatory element RNA, designing a DCL based on disulfide chemistry.^[23]

In 2010 Bradley, Diaz-Mochon and co-workers described a dynamic approach to the identification of single-nucleotide polymorphisms (SNPs).^[24] As shown in Figure 5, recognition of a single, unknown, nucleobase in a DNA strand was achieved starting with a peptide nucleic acid (PNA) strand. The PNA strand was complementary to the target DNA except in the target position, where the PNA contained only an amine or “blank”. Four aldehyde-modified nucleobases were then added which were capable of reversibly binding in the blank position of the PNA strand *via* an imine bond. The base pair corresponding to the unknown nucleotide was amplified in binding to the blank position. Reduction of the unstable imine bond followed by MALDI-TOF MS analysis led to the detection of the unknown nucleobase. The approach was successfully employed to identify SNPs associated with cystic fibrosis.

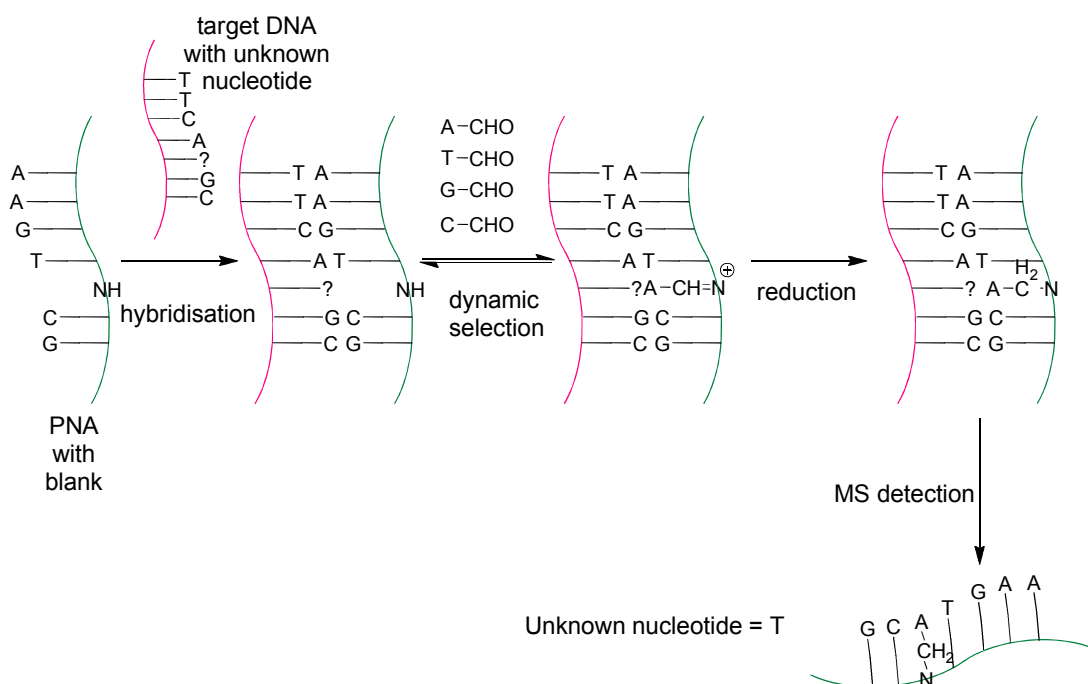


Figure 5: Identification of SNPs using reversible dynamic imine chemistry.

To face the challenge posed by many of nature's drug targets, a number of strategies have been described that couple the power of DCC with an additional drug discovery technique. Erlanson and co-workers combined DCC with a tethering approach to unearth new inhibitors for the aurora A kinase.^[25] One member of the DCL was designed to form a disulfide linkage with a cysteine side chain located in close proximity to the binding pocket in order to anchor the fragment. This fragment was decorated with a second thiol capable of participating in a disulfide DCL, which was used to probe a second binding pocket. A similar tethering approach is used by Schofield *et al.* in their discovery of novel metallo- β -lactamase inhibitors.^[26] Again, one library member was a dithiol, with one thiol designed to complex with one of the two zinc ions located in the active site of the enzyme and the other thiol reacting in a disulfide DCL. ESI-MS was used to directly measure any library members complexed with the enzyme.

Ciulli *et al.* employed a different approach to ligand design by creating a DCL based on a thiol-derivatised adenosine building block to probe an adenosine-binding enzyme, *M. tuberculosis* pantothenate synthetase, to identify new binding species.^[27] X-ray crystallography was then utilized to optimize the best binding component, identified in the DCC, to create a highly-potent inhibitor.

1.3.2 Recent developments in DCC – (ii) Compound detection in large DCLs

A limiting factor in the use of DCC is the size of the library that can be monitored. Currently, the vast majority of examples of DCC make use of small libraries, typically of ten to one hundred components, to allow the separation and observation of each library component by liquid chromatography. Work within the Otto group has focused on identifying the optimum library size for DCC. In a report describing the identification of ephedrine receptors, a library of at least 9000 components was set up using the disulfide exchange reaction between fragments with multiple sulfide groups.^[28] From this large library, receptors with strong binding interactions with ephedrine were identified from only two vials and two LC-MS-MS analyses. The Otto group has subsequently gone on to computationally simulate the outcome of a range of DCLs of varying size in an attempt to determine the optimum library size.^[29] The models showed that whilst the probability of detecting strongly binding library members drops off rapidly with an increase in library size, the probability of producing a good binding component increases at a much faster rate. This suggests that using larger DCLs is more advantageous.

In an attempt to avoid the necessity for monitoring the entire DCL, another approach sees one of the components of the DCL tethered to a polymer bead enabling simple separation of the binding ligands from the rest of the library. Miller *et al.* described an early example of this approach in their identification of DNA-binding compounds from a metal complex DCL using an immobilized DNA receptor as the biological template.^[16] The disulfide bond formation strategy, described by Ramström and Lehn, utilized a polymer-bound form of the biological target lectin concanavalin A to remove the best-binding ligands from the DCL to allow for easier identification.^[14] More recently, Liu and co-workers have reported a strategy of immobilizing the biological template onto resin in their identification of novel vancomycin analogues.^[30] The resin-bound peptide target is added to the DCL and the best ligand binds to the peptide on the resin and so it is easily separated from the rest of the DCL. The ligand is then eluted and identified by LC/MS, as described in Figure 6.

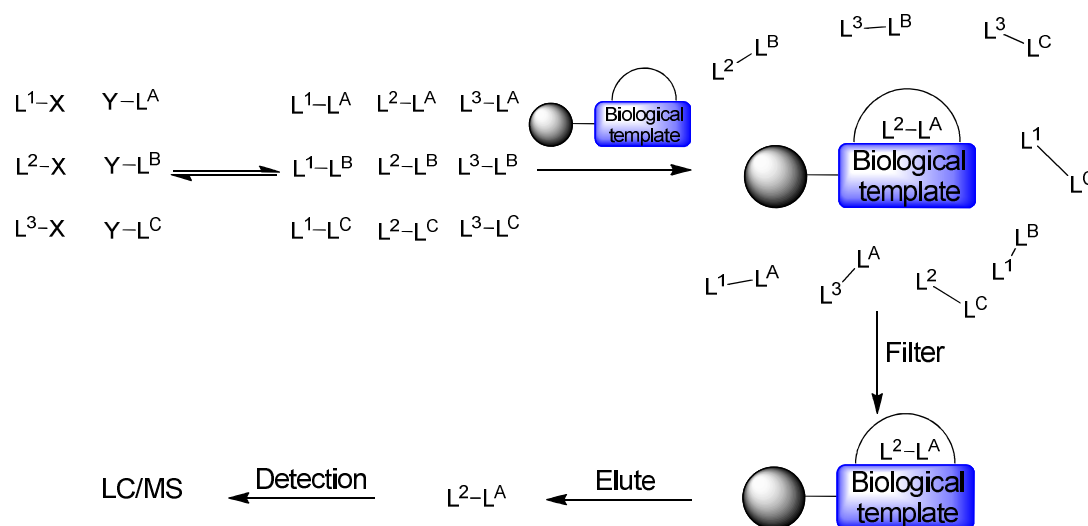


Figure 6: Diagram outlining DCC with a solid phase-bound biological template.

Miller *et al.* chose to immobilize one of the reactive fragments on to a polymer bead in order to identify novel ligands for RNA structures involved in myotonic dystrophy type 1.^[22] Disulfide exchange was utilized to set up a DCL of components from 150 polymer-bound monomers and 150 solution-phase monomers leading to a potential 11325-member resin-bound DCL. The RNA target, which was fluorescently-tagged, was added to the DCL and successful interactions were identified through the removal of fluorescent beads followed by mass spectral analysis, as outlined in Figure 7.

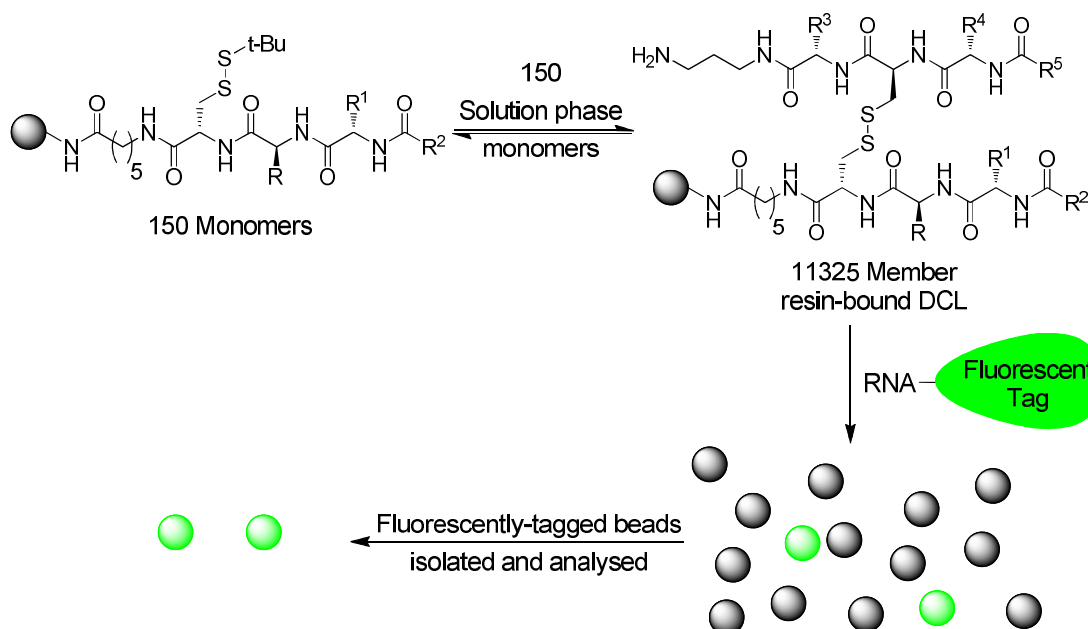


Figure 7: Diagram outlining DCC with a solid phase-bound fragment library member.

1.3.3 Recent developments in DCC – (iii) Novel DCC reactions

Work within the Greaney and Campopiano groups has focused on the development and application of novel reversible reactions to be used in biologically-templated DCC reactions. In 2005 they reported the conjugate addition of thiols to enones as a suitable reaction with equilibration of a small five-membered library occurring within one hour at pH 8.^[31] The methodology was used to identify novel glutathione transferase (GST) inhibitors based on the reaction between a thiol-containing tripeptide glutathione **2** and a series of analogues of ethacrynic acid, a known substrate of GSTs, Figure 8.^[32]

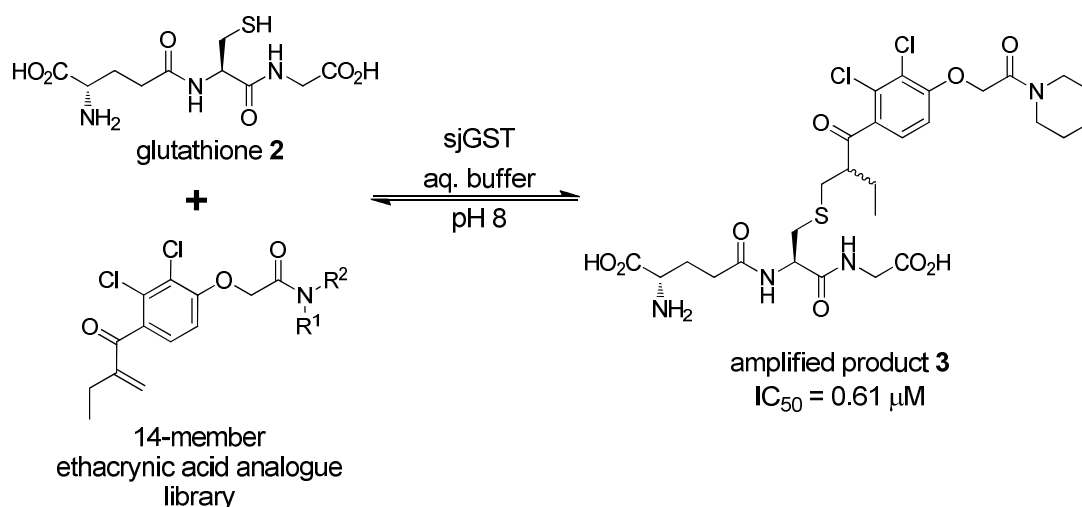
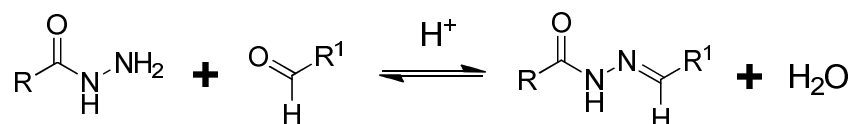


Figure 8: Application of the reversible conjugate addition of a thiol to an enone to the discovery of novel GST inhibitors.

In 2010, the Greaney and Campopiano group reported the successful application of the acyl hydrazone formation reaction in biologically-templated DCC, Scheme 3.^[33]



Scheme 3: Reversible formation of an acylhydrazone from the reaction between an aldehyde and a hydrazide.

Based on the original DCC reaction of imine-formation from an aldehyde and an amine, the acylhydrazone reaction between an aldehyde and hydrazide was developed by the Sanders group as, unlike the imine, the acylhydrazone product is

stable, so a secondary *in situ* step is not required. Many examples of DCC with this reaction exist in the literature,^[34-39] but as rapid equilibration only occurs at acidic pH (approx. pH 4) there are few reports of this reaction being used in conjunction with a biological template. Lehn reports the use of acylhydrazone formation to generate a library of ligands at pH 4. This library is subsequently screened in a standard binding assay to find potent inhibitors. This approach has been applied to discover ligands for acetylcholinesterase,^[40] the bifunctional enzyme HPr kinase/phosphatase from *Bacillus subtilis*,^[41] and the lectin concanavalin A.^[42] The only direct application of acylhydrazone formation with a biological template, prior to the work from Greaney and Campopiano *et al.*, came from Poulsen, who took advantage of the slow equilibration of acyl hydrazones at pH 7, taking one week, to demonstrate an acceleration in the formation of the best-binding library components in the presence of the biological template carbonic anhydrase II.^[43] In order to make the formation of acylhydrazones a reaction that could be compatible with biological templates, Greaney and Campopiano *et al.* took inspiration from work by Jencks in the 1960s, who applied nucleophilic catalysis to semicarbazone formation using aniline derivatives.^[44] Nucleophilic catalysis has recently been applied to hydrazone and oxime formation in peptide ligation systems by Dawson.^[45, 46] The Greaney and Campopiano groups demonstrated that a ten-member acylhydrazone library that took five days to equilibrate at pH 6.2, could equilibrate in 6 hours in the presence of aniline, which was acting as a nucleophilic catalyst.^[33] The compatibility of the aniline-catalysed acylhydrazone formation with a biological template was then demonstrated in the discovery of novel GST inhibitors. GST isoform specificity was demonstrated using a library based on an aldehyde analogue **ALD01** of 1-chloro-2,4-dinitrobenzene (CDNB), the universal substrate of GSTs, with a selection of hydrophobic hydrazides that were designed to bind into the hydrophobic binding pocket of GST, as highlighted in Figure 9. The presence of the aniline catalyst was shown to have no effect on the selection of the best binding inhibitors.

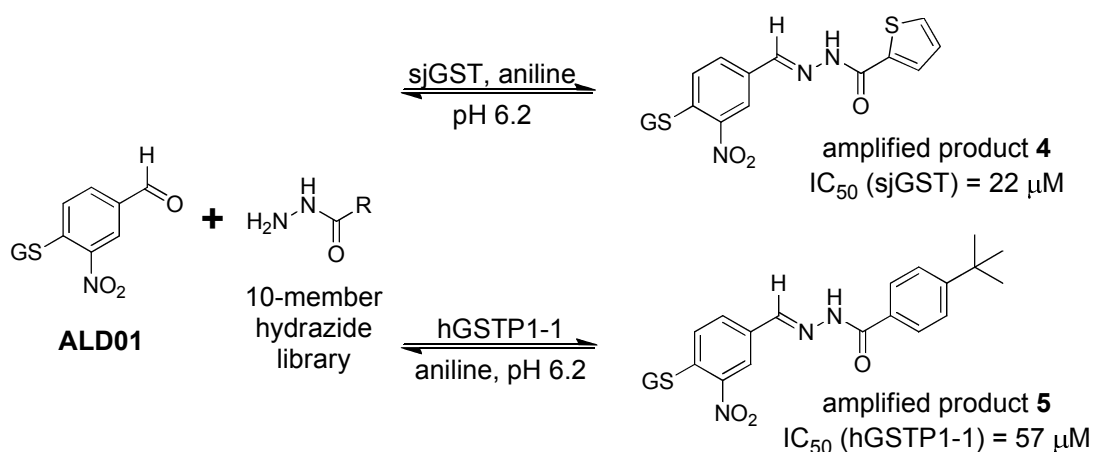


Figure 9: Application of reversible acylhydrazone formation to the discovery of novel GST inhibitors.

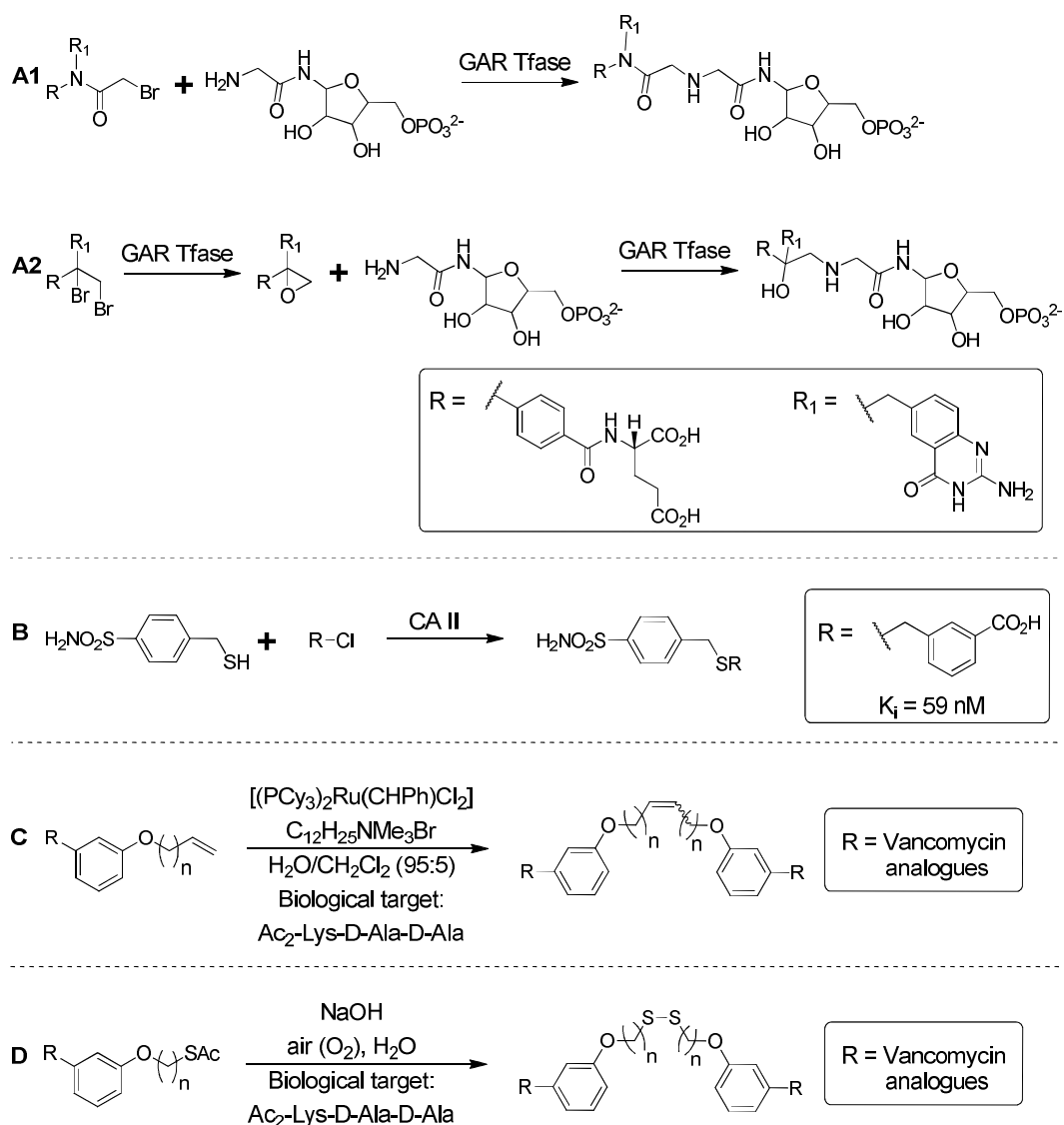
1.4 Kinetic target-guided synthesis

There are two main reasons for kinetic TGS being somewhat overlooked in favour of DCC approaches. Firstly, the limiting requirements placed on the type of reaction suitable for kinetic TGS. The reaction would ideally be very slow under standard conditions allowing any protein-templated reactions to be clearly observed. Furthermore, the reactive fragments need to be stable under conditions amenable to the presence of the biological target; typically aqueous buffer, pH ~ 7, temperature 25-37 °C. The second reason for the lack of kinetic TGS strategies is the limitations of the detection methods available to measure the products formed in the protein-templated reactions. This has been addressed with the introduction of liquid chromatography with mass spectrometry in selected ion monitoring (LC/MS-SIM) mode, which is discussed in greater detail later in this introduction.^[47]

Early examples of kinetic TGS in the literature, which have been well reviewed,^[48-50] include work by Benkovic and Bodger on multisubstrate inhibitors of the enzyme glycylamide ribonucleotide transformylase (GAR Tfase).^[51-53] Inhibitors were developed using an enzyme-templated alkylation of substrates with a folate-derived electrophile, Scheme 4 A1 and A2. Hu and Nguyen have described alkylation of thiols with α -chloroketones to generate inhibitors of carbonic anhydrase II, Scheme 4B.^[54]

A variation on kinetic TGS is catalyst-accelerated TGS. In this case a catalyst

is required for the reaction to proceed in addition to the protein template. This form of TGS is best exemplified in work described by Nicolau *et al.* in their development of novel vancomycin-based multidentate ligands for vancomycin-dependent bacteria using olefin metathesis in the presence of Grubb's catalyst, Scheme 4C.^[55, 56] In this work, Nicolau *et al.* also demonstrate the use of a disulfide bond-forming reaction as an alternative method for coupling the vancomycin analogues, Scheme 4D.^[55, 56] In this example the disulfide reaction is not set up under equilibrating conditions so the kinetic response is therefore observed.

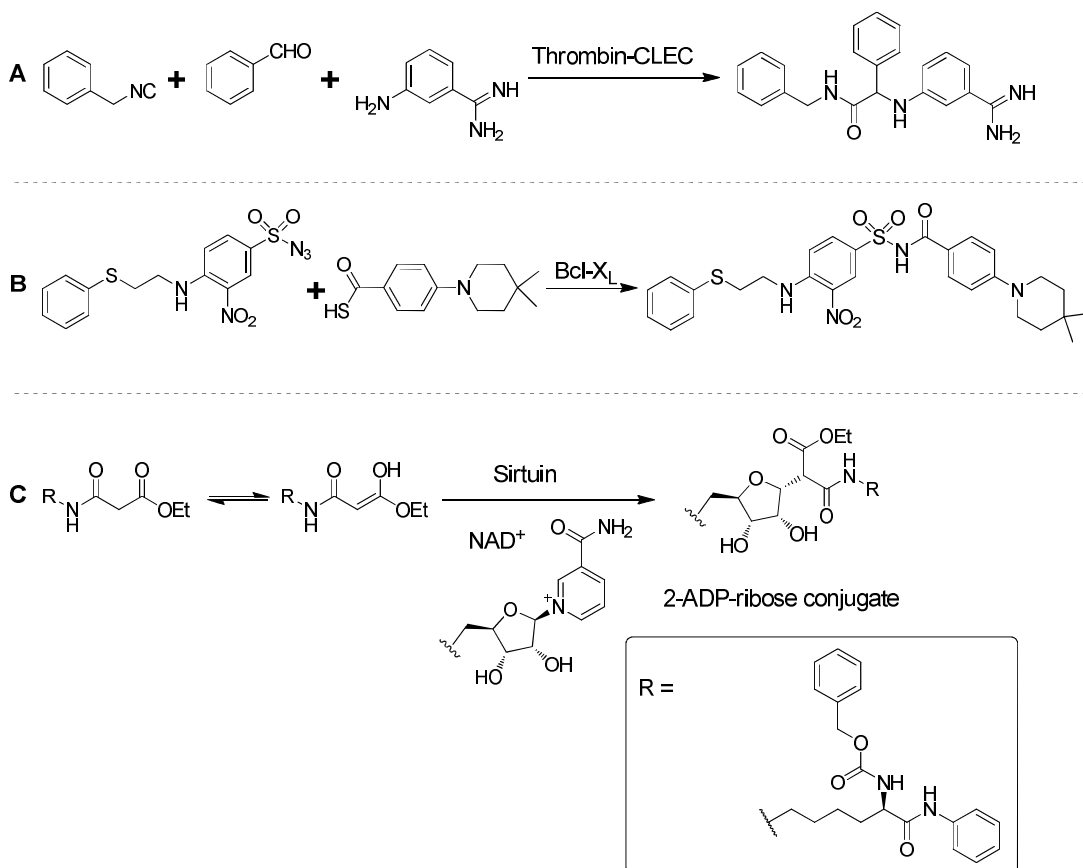


Scheme 4: Reactions used in kinetic TGS. **A1** and **A2**) nucleophilic substitution for synthesis of GAR Tfase inhibitors, **B**) thiol alkylation for synthesis of carbonic anhydrase II inhibitors, **C**) olefin metathesis for synthesis of vancomycin dimers, **D**) disulfide bond formation for synthesis of vancomycin dimers.

More recent reports of kinetic TGS include Weber's 2004 report outlining the use of the Ugi reaction for the development of thrombin-CLEC inhibitors.^[57] In this approach multiple different reactive fragments are combined in a complex equilibrium of reactions, with the final product removed from the equilibrium by a final irreversible step, which can be amplified in the presence of the biological template, Scheme 5A.

Manetsch and co-workers have demonstrated the reaction between a thio acid and sulfonyl azide as a viable tool in kinetic TGS, Scheme 5B.^[58] More importantly they use this reaction to identify novel inhibitors for the antiapoptotic protein Bcl-X_L at the BH3 peptide binding domain, in the first example of a kinetic TGS strategy for the development of inhibitors for protein-protein interactions. This is a key advance in the application of kinetic TGS as a drug discovery tool as protein-protein interactions are such an important target for the pharmaceutical chemist.^[59] As protein-protein interactions generally occur over a large surface area, this work proves that the well defined binding pockets associated with enzymes and protein-substrate interactions are not an essential requirement for a biological template employed in TGS.

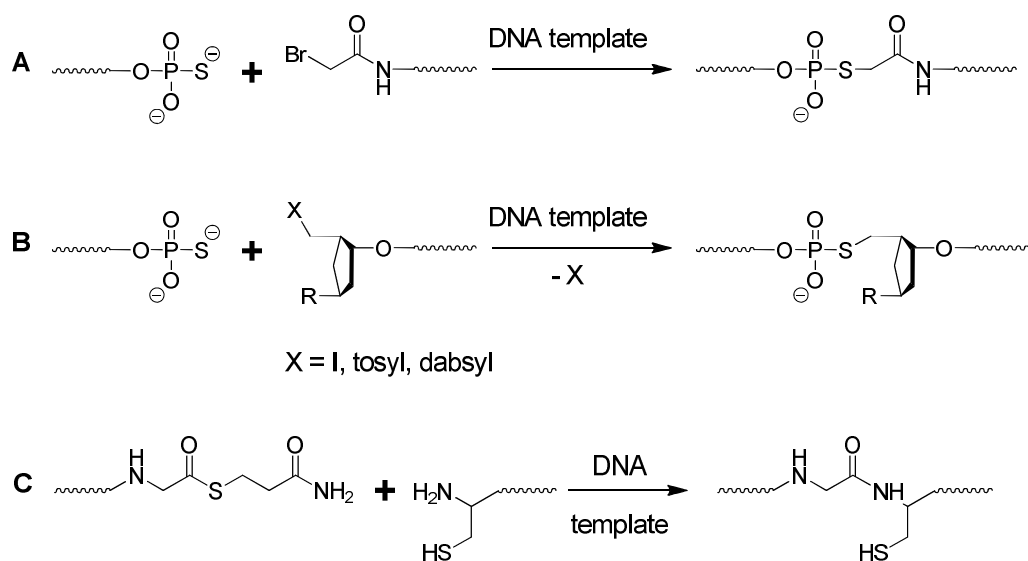
Small-molecule inhibitors of sirtuins were developed by taking advantage of the known catalytic mechanism of the enzyme. The role of sirtuins in the cell is to catalyse the deacetylation of acetylated lysine residues using NAD⁺ as a cofactor.^[60] The first step of the mechanism sees the oxygen of the acetylated lysine residue attack the NAD⁺.^[61] Based on this step, Suzuki and Miyata *et al.* designed a series of acetylated lysine analogues with a nucleophilic α -carbon, which combined *via* a carbon-carbon bond in the sirtuin active site to yield a stable substrate-ADP-ribose inhibitor, Scheme 5C.^[62]



Scheme 5: More reactions used in kinetic TGS. **A)** Ugi reaction for synthesis of thrombin-CLEC inhibitors, **B)** reaction between thio acid and sulfonyl azide for synthesis of Bcl-X_L inhibitors, **C)** carbon-carbon bond formation for synthesis of sirtuin inhibitors.

Single-strand DNA has been used to template a number of kinetically-controlled reactions designed to identify mutations within the template DNA strand. In these reactions two short sequences of DNA or PNA are coupled when a match is made for the template DNA. The first reported reaction to couple the two short strands of DNA was the reaction between a 3'-phosphorothioate and an oligonucleotide containing an electrophilic 5'-bromoacetyl group described by the Letsinger laboratory, Scheme 6A.^[63] Due to the difficulty in preparing the 5'-bromoacetyl oligonucleotides, the Letsinger group also explored the use of a 5'-tosyl modification as an alternative electrophilic leaving group, Scheme 6B.^[64] The Kool group have described autoligations between 3'-phosphorothioate oligonucleotides with 5'-iodo modified oligonucleotides in response to the observed degradation of the tosyl group, whilst in storage and in use.^[65] They later went on to describe the use of 5'-dimethylaminoazobenzenesulfonyl (dabsyl) modified

oligonucleotides in place of the 5'-iodo modification.^[66] As the dabsyl group also acts as a quencher of fluorescence, the use of this group allows the combination of chemical ligation and the switching on of fluorescence into one step. The Seitz group have reported a similar approach to couple short PNA sequences together using native chemical ligation, the reaction between a thioester and the thiol group of an N-terminal cysteine residue to yield a new amide bond, Scheme 6C.^[67] All of these approaches have been successfully used to identify mutations in specific DNA sequences, as the presence of a mutation results in a decrease in the ligation of the reacting DNA or PNA short sequences.

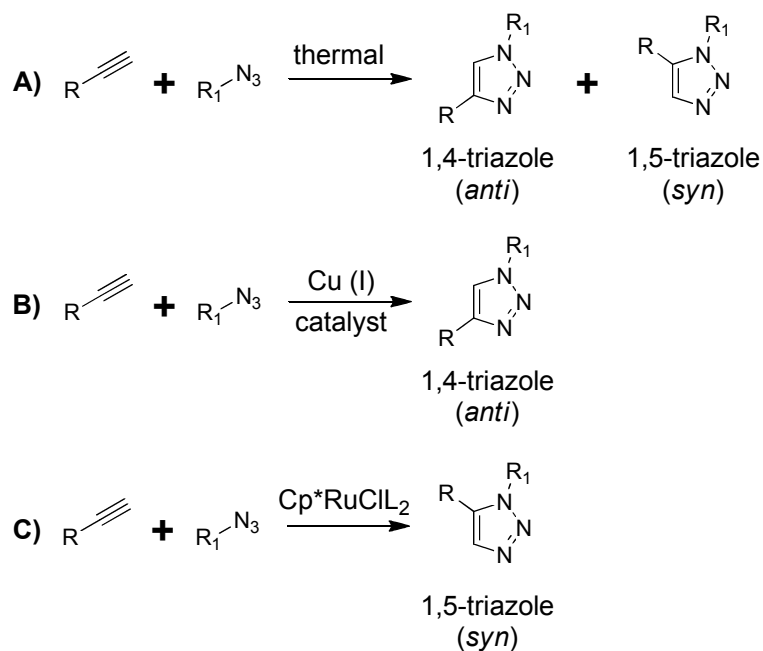


Scheme 6: DNA-templated reactions between short, single-strand DNA or PNA sequences. **A)** Reaction between 3'-phosphorothioate and 5'-bromoacetyl group. **B)** Reaction between 3'-phosphorothioate and 5'-iodo or tosyl or dabsyl leaving group. **C)** Reaction between thioester and N-terminal cysteine.

1.5 *In situ* click chemistry

There are a number of drawbacks with the TGS approaches described above. Some of the reactions used are between functional groups that are capable of unwanted side reactions with the protein itself, some of the reactions require the addition of an extra catalyst and other reactions have an observable background reaction which can complicate the analysis. One reaction which overcomes such problems, and so has expanded the scope of kinetic TGS in the last decade, is the Huisgen 1,3 dipolar cycloaddition reaction of an azide and an alkyne to produce a

1,2,3 triazole (Scheme 7). Both the azide and alkyne functional groups are biologically inert, insofar as they will not react with the protein target itself.



Scheme 7: General reaction scheme for the Huisgen 1,3 dipolar cycloaddition reaction. **A)** thermal reaction yields both regioisomers, **B)** copper (I)-catalysed reaction yields 1,4-triazole, **C)** ruthenium-catalysed reaction yields 1,5-triazole.

The Huisgen 1,3 dipolar cycloaddition reaction has become a tour de force in organic chemistry over the last 10 years, with the development of the regioselective copper (I)-catalysed transformation to yield purely the 1,4- (*anti*) triazole product, reported firstly by the Meldal group and later in a paper from the groups of Sharpless and Fokin in 2002 (Scheme 7B).^[68, 69] The reaction has been coined a “click reaction”, due to its high efficiency as well as the absence of unwanted side products. Click chemistry’s uses in chemistry today are numerous and far reaching with applications in organic synthesis, bioorganic chemistry and materials and polymer science.^[70-72] A thorough review of the applications of click chemistry has been published in 2010 in a themed issue by Chemical Society Reviews.^[73] The discovery of a ruthenium catalyst, described in Scheme 7C, which yields the complementary product of pure 1,5- (*syn*) triazole was reported in 2005, again by the groups of Sharpless and Fokin.^[74]

In the absence of a copper or ruthenium catalyst the reaction’s high activation barrier under standard conditions makes the click reaction an ideal candidate for a

kinetic TGS approach. No background reaction should be observed, and thus any product formed will be the result of two reacting fragments (azide and alkyne) being brought into close proximity, *via* binding to the biological target, allowing these fragments to react. The triazole products formed should, therefore, demonstrate a good binding affinity for the biological target themselves.

1.5.1 Cucurbituril

In 1983 Mock *et al.* used the nonadecacyclic cage compound, cucurbituril **6** to catalyse a triazole-forming cycloaddition reaction.^[75, 76] This was achieved as the cucurbituril encapsulated the reactants, bringing them into close proximity thus promoting reactivity, Figure 10. Not only was the reaction regioselective, with only 1,4 disubstituted (*anti*-) triazole formed in the cucurbituril cavity, but the rate of reaction was accelerated approximately 10^5 -fold over the non-templated reaction. This result provided promising evidence that biological targets could also be used as a template to promote cycloaddition reactions.

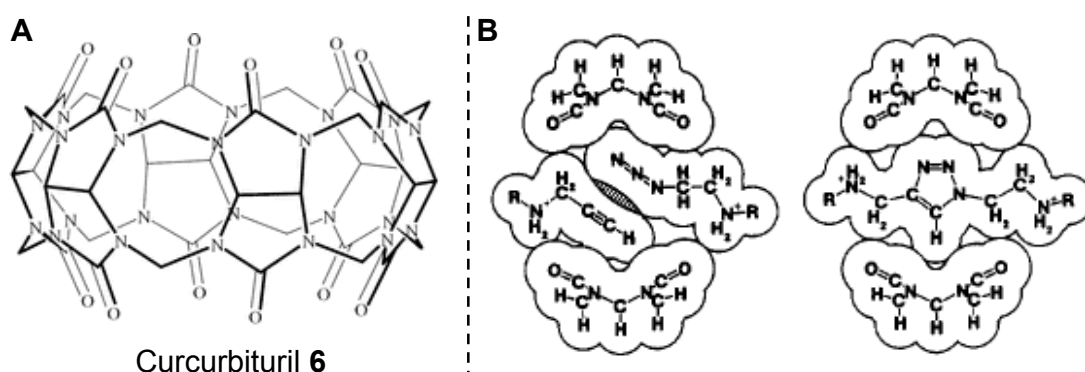


Figure 10: A) nonadecacyclic cage cucurbituril, B) proposed cross-sectional representation of the cucurbituril-templated click reaction. Outlines drawn to van der Waals radii.^[76]

1.5.2 Acetylcholinesterase

The first successfully reported example of *in situ* click chemistry (ISCC) came in 2002, published by the Sharpless group. Taking inspiration from the work described by Mock, it was proposed that a protein binding pocket could be used as the host or template to promote the click reaction for the best binding reactive fragments.^[77] Acetylcholinesterase (AChE), an enzyme which plays an important role in neurotransmitter hydrolysis in the central and peripheral nervous system,^[78]

provided an ideal model to develop ISCC. This was because AChE contains a secondary “peripheral” binding pocket in close proximity to the substrate binding site, Figure 11A. Proof of principle was described by using known inhibitors of each binding pocket (Figure 11B); the tacrine motif, **7** and **8**, binds to the active centre and the phenanthridinium motif, **9** and **10**, binds in the peripheral pocket.^[77] Each motif was functionalised with either azide or alkyne groups both on alkyl chains of varying length. Incubation of AChE with this mixture of azides and alkynes led to the identification of only one triazole product **11** from a possible 98 combinations. Products were identified using desorption/ionisation on silicon mass spectrometry (DIOS-MS) to detect triazole formation. The triazole product observed in the *in situ* reaction had a K_d of 99 fM and could therefore be considered as irreversibly bound to the enzyme during the course of the *in situ* experiment. As a result, the amount of triazole that can be produced is limited to being no more than the amount of AChE present in the reaction, as once the binding pocket has acted as a template it is then blocked by the triazole product. To ensure the identification of any successful click reactions the analysis technique was optimised, leading to the use of liquid chromatography with mass spectrometry in selected ion monitoring (LC/MS-SIM) mode. Products are identified based on their molecular weight and retention time. Use of LC/MS-SIM identified three further triazoles **12**, **13**, **14**, along with the previously identified triazole **11** from the tacrine/phenanthridinium library, now containing an additional fragment **9** ($n = 5$) with an alkyl chain length of five carbons.^[47] Reference samples of each triazole regioisomer observed in the ISCC reaction were independently synthesised in order to determine which regioisomer was formed in the presence of AChE. For all four triazoles it was found that the *syn*-isomer was exclusively produced. The binding affinity for AChE for both regioisomers was measured and, as expected, the *syn*-triazoles displayed much stronger binding affinity, for AChE compared with the *anti*-triazoles, typically 15-100 fold better, as displayed in Figure 11B.

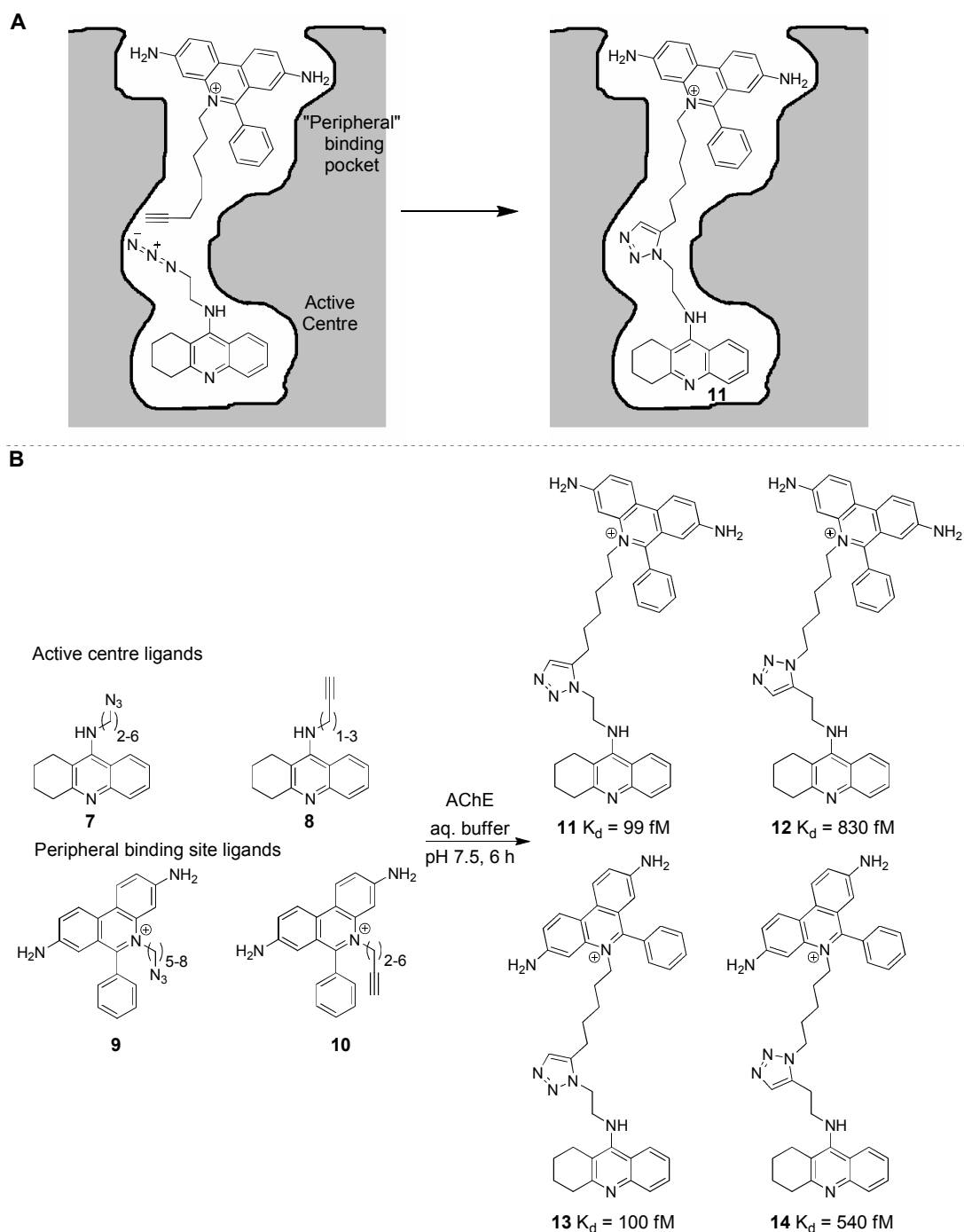


Figure 11: **A)** diagram outlining the template effect of the enzyme binding pocket on the tacrine and phenanthridinium fragments, **B)** ISCC library to find novel inhibitors of AChE.

A crystal structure of *syn*-**11** bound to mouse AChE was obtained and displayed the protein in a novel conformational form, never previously characterised.^[79] Furthermore the crystal structure revealed that the triazole ring formed in the ISCC reaction also plays an important role in the binding of **11** through hydrogen bonding and π -stacking interactions. The triazole ring was best located

approximately two methylene units from the tacrine fragment. Thus the azide-decorated tacrine fragment with a linker of two methylene groups, **15**, was screened with a novel library of acetylene compounds, not previously known to bind to AChE, to probe the peripheral binding site, Figure 12.^[80] Two new triazole products, **16** and **17**, were detected using LC/MS-SIM and furthermore these triazole inhibitors were selectively synthesised from a one pot mixture of several acetylene compounds. Referencing with authentic samples of triazole products determined that the reaction selectively produced the *syn*-triazole. Importantly, these triazole products, detected using ISCC, had low femtomolar binding affinity for AChE.

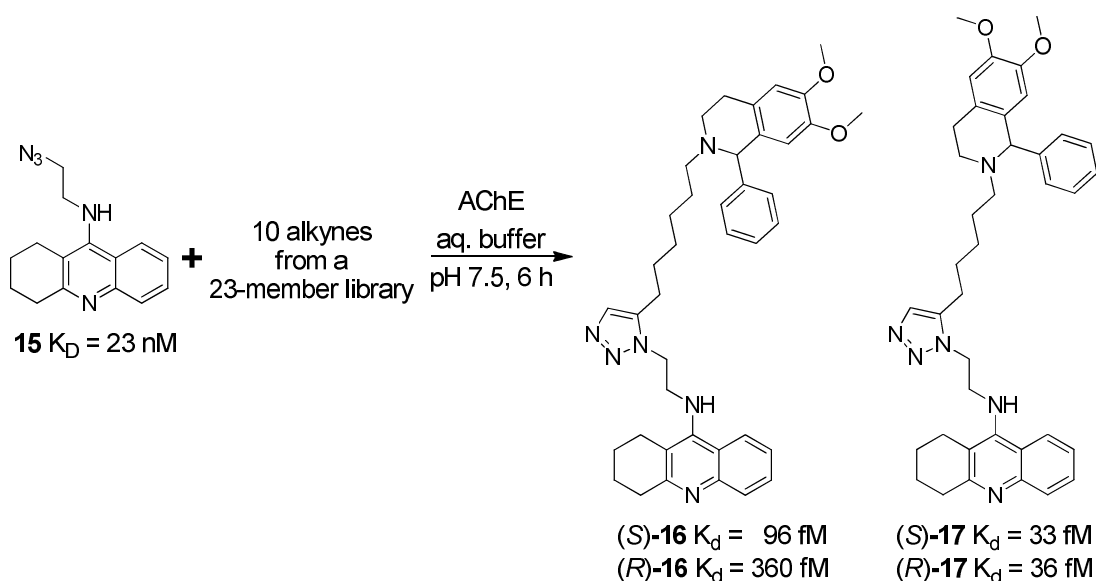


Figure 12: A second ISCC library to find novel inhibitors of AChE.

1.5.3 Carbonic anhydrase II

With the success of AChE, the approach was applied to developing novel inhibitors for carbonic anhydrase II (CAII).^[81] An acetylenic benzenesulfonamide **18**, based on known inhibitors with nanomolar binding affinity was selected as a basis to begin the screen of a library of azides, Figure 13. From a possible 24 triazole products that could have been made, 11 triazoles were detected. All of these compounds showed a greater binding affinity than that of the parent acetylene **18**. Several of the unsuccessful triazoles were also synthesised and their binding affinity for CAII was measured. Six out of ten compounds that were not formed in the presence of CAII still exhibited dissociation constants of less than 7.5 nM, which is

comparable with those triazoles that were a hit in the ISCC reaction. This highlights one drawback of the ISCC approach – the possibility of false negatives. If the alkyne and azide groups are not bound within close proximity in just the right geometry then the triazole will not form, even if the product itself has got good binding affinity.

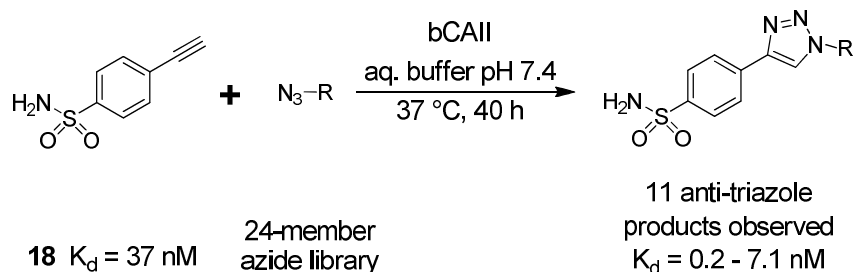


Figure 13: ISCC library to find novel inhibitors of CAII.

1.5.4 HIV-1 protease

In both the cases of AChE and carbonic anhydrase II, the starting azide and/or acetylene species already exhibited a high affinity for the target ($K_i < 37 \text{ nM}$).^[47, 77, 80, 81] ISCC has been applied to make HIV-1 protease inhibitors from fragments with a low binding affinity.^[82] Triazole **21** was a known nanomolar inhibitor of HIV-1 protease, however the azide **19** and alkyne **20** fragments only exhibited binding at micromolar quantities.

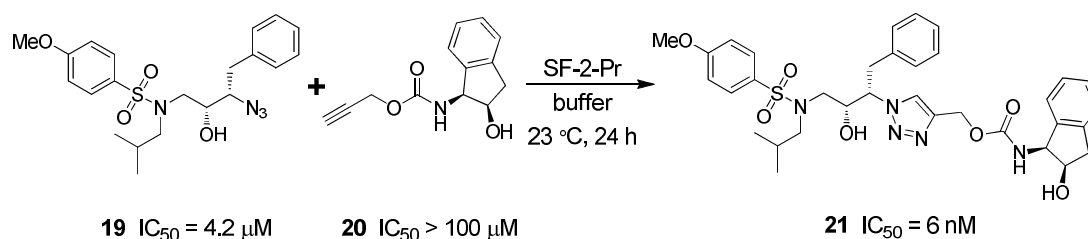


Figure 14: ISCC library to find novel inhibitors of HIV-1 protease.

Even with such weak binding, the triazole was not only formed *in situ*, but was also selectively produced from a mixture of fragments, Figure 14. As a result of the azide and acetylene fragments binding with a much weaker affinity, a higher concentration had to be used in the ISCC reaction. This resulted in the observation of trace amounts of triazole product forming in the background reaction. The triazole product **21** with the best binding affinity was the only one that was amplified in the presence of the enzyme template.

1.5.5 DNA binding groove

Poulin-Kerstien and Dervan have shown that DNA can also be used to template the click reaction.^[83] Hairpin polyamides are made from three aromatic amino acids, *N*-methylpyrrole, *N*-methylimidazole and *N*-methyl-3-hydroxypyrrole, and are capable of binding to specific sequences of DNA.^[84] In order to develop polyamide compounds capable of recognizing longer sequences of DNA an ISCC approach was utilized to couple together two polyamides that were each capable of recognizing a five base pair sequence.^[83] One polyamide was decorated with an alkyne group with two different linker lengths **22a-b** and the other polyamide was functionalized with an azide group, again *via* two different linker lengths **23a-b**. Three DNA templates were utilized with zero (A), one (B) or two (C) base pairs separating the polyamide binding sites. Only DNA template A, with no spacer between binding sites, resulted in formation of triazole for each polyamide pairing, as detected by HPLC and verified with MALDI-TOF MS, Figure 15. The length of the linker between the azide/alkyne functionality and the polyamide affected the rate of observed triazole formation. The introduction of base pair mismatches in the DNA template sequence resulted in a dramatic decrease in the rate of triazole formation.

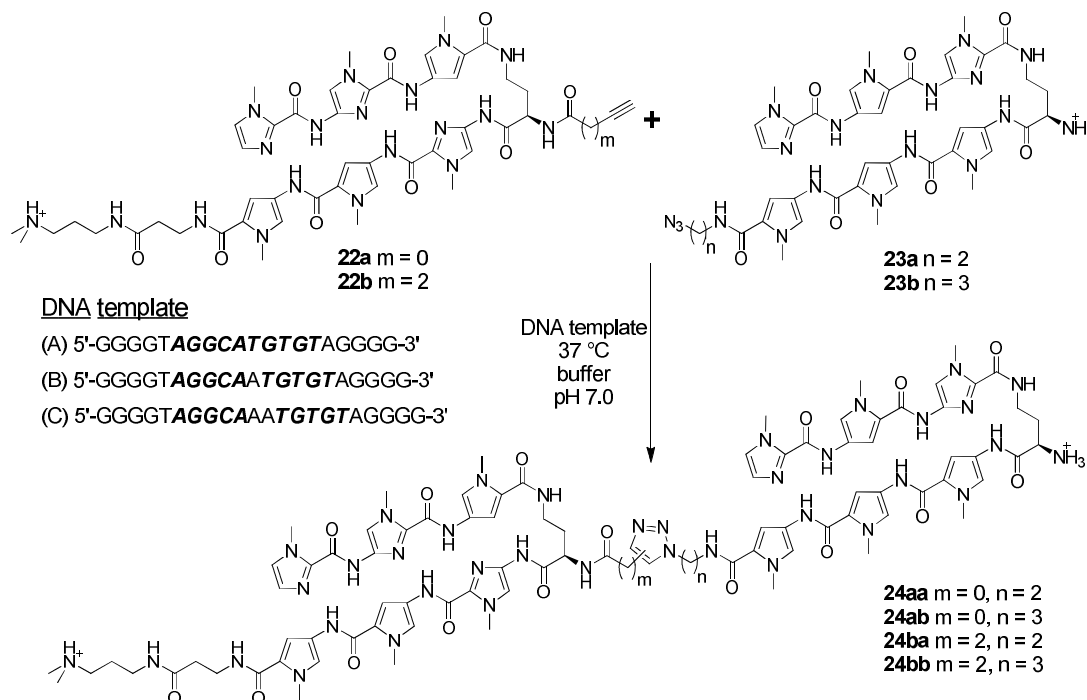


Figure 15: ISCC approach to designing novel linked hairpin polyamides capable of binding in the DNA binding groove.

1.5.6 Chitinase

In 2009, Omura *et al.* demonstrated the use of ISCC to identify a novel and potent chitinase inhibitor.^[85] Chitinases are the target for antifungal, insecticidal and antiparasitic agents.^[86] Azide fragment **26** was designed based on the natural cyclic peptide Argifin, **25**, which is known to demonstrate good chitinase inhibition.^[87, 88] This azide was subjected to an ISCC screen using a 71-member library of diverse alkynes in the presence of three chitinases, Figure 16.^[85] Only one alkyne, **27** resulted in the observation of a triazole product in the presence of the biological template, with the triazole product **28** detected using LC/MS-SIM. Both the *anti*- and *syn*- triazole regioisomers of this product were synthesised and their inhibitory activity measured for each chitinase. The *syn*-triazole **28** displayed a 30-fold stronger inhibition against *Serratia marcescens* Chitinase B, *SmChiB*, compared with the starting azide and a 300-fold increase in potency over the natural product argifin **25**. The initial triazole detection conditions were unable to separate the two triazole regioisomers so the reaction between azide **26** and alkyne **27** was repeated in the presence of *SmChiB* with improved LC conditions allowing the observation of both regioisomers. These results showed that whilst a small amount of the *anti*-isomer was observed in the background reaction as well as in the presence of the chitinase, the *syn*-isomer was only detected in the presence of chitinase, and the amount of product observed was dependent on the amount of chitinase present.

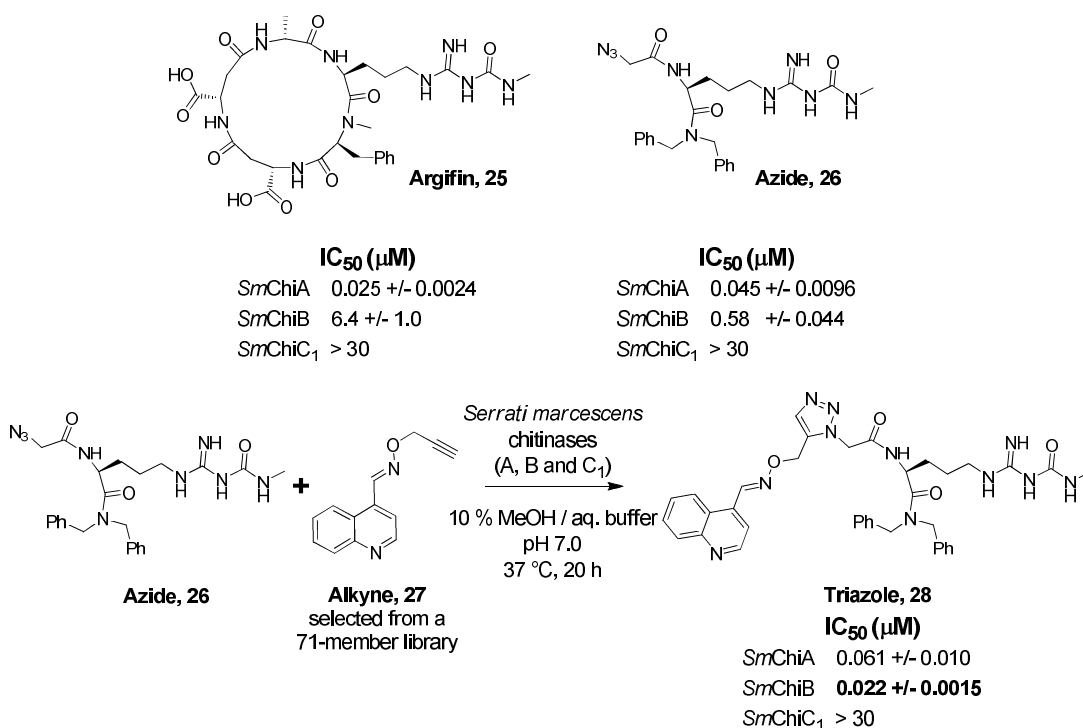


Figure 16: ISCC library based on argifin designed to identify novel chitinase inhibitors.

1.5.7 Protein-capture agents for carbonic anhydrase II

All of the previous examples of ISCC used a well defined biological template, either an enzyme binding pocket or DNA binding groove. Heath and co-workers have demonstrated that a well defined template is not a prerequisite for ISCC in their design of novel protein-capture agents towards bovine and human carbonic anhydrase II (b(h)CAII).^[89] The work also further supports the argument that the starting anchor fragment does not need to have a strong affinity for the biological target itself. A one bead one compound (OBOC) library of heptapeptides made from D-amino acids with a terminal D-propargylglycine were screened against fluorescently-tagged bCAII. Hit beads were identified by the increased fluorescence intensity and the heptapeptide sequence was determined by Edman degradation, as shown in Figure 17, 1st ligand screen. The peptide **29** that was identified was then used in a second screen of an OBOC library, this time decorated with an alkyne group **29a**, in the presence of bCAII. The heptapeptide **29** used as an anchor fragment had approximately 500 μM binding affinity for the protein. The second ligand screen identified triazole product **30**, which had a binding affinity of 3 μM for bCAII. The triazole **30a** was utilized as the anchor ligand in a third ligand screen of

the OBOC library, which identified triligand **31**. The bistriazole **31** exhibited 64 nM binding affinity for bCAII, however it was shown to have no effect on the activity of bCAII proving that it was not binding in the active site of the protein. Another significant achievement demonstrated in this work is that no knowledge of the protein structure was required to develop a high affinity-binding ligand.

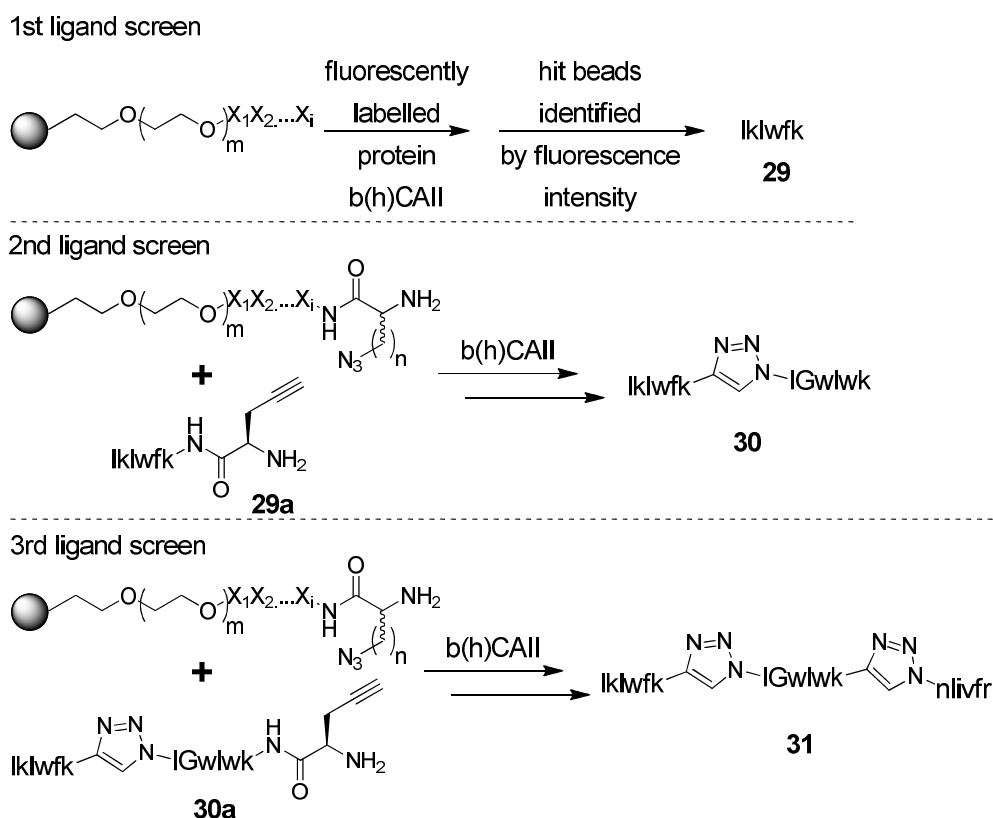


Figure 17: A solid phase ISCC approach to the discovery of protein capture agents for b(h)CAII.

In 2010, Lee, Heath and co-workers reported a further advance in the detection of protein-capture agents for CAII, by replacing the use of the Edman degradation to identify the peptide sequence of successful beads with a matrix-assisted laser desorption/ionisation tandem mass spectrometry (MALDI-MS/MS) sequencing approach.^[90] This new step alongside the ISCC screen resulted in the identification of two novel biligand capture agents for bCAII.

1.5.8 Myelin-associated glycoprotein – combinatorial approach with NMR spectroscopy

ISCC has been combined with NMR spectroscopy techniques to identify

novel high affinity ligands for myelin-associated glycoprotein (MAG), a protein with no available crystal structure.^[91] With little information about the 3D structure of the protein, NMR spectroscopy was used to identify small molecule fragments with weak binding affinity for the MAG protein. A sialic acid derivative based on ligand **32** without the added acetylene linker, which was a known ligand of MAG was tagged with a TEMPO spin label. This allowed fragments that bound to the protein in close proximity to the sialic acid ligand to be identified by NMR techniques monitoring paramagnetic relaxation effects. Fragment **33** was observed to be binding in close proximity to the sialic acid ligand. These fragments were then utilized in an ISCC approach to synthesise a novel biligand. The fragments were furnished with acetylene and azide linkers, as demonstrated in Figure 18, and in the presence of the MAG protein only one triazole product *syn*-**34** was observed, exhibiting a binding affinity of 2 μ M, which was a greater than 50-fold improvement in binding affinity compared with the starting fragments. The ability of ISCC to successfully identify novel triazole products from very weakly binding starting fragment was again demonstrated.

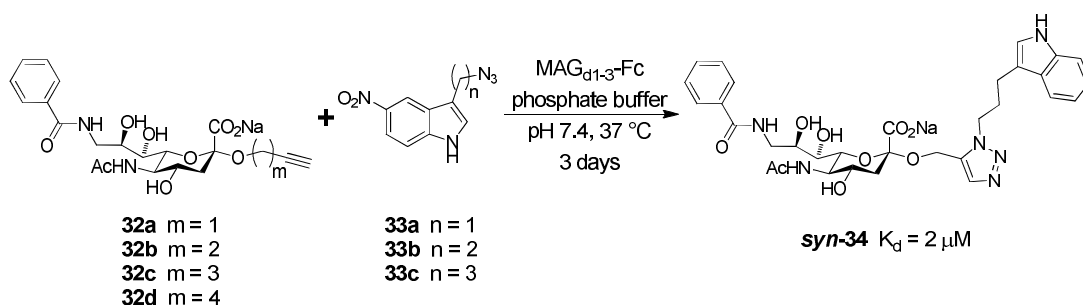


Figure 18: ISCC library of fragments discovered by NMR techniques for the identification of novel MAG protein biligands.

1.5.9 Mycobacterial monooxygenase EthR

Further to the observation by the Sharpless group that acetylcholinesterase adopted a previously unseen conformation upon the binding of the *syn*-triazole formed through an ISCC approach,^[79] Willand, Deprez *et al.* have again demonstrated the ability of ISCC to capture previously unknown protein conformations.^[92] They set up an ISCC screen to develop inhibitors for the mycobacterial monooxygenase EthR. Starting from the scaffold azide **35**, identified

by X-ray crystallography, six batches of ten alkynes were screened in the presence of biological target EthR. Only one triazole product was observed and this was identified as being the *anti*-triazole **36**, Figure 19. The triazole exhibited a 10-fold increase in potency for EthR, but more importantly crystal structures of the triazole bound to EthR revealed a novel protein conformation, with access to a previously buried hydrophobic pocket. This highlights the ability of ISCC to not only identify novel inhibitors but also to investigate the conformational space of flexible protein binding pockets.

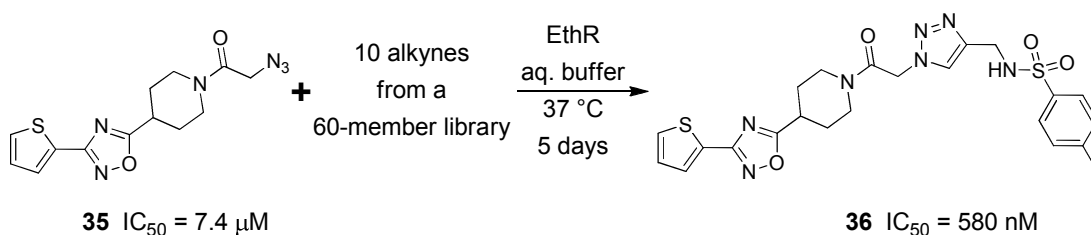


Figure 19: ISCC approach to discover novel inhibitors of EthR.

1.6 Summary – DCC versus kinetic TGS

The last ten years have seen a rapid increase in the use of TGS approaches in the development of novel inhibitors. The leading light in terms of a thermodynamic DCC approach is the reversible disulfide-forming reaction. The kinetic approach has recently caught up with its thermodynamic counterpart with the introduction of the *in situ* click reaction between an azide and terminal alkyne. Whilst these two reactions have led the attack in understanding the power of TGS, researchers remain committed to developing new connections to add to the, at present, small TGS reaction pool. In addition to the lack of diverse reactions suitable for use with a biological template, another notable limitation that both DCC and kinetic TGS currently share is the problem of product detection and analysis. As the field of mass spectrometry has developed to become more sensitive, the problems of product detection in TGS have decreased. The introduction of LC/MS-SIM has provided a great improvement to product detection in TGS.^[47] Solid phase techniques have also been employed in both DCC and ISCC approaches in a bid to further reduce the problems associated with product identification.^[14, 16, 22, 30, 89]

Efforts in TGS are now being directed towards coupling the strategy with another area of drug discovery known as fragment-based drug design. In this

approach, low molecular weight fragments are identified as possessing some binding affinity for a given biological target and based on this knowledge larger lead compounds are built up, Figure 20. Various fragment-based drug discovery approaches have been developed over the last fifteen years and are well reviewed in the literature.^[7, 93-95] By coupling this approach with TGS, the fragments that have been identified in the fragment-based approach can then be built up using a TGS strategy.

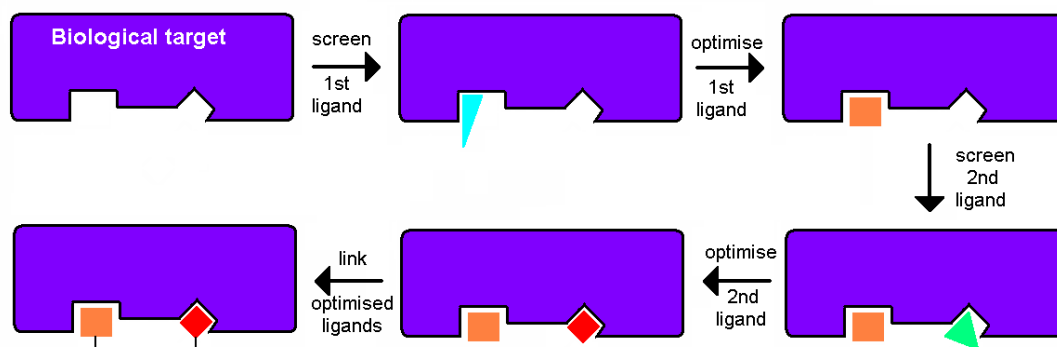


Figure 20: Cartoon outlining the principles of fragment-based drug design.

The first reported case of fragment-based drug design was described by Fesik and co-workers, who used NMR to identify fragments with binding affinity for the target FK506 binding protein, a potent immunosuppressant.^[96] ^{15}N -HSQC NMR was used to identify shifts in peaks characteristic of a binding event on the addition of ^{15}N -labelled biological target to the fragment of interest. Fragments that bind to adjacent sites within the protein are then coupled to create the final high-affinity ligand. Drawbacks associated with this technique are that the protein target should be less than 40 kDA, significant quantities of the ^{15}N -labelled protein are required (> 200 mg) and the fragment compounds must be soluble at high concentrations (> 0.2 mM). NMR-based approaches have been further developed by the introduction of a “spin label” onto a fragment with a known binding site.^[97, 98] Fragments which bind to an adjacent site will experience a paramagnetic relaxation enhancement caused by the spin label which can be readily detected by NMR. Sensitivity is improved using the spin label, which results in a reduction in the amount of protein required (2 mg). This approach was successfully employed to identify fragments that

bind to adjacent products in myelin-associated glycoprotein and these fragments were subsequently linked using an ISCC strategy as described in section 1.5.8.^[91]

A second approach used to identify fragments that bind to a biological target is X-ray crystallography. X-ray crystallography gives the clearest picture of the fragment binding to the target, allowing not only the discovery of the best fragments but also the optimum location for the addition of the linker unit and further information on the environment around the fragment to allow for extra development. Researchers at Abbott Laboratories were the first group to report the successful use of X-ray crystallography in the identification of a novel inhibitor for the anticancer target urokinase.^[99] Protein crystals were soaked in a mixture of fragment compounds. Those fragments which bound to the protein were determined by examining the shape of the electron density map. A hit was identified, 8-hydroxy-2-aminoquinoline, and the crystal structure showed that there was an adjacent pocket that the fragment could be extended in to. The modification to extend the fragment into this pocket resulted in a 1000-fold increase in potency (from $K_i = 56$ to $0.37 \mu\text{M}$) over the original fragment. There are now several companies using this technology to drive their drug discovery programmes including Abbott Laboratories, Astex Therapeutics and Plexxikon. The inhibitors that have been discovered by these companies are well reviewed in the literature.^[7, 95]

Mass spectrometry, developed by Ibis Therapeutics,^[100] and direct binding assays at high concentration, developed by the Ellman group,^[101] have also been employed to identify binding fragments. Unlike the NMR and crystallography-based approaches described above, mass spectrometry and binding assays do not give any information about the binding mode of the fragment within the biological target.

All of the above approaches to fragment-based drug design utilise various detection methods to identify the fragments with the best binding affinity for the biological target. As the interactions between the small fragments and the biological target are typically weak, large concentrations are required to ensure the detection of any binding events. In response to this, Wells and co-workers developed a fragment approach known as tethering.^[102] Based on a similar concept to DCC, fragments are “tethered” to the biological target via a reversible disulfide linkage with a cysteine

residue located in close proximity to the binding pocket. The fragment with the best fit for the binding pocket will form a disulfide bond with the cysteine thiol and can then be detected at much lower concentrations, typically by mass spectrometry. A cysteine residue can be introduced by site-directed mutagenesis if there is not a naturally occurring cysteine residue located near the binding pocket. Some of these fragment-based approaches, including tethering,^[25, 26] X-ray crystallography,^[27, 92] and fragment-based NMR screening,^[91] have already been successfully coupled with DCC and ISCC to identify novel inhibitors.

The main difference between DCC and kinetic TGS is that DCC should always result in the identification of the best binding compound from any given library, even if this compound is not a highly potent inhibitor itself. Depending on the choice of fragments in a kinetic TGS library, there can be a much smaller chance of finding a hit; however this hit is much more likely to be a potent inhibitor. Kinetic TGS has the added drawback that false negatives can occur. If the reactive functional groups are not positioned in the correct orientation, despite a fragment binding to the biological template, then the reaction cannot occur and the fragment will not be identified as being a good inhibitor. Current literature suggests that neither approach is significantly better than the other. The choice between the two comes down to the ease with which a fragment library can be synthesized from the selected scaffold fragment.

One area of TGS that remains relatively unexplored is the possibility of synthesising multivalent ligands by having multiple reactive groups on one fragment and therefore allowing the exploration of more chemical space in one TGS screen. This thesis aims to explore the application of a multivalent approach to both DCC and ISCC and compare the results observed from both systems.

1.7 References

- [1] Did you know? Facts and figures about the pharmaceutical industry in the UK, 2011, pp. <http://www.abpi.org.uk/our>.
- [2] S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, A. L. Schacht, *Nat Rev Drug Discov* **2010**, 9, 203.

- [3] R. M. Sanchez-Martin, S. Mittoo, M. Bradley, *Curr. Top. Med. Chem.* **2004**, *4*, 653.
- [4] R. Macarron, *Drug Discovery Today* **2006**, *11*, 277.
- [5] X. Barril, R. Soliva, *Molecular BioSystems* **2006**, *2*, 660.
- [6] J. J. Diaz-Mochon, G. Tourniaire, M. Bradley, *Chem. Soc. Rev.* **2007**, *36*, 449.
- [7] P. J. Hajduk, J. Greer, *Nat Rev Drug Discov* **2007**, *6*, 211.
- [8] C. E. Housecroft, E. C. Constable, *Chemistry: an introduction to organic, inorganic, and physical chemistry*, Prentice Hall, **2002**.
- [9] J. F. Chase, P. K. Tubbs, *Biochem. J.* **1969**, *111*, 225.
- [10] W. P. Jencks, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4046.
- [11] D. Rideout, *Science* **1986**, *233*, 561.
- [12] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652.
- [13] I. Huc, J.-M. Lehn, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 2106.
- [14] O. Ramström, J.-M. Lehn, *ChemBioChem* **2000**, *1*, 41.
- [15] M. F. Greaney, V. T. Bhat, *Protein-Directed Dynamic Combinatorial Chemistry*, John Wiley & Sons, Inc., New York, **2010**.
- [16] B. Klekota, M. H. Hammond, B. L. Miller, *Tetrahedron Lett.* **1997**, *38*, 8639.
- [17] R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 3405.
- [18] R. Larsson, Z. Pei, O. Ramström, *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 3716.
- [19] J. R. Thomas, P. J. Hergenrother, *Chem. Rev.* **2008**, *108*, 1171.
- [20] F. Aboul-ela, *Future Med. Chem.* **2010**, *2*, 93.
- [21] A. Bugaut, K. Jantos, J.-L. Wietor, R. Rodriguez, J. K. M. Sanders, S. Balasubramanian, *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 2677.
- [22] P. C. Gareiss, K. Sobczak, B. R. McNaughton, P. B. Palde, C. A. Thornton, B. L. Miller, *J. Am. Chem. Soc.* **2008**, *130*, 16254.
- [23] P. López-Senín, I. Gómez-Pinto, A. Grandas, V. Marchán, *Chem.-Eur. J.* **2011**, *17*, 1946.

- [24] F. R. Bowler, J. J. Diaz-Mochon, M. D. Swift, M. Bradley, *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 1809.
- [25] M. T. Cancilla, M. M. He, N. Viswanathan, R. L. Simmons, M. Taylor, A. D. Fung, K. Cao, D. A. Erlanson, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3978.
- [26] B. M. R. Liénard, R. Hüting, P. Lassaux, M. Galleni, J.-M. Frère, C. J. Schofield, *J. Med. Chem.* **2008**, *51*, 684.
- [27] D. E. Scott, G. J. Dawes, M. Ando, C. Abell, A. Ciulli, *ChemBioChem* **2009**, *10*, 2772.
- [28] R. F. Ludlow, S. Otto, *J. Am. Chem. Soc.* **2008**, *130*, 12218.
- [29] R. F. Ludlow, S. Otto, *J. Am. Chem. Soc.* **2010**, *132*, 5984.
- [30] Y.-Y. Chen, W.-Y. He, Y. Wu, C.-Q. Niu, G. Liu, *J. Combin. Chem.* **2008**, *10*, 914.
- [31] B. Shi, M. F. Greaney, *Chem. Commun.* **2005**, 886.
- [32] B. Shi, R. Stevenson, D. J. Campopiano, M. F. Greaney, *J. Am. Chem. Soc.* **2006**, *128*, 8459.
- [33] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat Chem* **2010**, *2*, 490.
- [34] G. R. L. Cousins, S.-A. Poulsen, J. K. M. Sanders, *Chem. Commun.* **1999**, 1575.
- [35] R. L. E. Furlan, Y.-F. Ng, S. Otto, J. K. M. Sanders, *J. Am. Chem. Soc.* **2001**, *123*, 8876.
- [36] S. L. Roberts, R. L. E. Furlan, G. R. L. Cousins, J. K. M. Sanders, *Chem. Commun.* **2002**, 938.
- [37] J. Liu, K. R. West, C. R. Bondy, J. K. M. Sanders, *Org. Biomol. Chem.* **2007**, *5*, 778.
- [38] M. G. Simpson, M. Pittelkow, S. P. Watson, J. K. M. Sanders, *Org. Biomol. Chem.* **2010**, *8*, 1173.
- [39] M. G. Simpson, M. Pittelkow, S. P. Watson, J. K. M. Sanders, *Org. Biomol. Chem.* **2010**, *8*, 1181.
- [40] T. Bunyapaiboonsri, O. Ramström, S. Lohmann, J.-M. Lehn, L. Peng, M. Goeldner, *ChemBioChem* **2001**, *2*, 438.

- [41] T. Bunyapaiboonsri, H. Ramström, O. Ramström, J. Haiech, J.-M. Lehn, *J. Med. Chem.* **2003**, *46*, 5803.
- [42] O. Ramström, S. Lohmann, T. Bunyapaiboonsri, J.-M. Lehn, *Chem.-Eur. J.* **2004**, *10*, 1711.
- [43] S.-A. Poulsen, *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1074.
- [44] E. H. Cordes, W. P. Jencks, *J. Am. Chem. Soc.* **1962**, *84*, 826.
- [45] A. Dirksen, S. Dirksen, T. M. Hackeng, P. E. Dawson, *J. Am. Chem. Soc.* **2006**, *128*, 15602.
- [46] A. Dirksen, P. E. Dawson, *Bioconj. Chem.* **2008**, *19*, 2543.
- [47] R. Manetsch, A. Krasinski, Z. Radic, J. Raushel, P. Taylor, K. B. Sharpless, H. C. Kolb, *J. Am. Chem. Soc.* **2004**, *126*, 12809.
- [48] A. V. Eliseev, *Drug Discovery Today* **2004**, *9*, 348.
- [49] G. E. Boldt, T. J. Dickerson, K. D. Janda, *Drug Discovery Today* **2006**, *11*, 143.
- [50] X. Hu, R. Manetsch, *Chem. Soc. Rev.* **2010**, *39*, 1316.
- [51] J. Inglese, S. J. Benkovic, *Tetrahedron* **1991**, *47*, 2351.
- [52] D. L. Boger, N.-E. Haynes, M. S. Warren, J. Ramcharan, P. A. Kitos, S. J. Benkovic, *Bioorg. Med. Chem.* **1997**, *5*, 1853.
- [53] S. E. Greasley, T. H. Marsilje, H. Cai, S. Baker, S. J. Benkovic, D. L. Boger, I. A. Wilson, *Biochemistry* **2001**, *40*, 13538.
- [54] R. Nguyen, I. Huc, *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 1774.
- [55] K. C. Nicolaou, R. Hughes, S. Y. Cho, N. Winssinger, C. Smethurst, H. Labischinski, R. Endermann, *Angew. Chem., Int. Ed. Engl.* **2000**, *39*, 3823.
- [56] K. C. Nicolaou, R. Hughes, S. Y. Cho, N. Winssinger, H. Labischinski, R. Endermann, *Chem.-Eur. J.* **2001**, *7*, 3824.
- [57] L. Weber, *Drug Discovery Today: Technologies* **2004**, *1*, 261.
- [58] X. Hu, J. Sun, H.-G. Wang, R. Manetsch, *J. Am. Chem. Soc.* **2008**, *130*, 13820.
- [59] M. R. Arkin, J. A. Wells, *Nat Rev Drug Discov* **2004**, *3*, 301.
- [60] P. Oberdoerffer, D. A. Sinclair, *Nat Rev Mol Cell Biol* **2007**, *8*, 692.
- [61] S.-i. Imai, C. M. Armstrong, M. Kaerberlein, L. Guarente, *Nature* **2000**, *403*, 795.

- [62] T. Asaba, T. Suzuki, R. Ueda, H. Tsumoto, H. Nakagawa, N. Miyata, *J. Am. Chem. Soc.* **2009**, *131*, 6989.
- [63] M. K. Herrlein, R. L. Letsinger, *Nucleic Acids Res.* **1994**, *22*, 5076.
- [64] M. K. Herrlein, J. S. Nelson, R. L. Letsinger, *J. Am. Chem. Soc.* **1995**, *117*, 10151.
- [65] Y. Xu, E. T. Kool, *Nucleic Acids Res.* **1999**, *27*, 875.
- [66] S. Sando, E. T. Kool, *J. Am. Chem. Soc.* **2002**, *124*, 2096.
- [67] S. Ficht, A. Mattes, O. Seitz, *J. Am. Chem. Soc.* **2004**, *126*, 9970.
- [68] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2596.
- [69] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057.
- [70] H. C. Kolb, K. B. Sharpless, *Drug Discovery Today* **2003**, *8*, 1128.
- [71] J. E. Moses, A. D. Moorhouse, *Chem. Soc. Rev.* **2007**, *36*, 1249.
- [72] G. C. Tron, T. Pirali, R. A. Billington, P. L. Canonico, G. Sorba, A. A. Genazzani, *Med. Res. Rev.* **2008**, *28*, 278.
- [73] Guest Editors, M. G. Finn, V. Fokin, *Chem. Soc. Rev.* **2010**, *39*, 1221.
- [74] L. Zhang, X. G. Chen, P. Xue, H. H. Y. Sun, I. D. Williams, K. B. Sharpless, V. V. Fokin, G. C. Jia, *J. Am. Chem. Soc.* **2005**, *127*, 15998.
- [75] W. L. Mock, T. A. Irra, J. P. Wepsiec, T. L. Manimaran, *J. Org. Chem.* **1983**, *48*, 3619.
- [76] W. L. Mock, T. A. Irra, J. P. Wepsiec, M. Adhya, *J. Org. Chem.* **1989**, *54*, 5302.
- [77] W. G. Lewis, L. G. Green, F. Grynszpan, Z. Radic, P. R. Carlier, P. Taylor, M. G. Finn, K. B. Sharpless, *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 1053.
- [78] P. Taylor, Z. Radic, *Annu. Rev. Pharmacol. Toxicol.* **1994**, *34*, 281.
- [79] Y. Bourne, H. C. Kolb, Z. Radic, K. B. Sharpless, P. Taylor, P. Marchot, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 1449.
- [80] A. Krasinski, Z. Radic, R. Manetsch, J. Raushel, P. Taylor, K. B. Sharpless, H. C. Kolb, *J. Am. Chem. Soc.* **2005**, *127*, 6686.
- [81] V. P. Mocharla, B. Colasson, L. V. Lee, S. Roper, K. B. Sharpless, C. H. Wong, H. C. Kolb, *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 116.

- [82] M. Whiting, J. Muldoon, Y. C. Lin, S. M. Silverman, W. Lindstrom, A. J. Olson, H. C. Kolb, M. G. Finn, K. B. Sharpless, J. H. Elder, V. V. Fokin, *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 1435.
- [83] A. T. Poulin-Kerstien, P. B. Dervan, *J. Am. Chem. Soc.* **2003**, *125*, 15811.
- [84] P. B. Dervan, B. S. Edelson, *Curr. Opin. Struct. Biol.* **2003**, *13*, 284.
- [85] T. Hirose, T. Sunazuka, A. Sugawara, A. Endo, K. Iguchi, T. Yamamoto, H. Ui, K. Shiomi, T. Watanabe, K. B. Sharpless, S. Omura, *J. Antibiot.* **2009**, *62*, 277.
- [86] O. A. Andersen, M. J. Dixon, I. M. Eggleston, D. M. F. van Aalten, *Nat. Prod. Rep.* **2005**, *22*, 563.
- [87] K. Shiomi, N. Arai, Y. Iwai, A. Turberg, H. Kölbl, S. mura, *Tetrahedron Lett.* **2000**, *41*, 2141.
- [88] F. V. Rao, D. R. Houston, R. G. Boot, J. M. F. G. Aerts, M. Hodgkinson, D. J. Adams, K. Shiomi, S. Omura, D. M. F. van Aalten, *Chemistry & biology* **2005**, *12*, 65.
- [89] H. D. Agnew, R. D. Rohde, S. W. Millward, A. Nag, W.-S. Yeo, J. E. Hein, S. M. Pitram, A. A. Tariq, V. M. Burns, R. J. Krom, V. V. Fokin, K. B. Sharpless, J. R. Heath, *Angew. Chem., Int. Ed. Engl.* **2009**, *48*, 4944.
- [90] S. S. Lee, J. Lim, S. Tan, J. Cha, S. Y. Yeo, H. D. Agnew, J. R. Heath, *Anal. Chem.* **2010**, *82*, 672.
- [91] S. V. Shelke, B. Cutting, X. Jiang, H. Koliwer-Brandl, D. S. Strasser, O. Schwardt, S. Kelm, B. Ernst, *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 5721.
- [92] N. Willand, M. Desroses, P. Toto, B. Dirié, Z. Lens, V. Villeret, P. Rucktooa, C. Loch, A. Baulard, B. Deprez, *ACS Chem. Biol.* **2010**, *5*, 1007.
- [93] D. A. Erlanson, R. S. McDowell, T. O'Brien, *J. Med. Chem.* **2004**, *47*, 3463.
- [94] D. A. Erlanson, J. A. Wells, A. C. Braisted, *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 199.
- [95] R. A. E. Carr, M. Congreve, C. W. Murray, D. C. Rees, *Drug Discovery Today* **2005**, *10*, 987.
- [96] S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **1996**, *274*, 1531.

- [97] W. Jahnke, L. B. Perez, C. G. Paris, A. Strauss, G. Fendrich, C. M. Nalin, *J. Am. Chem. Soc.* **2000**, *122*, 7394.
- [98] W. Jahnke, A. Florsheimer, M. J. J. Blommers, C. Paris, J. Heim, C. M. Nalin, L. B. Perez, *Curr. Top. Med. Chem.* **2003**, *3*, 69.
- [99] V. L. Nienaber, P. L. Richardson, V. Klighofer, J. J. Bouska, V. L. Giranda, J. Greer, *Nat Biotech* **2000**, *18*, 1105.
- [100] S. A. Hofstadler, R. H. Griffey, *Chem. Rev.* **2001**, *101*, 377.
- [101] D. J. Maly, I. C. Choong, J. A. Ellman, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 2419.
- [102] D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9367.

Chapter 2: A dynamic combinatorial chemistry approach to identify novel bivalent inhibitors of glutathione transferases

The first target-guided synthesis (TGS) method to be investigated was the thermodynamic approach using dynamic combinatorial chemistry (DCC). The biological templates that were selected were glutathione transferases.

2.1 Glutathione transferases

Glutathione transferases (GST) are a family of enzymes primarily responsible for the conjugation of glutathione (γ -Glu-Cys-Gly, GSH) **2** to a number of substrates in the cell.^[1-6] This is the first step in the mercapturic acid pathway, a key pathway in the metabolism of xenobiotic electrophiles.^[7] Conjugate addition occurs between the thiol of GSH and an electrophilic substrate. The conjugation of GSH is promoted within two binding pockets in the enzyme; the glutathione binding region (“G-site”) and the adjacent hydrophobic acceptor binding region (“H-site”), Figure 1.

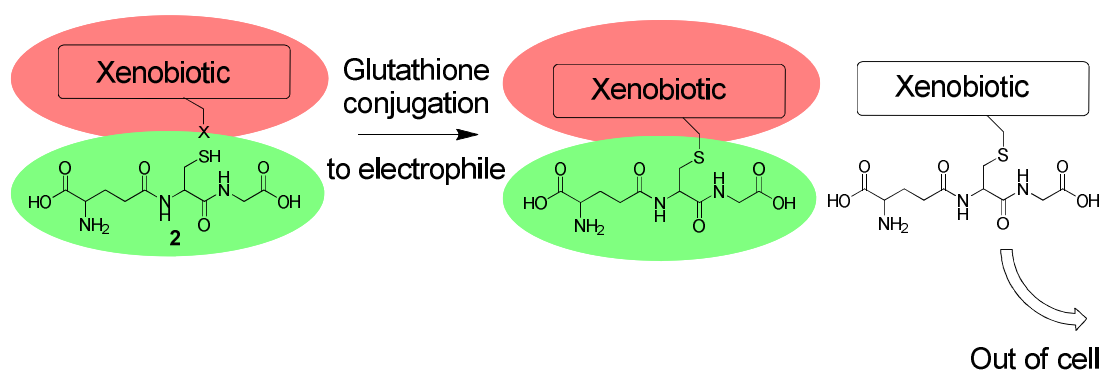


Figure 1: Glutathione-conjugation to xenobiotics. Green = glutathione binding site. Red = hydrophobic binding site.

GSTs can be separated into three different families; cytosolic, mitochondrial and membrane-bound microsomal. The cytosolic family can be split into a further seven classes; Alpha, Mu, Omega, Pi, Sigma, Theta and Zeta. The sequence identity within a class is generally quite high at above 60%, but the sequence identity between classes is much lower, at about 30%. The GST family is extensively reviewed in the literature.^[2-6] The three most abundant mammalian classes are the

alpha, mu and pi classes. Each class can be further sub-divided, with four subunits identified in human for the Mu class (M1 - M4), four subunits in the Alpha class (A1 - A4) and one subunit in the Pi class (P1). In all three classes the GST monomer unit is split into two domains. The N-terminal domain 1 is highly conserved between all of the cytosolic classes and forms the majority of the glutathione-binding G-site. The C-terminal domain 2 provides the majority of the residues that make up the hydrophobic pocket and is much less conserved between the mu, pi and alpha classes. It is these differences that are responsible for the variation in substrate specificity observed between the classes of GST. There are more structural features that are unique to each class of GST. The Mu class GSTs have the 'Mu loop' which results in a deeper active site cleft than those found in the Pi class.^[8] Domain 2 of the Alpha class contains an extra α -helix that folds onto the hydrophobic binding pocket, making it smaller and more hydrophobic than the H-site of the Pi and Mu class of enzymes.^[9] Representatives from all three classes have been selected to use as the biological template in this DCC approach in an effort to probe isoform specificity. The human form of the P1 subclass, hGSTP1-1, has been selected from the pi class, the mouse form of the M1 subclass, mGSTM1-1, has been selected as a representative of the mu class of GSTs and the mouse form of the A4 subclass, mGSTA4-4 has been chosen from the alpha class. The helminth GST *Schistosoma japonica* GST (SjGST) has also been selected as a fourth non-mammalian GST isoform to investigate. The three-dimensional structure of SjGST most closely resembles the Mu class of the mammalian GSTs and contains a shortened form of the Mu loop creating a more open hydrophobic binding pocket, as shown in Figure 2a.^[10-12]

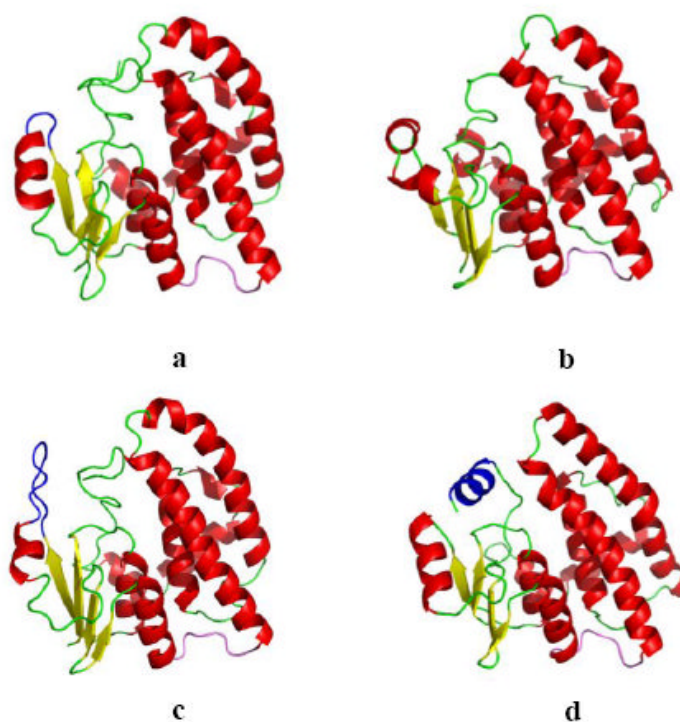


Figure 2: Three-dimensional structures of GST monomers. Unique structural features located near the hydrophobic binding pocket, the Mu loop or extra α -helix, are represented in blue. **a)** SjGST (PDB: 1Y6E),^[13] **b)** hGSTP1 (PDB: 1EOG),^[14] **c)** mGSTM1 (generated with the protein structure homology-modeling server SWISS-MODEL by Anne Caniard),^[15] **d)** mGSTA4 (PDB: 1GUK).^[16]

To become catalytically active the GST monomers form non-covalently bound homodimers. The active sites of each monomer lie at either end of a solvent-accessible cleft, as demonstrated in Figure 3. The distance between the active sites and the environment in the cleft vary between the GST classes.

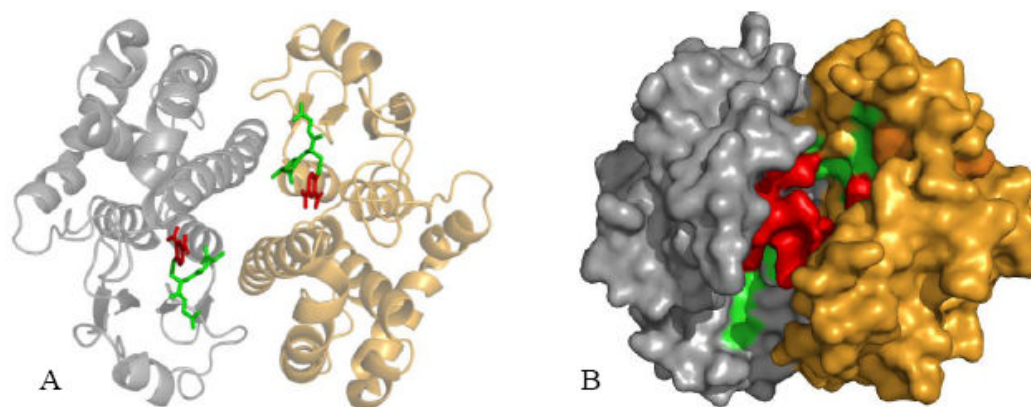
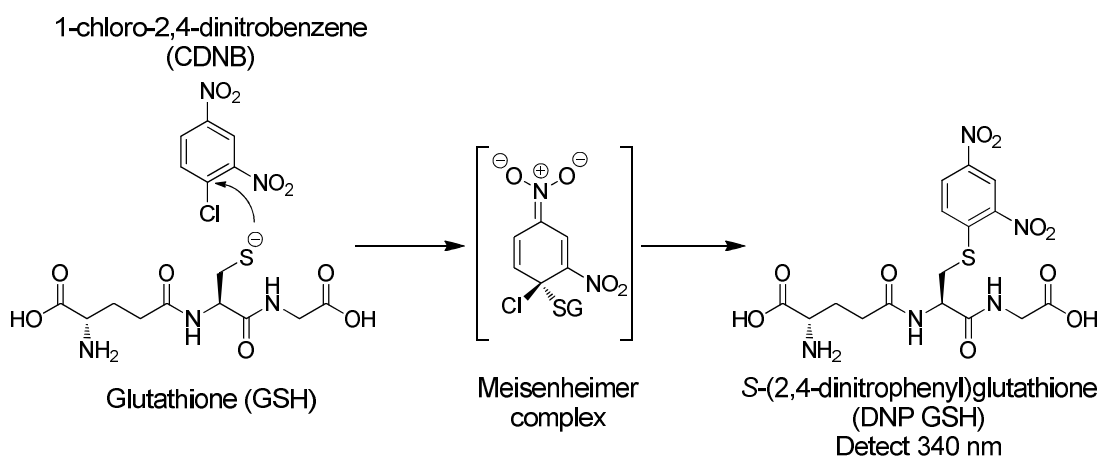


Figure 3: **A)** structure of hGSTM1-1 (PDB: 1XWK) with glutathione-conjugated xenobiotic bound in each active site.^[17] The grey and orange structures display the two monomer units. The green compound is glutathione and the red compound is the xenobiotic, **B)** the same structure illustrating the G-site in green and the H-site in red. The solvent-accessible cleft between the two monomer active sites can clearly be seen. This figure was generated using PyMOL by Anne Caniard.^[15]

GST activity is determined by measuring the rate of the conjugation of glutathione to a universal substrate 1-chloro-2,4-dinitrobenzene (CDNB), a substrate known to bind to all GST isoforms.^[7] The reaction proceeds via an S_NAr mechanism, forming an intermediate Meisenheimer complex, to yield the glutathione-conjugated product, *S*-(2,4-dinitrophenyl)glutathione (DNP GSH) as demonstrated in Scheme 1.^[18] DNP GSH absorbs at 340 nm, allowing the rate of product formation to be monitored using visible spectroscopy.



Scheme 1: GST enzymatic activity measured using the glutathione conjugation reaction to universal substrate CDNB.

A secondary role for GSTs appears to be the cellular uptake and distribution of large lipophilic molecules. These compounds, which include porphyrins, steroids and fatty acids, typically contain large planar aromatic regions, often with anionic functional groups. Such ligands are characterized by their display of non-competitive inhibition of glutathione conjugation.^[2] The site where these compounds bind to GST has been labelled the ligandin binding site. There is much debate over the location of the ligandin site and there is some suggestion that the location is variable between GST isoforms. Evidence suggests that in the GSTP1-1 isoform it is adjacent to the hydrophobic binding pocket and some overlap between the two sites can occur.^[19] A crystal structure of the drug praziquantel bound to SjGST demonstrates that the ligand binds in the inter-subunit cleft of the protein, as shown in Figure 4.^[11]



Figure 4: Crystal structure of SjGST dimer in complex with praziquantel (PDB: 1GTB).^[11] One monomer unit = orange, second monomer unit = grey, praziquantel = pink. Glutathione binding site is highlighted by the catalytic tyrosine residues (displayed in blue). This figure was generated using PyMOL by Anne Caniard.^[15]

GSTs have become important drug targets in cancer therapy in recent years after the observation of GST overexpression in cancer cell lines and the direct correlation this has with increased resistance to chemotherapeutic drugs.^[20-22] The resistance is thought to be the result of two different responses of GSTs. The first is that GSTs can be responsible for the metabolism of cancer drugs and their excretion from the tumour cells.^[23] The second response is as a result of the interaction of GSTs in the mitogen-activated protein (MAP) kinase pathway. GSTP1-1 has been shown to bind to and inhibit *c-Jun* NH₂-terminal kinase (JNK), which plays an

important role in proapoptotic signaling within the MAP kinase pathway.^[20-22, 24] Under oxidative or chemical stress GSTP1-1 dissociates from JNK, which is then activated leading to the induction of apoptosis. JNK activation is required for a number of anticancer drugs to have maximal cytotoxicity. GSTM1-1 has also been identified as playing a role in the MAP kinase pathway.^[25-27] GSTM1-1 binds to apoptosis signal-regulating kinase 1 (ASK1), which is a MAP kinase kinase kinase that activates both the JNK and p38 signalling pathways leading to apoptosis. As with the GSTP1-JNK complex, GSTM1-1 acts as an inhibitor of ASK1, which dissociates in response to cell stress resulting in the activation of ASK1.

Helminth GSTs are also important drug targets in parasitic diseases, such as malaria and schistosomiasis. Schistosomal GSTs are believed to be responsible for resistance to praziquantel, the major drug used to treat schistosomiasis.^[5]

A number of different approaches have been used to design inhibitors for GSTs and these have been well reviewed.^[22, 28] Modifications of the glutathione tripeptide typically result in a reduction in binding affinity, biological instability as a drug-like compound and, because of the highly conserved nature of the glutathione binding pocket, a lack of GST isoform specificity. A more successful example of a glutathione analogue is the compound with the addition of an α -phenyl glycine residue, **37** designed by Telik Technologies, which was highly potent ($K_i = 0.12 \mu\text{M}$) and demonstrated good selectivity for GSTP1-1.^[29] A second approach utilizes more hydrophobic inhibitors designed to bind in the hydrophobic pocket, an approach that has been utilized previously in the Greaney group to design novel inhibitors of the H-site using DCC.^[30, 31] Due to the large size of this binding pocket, this approach can also prove to be problematic in finding isoform specificity. Other approaches to GST inhibitor design include ligandin-type inhibitors and the use of pro-drugs designed to take advantage of the catalytic activity of the GST to release the active drugs. These approaches have been well reviewed in the literature.^[22, 28]

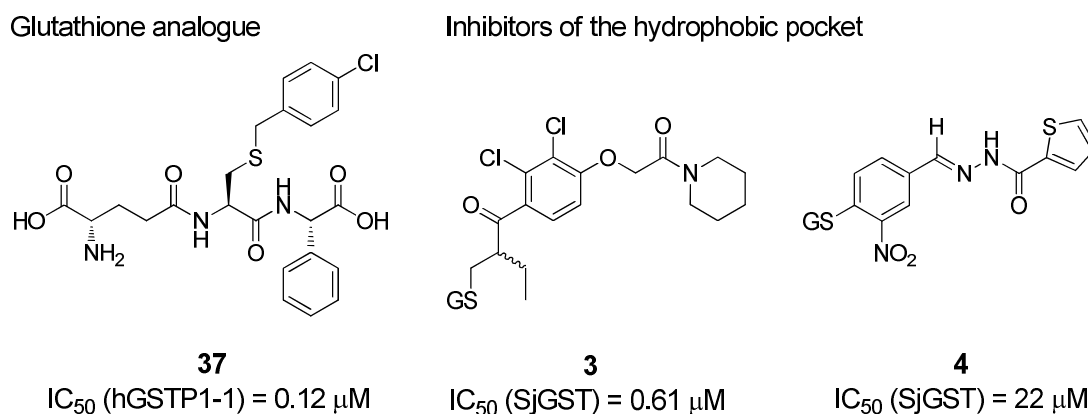
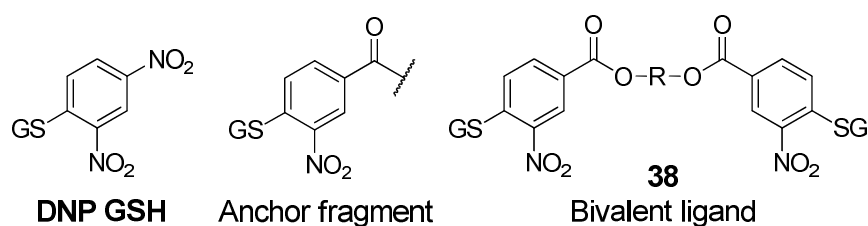


Figure 5: Literature examples of GST inhibitors.

In 2003, the Atkins group first reported a bivalent approach to the design of novel GST inhibitors.^[32] The use of multivalent inhibitors that bind to multiple binding sites on the biological target has been shown to increase binding affinity by several orders of magnitude over the original monovalent ligands.^[33] This is achieved by reducing the entropic price associated with ligand binding. The cost of translational entropy and rotational entropy is the same for binding of either one monovalent ligand or one bivalent ligand; therefore the entropic cost of binding two monovalent ligands is double the cost of binding one bivalent ligand. Careful design of a rigid linking group with an optimised spacing to match the ligands to the binding sites minimises the cost of a reduction in conformational entropy of the linking group. The necessity for optimised spacing also results in ligands with high selectivity for a given target.

The Atkins group took advantage of the dimeric nature of GSTs with the solvent-accessible channel between the active sites of each monomer unit to design bivalent inhibitors. Their strategy was to link two molecules of a known GST inhibitor with a spacer designed to bind in the solvent-accessible cleft. The first reported bivalent inhibitors used either an analogue of DNP GSH, the product of the glutathione conjugation of the universal substrate CDNB, or uniblue A, an analogue of cibacron blue as the anchor fragments.^[32] Cibacron blue is a dye known to bind to the H-site/ligandin pocket of GST. A small library of compounds was made using three spacers of increasing length and, in the example using the DNP GSH analogue, the binding data showed isoform specificity between A1 and P1 based on spacer length (Figure 6). The longest spacer in the library demonstrated a 10-fold increase

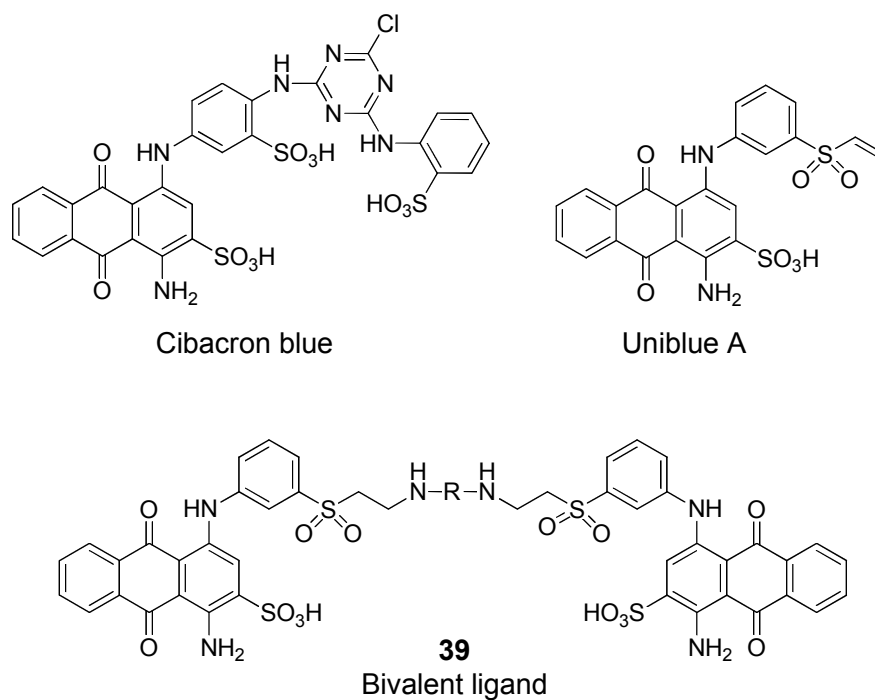
in binding affinity for GSTP1-1 over the GSTA1-1 isoform. This ligand also exhibited a 100-fold increase in inhibitory activity over the monovalent reference, highlighting the beneficial use of a multivalent approach.



O-R-O	GSTA1-1 IC ₅₀ (μM)	GSTP1-1 IC ₅₀ (μM)
38a (Monovalent reference)	30	21
38b 	29	4.5
38c 	1.6	2.9
38d 	3.3	0.28

Figure 6: Design of bivalent ligands based on the glutathione-conjugated product of CDNB and the inhibitory activity of the ligands against GSTA1-1 and GSTP1-1.

The bivalent ligands with terminal unibblue A groups showed little binding affinity for the Alpha GST but excellent affinity for GSTP1-1 with a 70-fold improvement in binding affinity for the best binding bivalent compound **39b** over the monomer unibblue A, Figure 7. At the time of the article, the best binding ligand was the most potent inhibitor of GSTP1-1 reported in the literature.




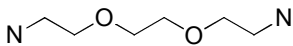
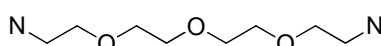
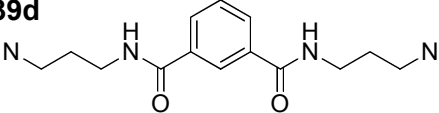
O-R-O	GSTA1-1 IC ₅₀ (μM)	GSTP1-1 IC ₅₀ (μM)
39a  (Monovalent reference)	> 100	5.0
39b 	> 100	0.044
39c 	> 100	0.072
39d 	> 100	0.440

Figure 7: Design of bivalent ligands based on Uniblue A, an analogue of the GST ligandin site binding dye Cibacron Blue, and the inhibitory activity of the ligands against GSTA1-1 and GSTP1-1.

The Atkins group later went on to report further bivalent inhibitors of GST using ethacrynic acid as the anchor ligand.^[34, 35] Ethacrynic acid is a non-specific GST inhibitor that binds into the hydrophobic binding pocket. Isoform specificity between GSTA1-1 and GSTP1-1 was achieved, firstly, by varying the length of the spacer and secondly, by using a peptidic spacer to exploit the differences in the

environment of the solvent-accessible cleft. In the first case, as demonstrated in Figure 8A, the shorter spacer length favoured binding to GSTA1-1 but increasing the length of the spacer improved the binding to GSTP1-1.^[34] In the second case a combinatorial library of peptide spacers was synthesized based on the design of **41** in Figure 8B. Negatively charged or phenyl side groups on the spacer resulted in much greater selectivity towards GSTA1-1, while hydrophobic groups showed an improved affinity towards the GSTP1-1 isoform.^[35] The spacer length resulted in dramatic changes in binding affinity between compounds with a 30-fold difference observed between **40a** and **40b** for GSTP1-1. The changes in side chain residue produced more subtle changes with a 3-fold difference in binding affinity for both GSTA1-1 and GSTP1-1.

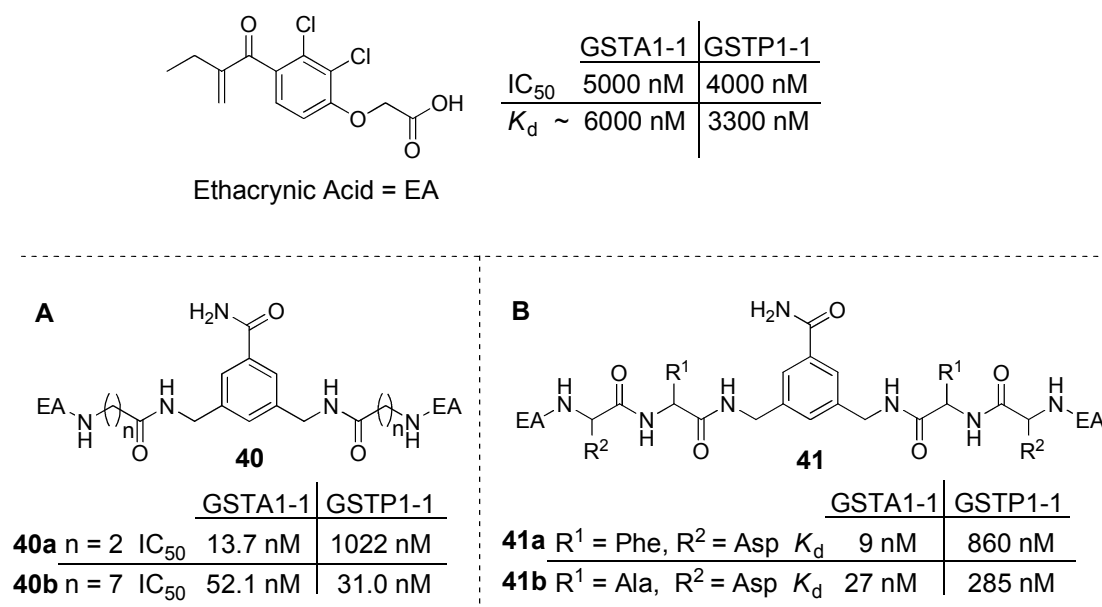


Figure 8: Bivalent ligands of GST based on an ethacrynic acid anchor fragment. **A)** Changes in binding affinity for GSTA1-1 and GSTP1-1 by varying spacer length. **B)** Changes in binding affinity for GSTA1-1 and GSTP1-1 by varying the side chains of the spacer.

The anchoring fragments at the terminus of each spacer, the differences in the spacer length and, perhaps to a lesser extent, the environment of the solvent-accessible cleft between GST isoforms clearly offer great potential to develop novel isoform specific GST inhibitors. The design of the ligands, with two

building blocks connected by a central spacer unit, offers an excellent opportunity to utilise TGS techniques to investigate these three areas.

2.2 Design and synthesis of initial dynamic combinatorial library

The reversible reaction that was chosen to create the linkage between the H-site anchoring fragments and the spacer unit was the formation of acylhydrazones from a hydrazide and an aldehyde species. The aniline-catalyzed acylhydrazone reaction is a novel tool for biologically-templated DCC developed within the Greaney group.^[31] As described in the introduction, the use of aniline as a catalyst allows equilibrium to be reached rapidly at pH 6 - 8, a pH window making the reaction suitable for use with biological agents. The Greaney group developed a DCL focusing on an aldehyde fragment based on DNP GSH (Figure 9B). This aldehyde fragment (**ALD01**) is similar to the DNP GSH-based fragments used in the first bivalent GST inhibitors described by Atkins **38** (Figure 9A).^[32] Replacing the ester link used in these inhibitors with an acylhydrazone group leads to a bivalent spacer design synthesised from a library of bishydrazide spacers and the DNP GSH analogue **ALD01**, as demonstrated in Figure 9C.

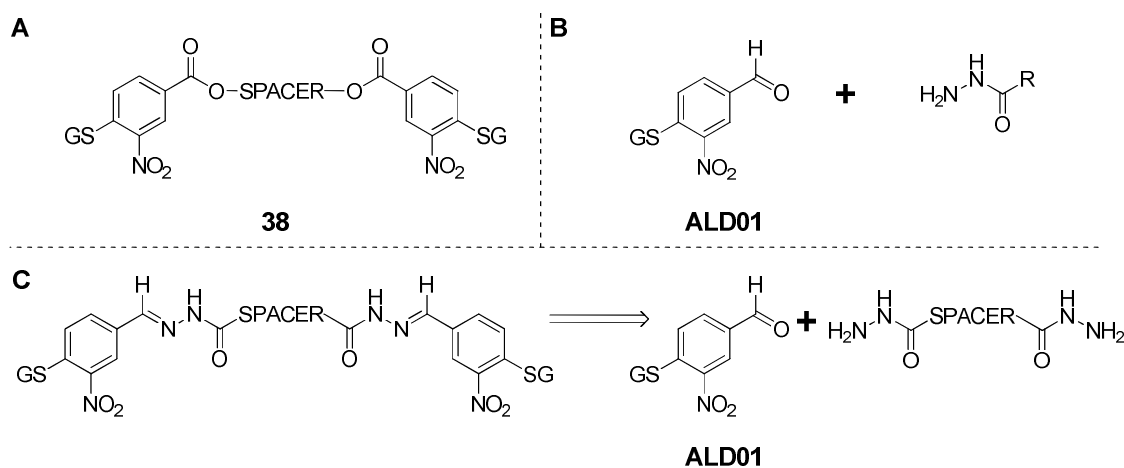


Figure 9: **A)** Literature example of bivalent ligand based on DNP GSH anchor.^[32] **B)** DCL designed in the Greaney group based on DNP GSH anchor fragment.^[31] **C)** Proposed bivalent ligand with acyl hydrazone linkage and the fragments from which these ligands will be synthesised.

Bishydrazide spacers were designed based on both the alkyl and peptidic spacers used by Atkins.^[32, 34, 35] The hydrazide functionality could be easily added to the spacers *via* the condensation of hydrazine onto ester groups. With this in mind, several spacers were designed as outlined in Figure 10.

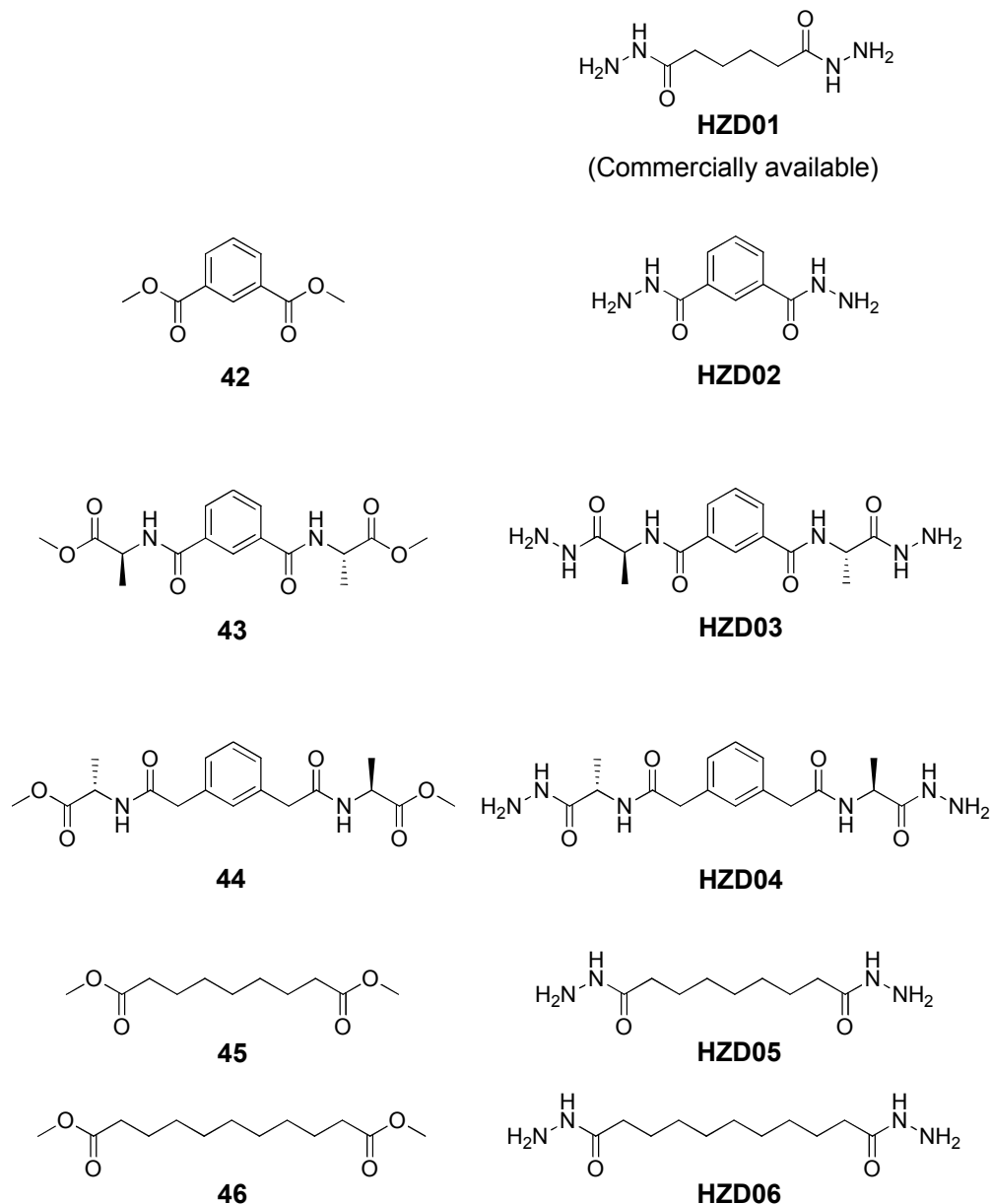
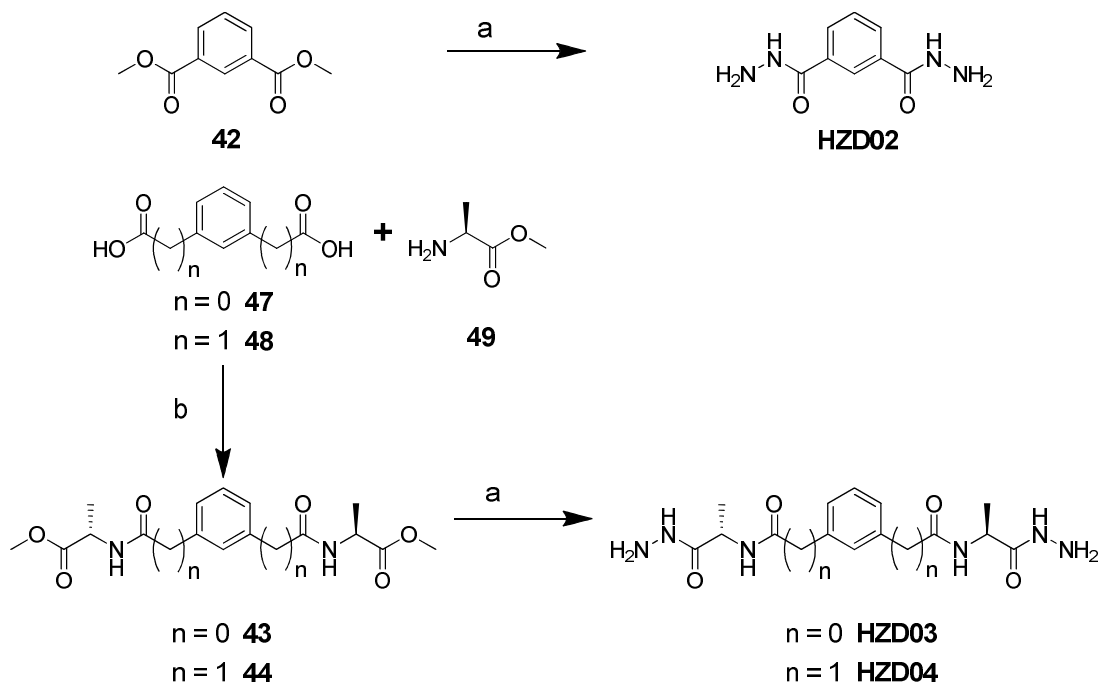


Figure 10: Design of hydrazide spacer library and the esters from which they will be synthesized.

The synthesis of hydrazide (**HZD**) spacers **HZD02** to **HZD04** is outlined in Scheme 2. For **HZD02**, only one step was required, the hydrazine condensation onto the commercially available ester dimethylisophthalate **42**. Two steps were required

for the synthesis of **HZD03** and **HZD04**. The first step was the amide coupling of alanine methyl ester **49** with the core dicarboxylic acid moiety **47** and **48**. This was followed by the condensation of hydrazine onto the ester group to yield the desired bisacylhydrazide spacers.



Scheme 2: Synthesis of hydrazide spacer library: a; hydrazine hydrate (6 eq.) and methanol, rt, overnight, b; alanine methyl ester (2 eq.), EDC.HCl (2.4 eq.), DMAP (0.4 eq.) and chloroform, rt, overnight.

For the alkyl spacers, **HZD01** was commercially available. The reaction to yield **HZD05** and **HZD06** was again the straight forward hydrazine condensation on to the ester group. The white solid products, however, were insoluble in most common solvents. The solids were insoluble in water, DMSO and ethanol which were the solvents that would be used in the DCC experiments. Because of this, these two hydrazide spacers were abandoned.

2.2.1 Optimisation and identification of the initial dynamic combinatorial library

The initial dynamic combinatorial library **DCL01** took a mixture of four hydrazide spacers, **HZD01-04** with one aldehyde, **ALD01** to yield a four member

acylhydrazone library, as demonstrated in Figure 11. The acylhydrazone (**HZN**) products have been labelled in the format **HZNYX-XX-YY**, where **XX** represents the code for the hydrazide spacer and **YY** represents the code for the aldehyde at each terminal position. As **DCL01** only contains one aldehyde **ALD01** in the reaction mixture, the label takes the format **HZN01-XX-01**.

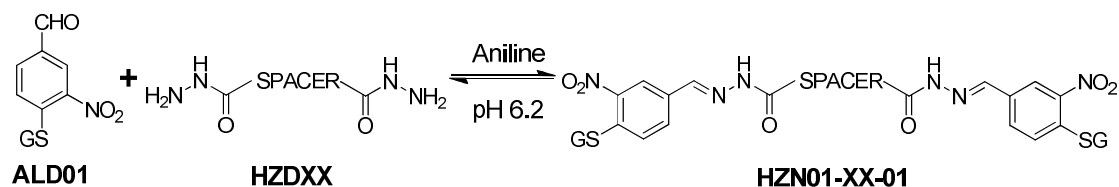
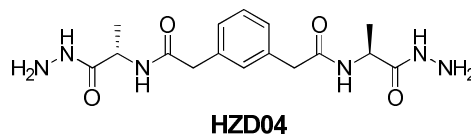
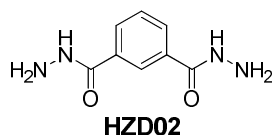
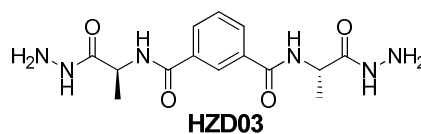
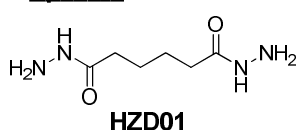
DCL01**Spacers**

Figure 11: DCL01 – library set up between aldehyde **ALD01** and hydrazide spacers **HZD01-HZD04**.

The library was set up and monitored over several hours to determine when equilibrium was reached. Figure 12 indicates that the equilibrium was reached within 8 hours in the presence of 5 mM aniline.

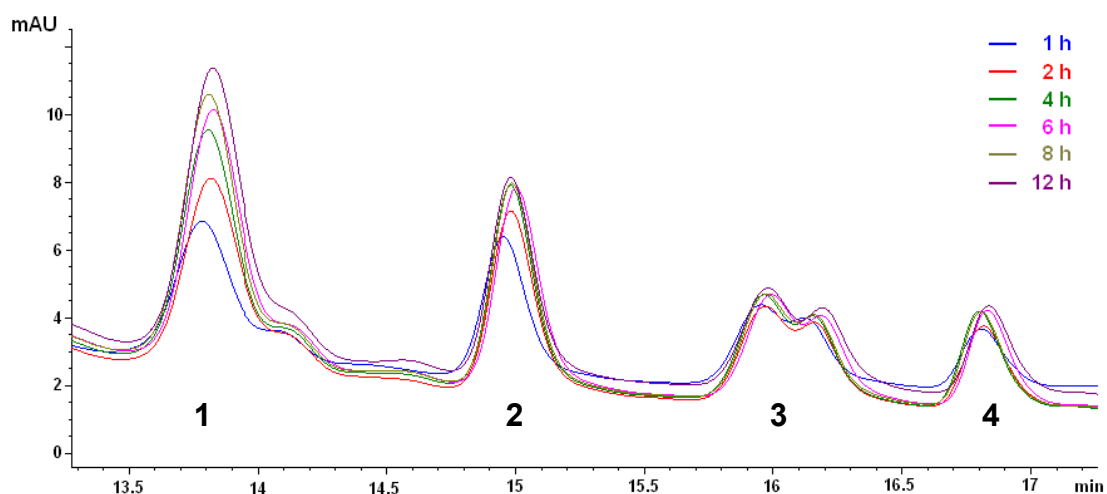


Figure 12: HPLC trace showing the distribution of acylhydrazone products over time to establish when equilibrium is reached. Experimental conditions: **HZD01** to **HZD04** (final conc. = 60 μ M each, DMSO), **ALD01** (60 μ M, water), aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (1.3 mL, 100 mM, pH 6.4) mixed at room temperature. 10 μ L of sample injected directly onto HPLC column at several time points. Products detected by UV at $\lambda = 254$ nm. HPLC gradient A (see chapter 5, section 5.2.1). Peak 1 = **HZN01-01-01**, 2 = **HZN01-04-01**, 3 = **HZN01-03-01**, 4 = **HZN01-02-01**.

The order of elution was established in three ways. Firstly individual mixtures were set up using one hydrazide with one aldehyde, as shown in Figure 13. Secondly samples of the acylhydrazone products that were prepared and isolated in the lab were run under the same chromatography conditions, see Appendix Figure 1. Finally the mixture was run through LC/MS with the mass of the four product peaks detected in selected ion monitoring (SIM) mode, see Appendix Figures 2-5.

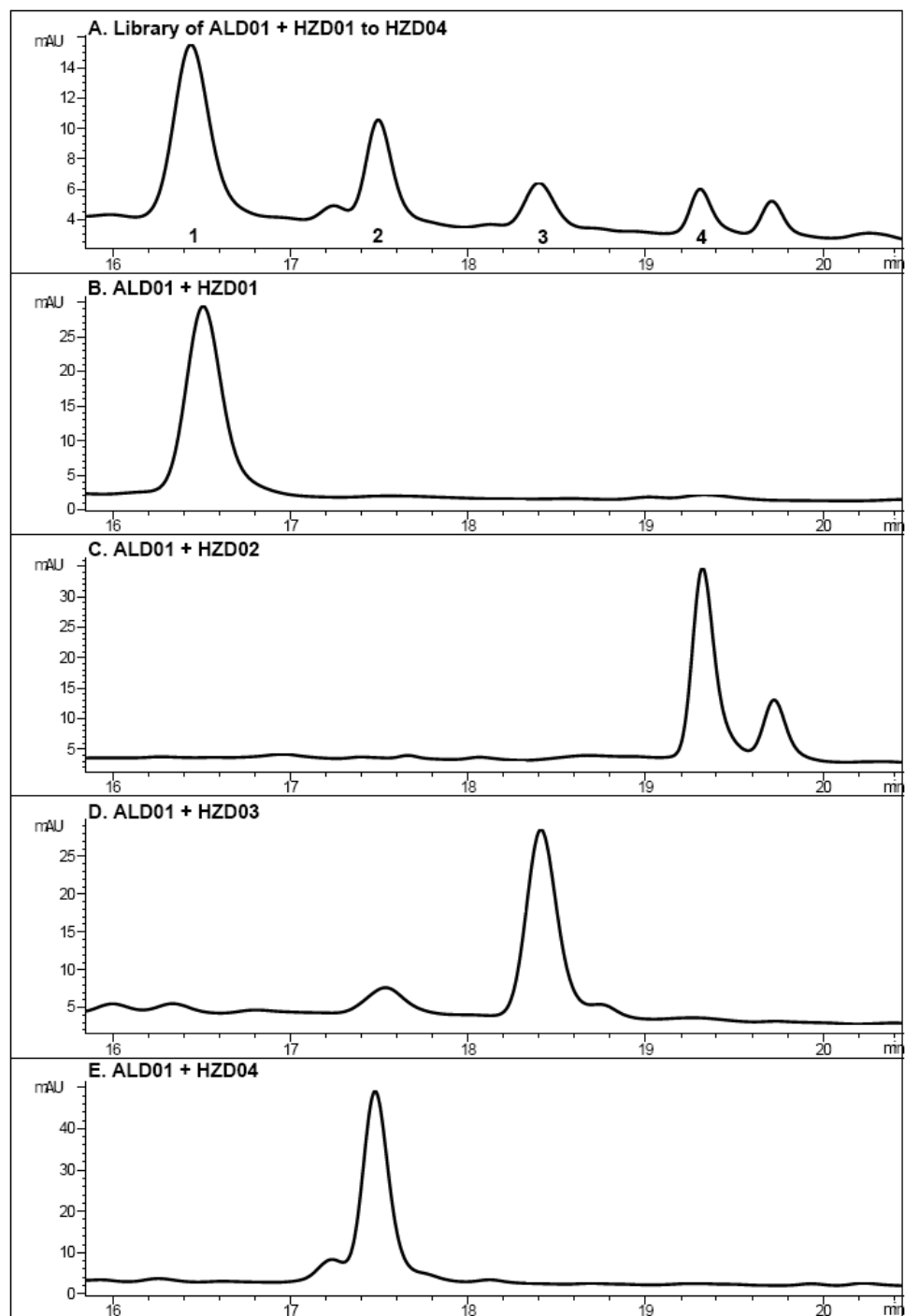


Figure 13: HPLC traces showing references to determine the products in the equilibrium based on individual reactions between **ALD01** and one hydrazide spacer. **A)** Full library of **ALD01** and **HZD01** to **HZD04**, **B)** reaction between **ALD01** and **HZD01**, **C)** reaction between **ALD01** and **HZD02**, **D)** reaction between **ALD01** and **HZD03**. **E)** reaction between **ALD01** and **HZD04**. Products are detected by UV detecting at $\lambda = 254$ nm. HPLC gradient A (see chapter 5, section 5.2.1). Peak 1 = **HZN01-01-01**, 2 = **HZN01-04-01**, 3 = **HZN01-03-01**, 4 = **HZN01-02-01**.

The conditions of the DCL, in particular the optimum concentration of aniline, were established previously in the lab.^[31] In order to confirm that **DCL01** was dynamic, the DCL was set up starting from one of the acylhydrazones **HZN01-04-01** and the other three hydrazide spacers **HZD01** to **HZD03**. No free aldehyde was added to the reaction mixture. If the DCL was truly dynamic, the same product distribution should still be reached. Several days were needed to reach the final equilibrium compared with the several hours it takes to reach equilibrium starting with aldehyde and hydrazide. The peak corresponding to the starting acylhydrazone **HZN01-04-01** decreases over time, as shown in Figure 14 and after ten days the formation of **HZN01-01-01**, **HZN01-02-01** and **HZN01-03-01** could be clearly identified, however the final equilibrium had still not been achieved. A possible explanation for this observation is described later in the chapter.

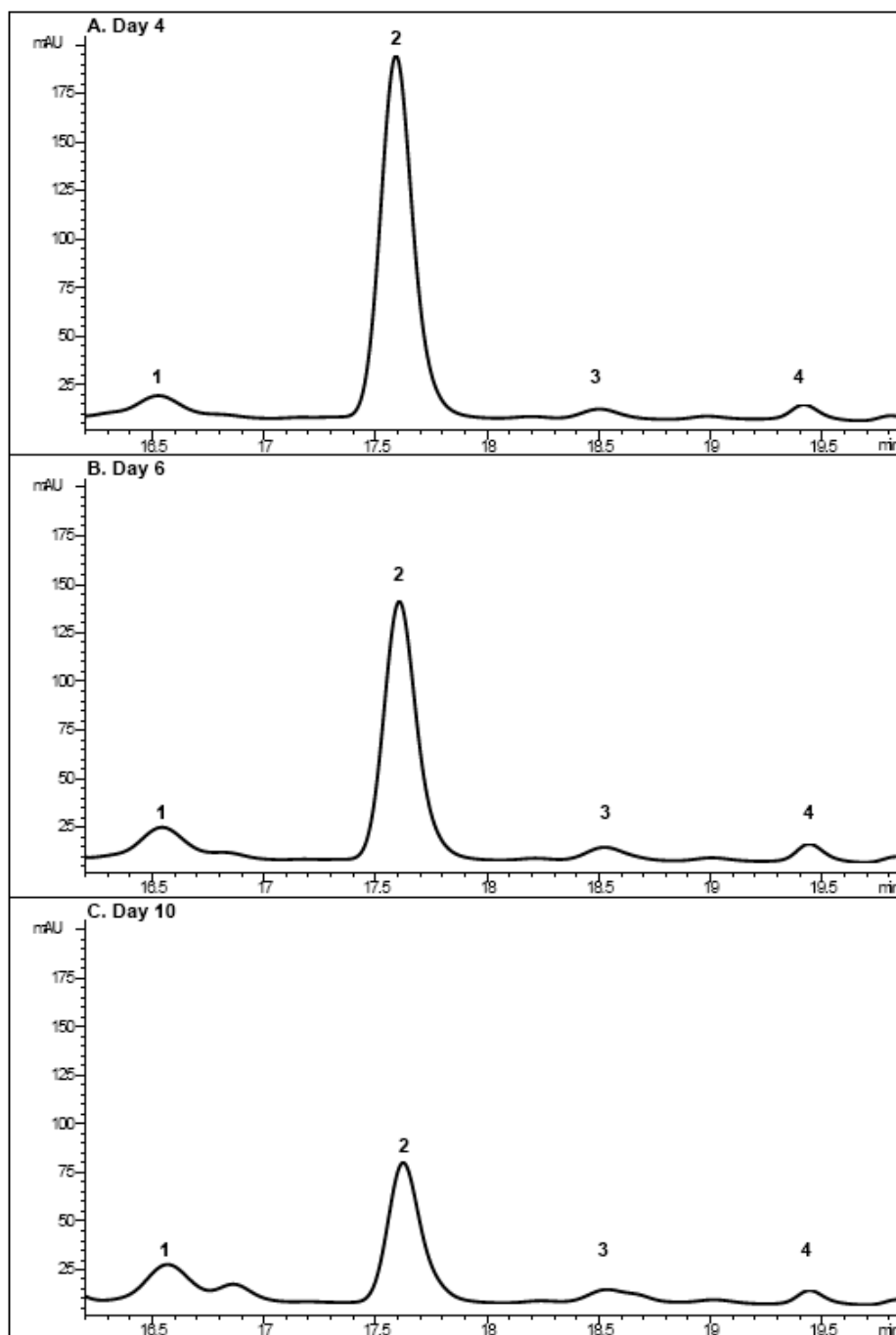


Figure 14: HPLC traces showing time course data for DCL library starting from **HZN01-04-01** (peak 2) and **HZD01**, **HZD02** and **HZD03**. **A)** Day 4, **B)** Day 6, **C)** Day 10. Experimental conditions: **HZD01** to **HZD03** (final conc. = 60 μ M each, DMSO), **HZD01-04-01** (60 μ M, water), aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (1.3 mL, 100 mM, pH 6.4) mixed at room temperature. 10 μ L of sample injected directly onto HPLC column at several time points. Products detected by UV at $\lambda = 254$ nm. HPLC gradient A (see chapter 5, section 5.2.1). Peak 1 = **HZN01-01-01**, 2 = **HZN01-04-01**, 3 = **HZN01-03-01**, 4 = **HZN01-02-01**.

With the equilibrium established and the products identified, the DCL was set up in the presence of the biological template GST, using the isoform mGSTM1-1. This was added both at the initial mixing of the library components and also after the equilibrium was reached. Conditions were again chosen based on previous work in the lab.^[31] It was shown in this work that GST would remain active under the DCL conditions and could tolerate 15 % by volume of DMSO and 5 mM aniline.

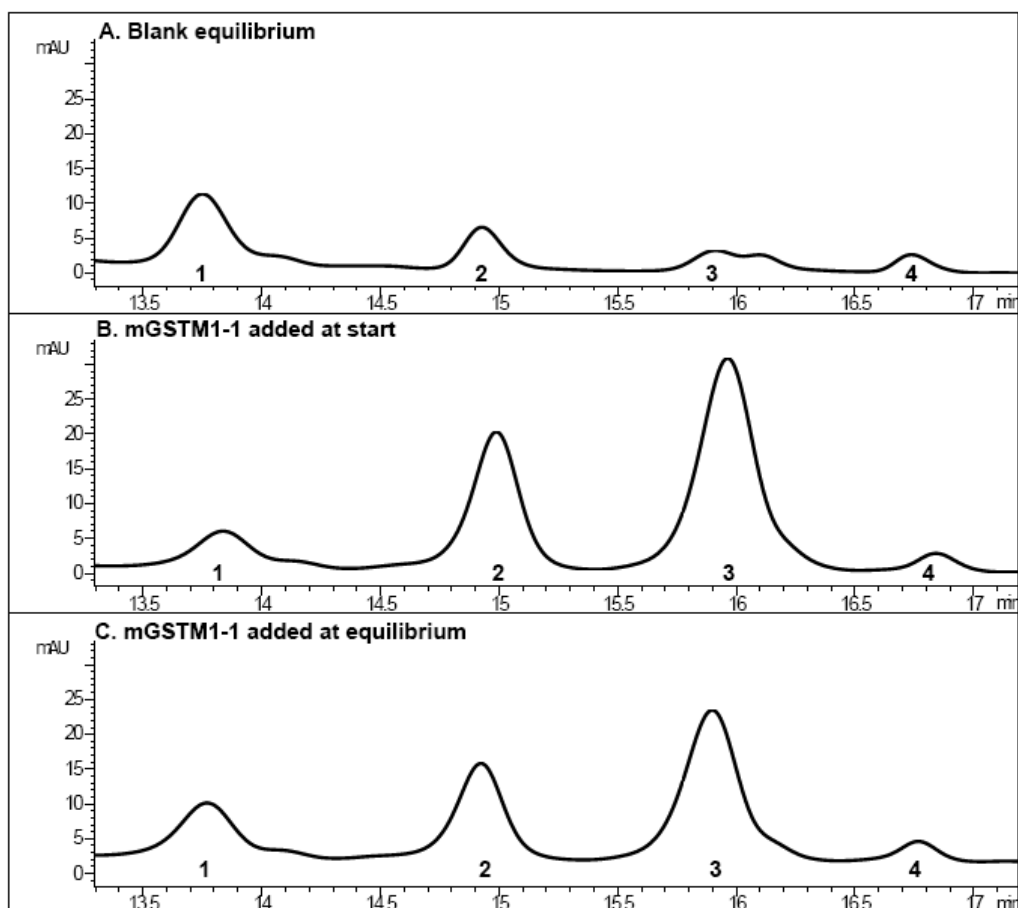


Figure 15: HPLC traces showing **DCL01** in the absence and presence of the biological template. Experimental conditions: **HZD01** to **HZD04** (final conc. = 60 μ M each, DMSO), **ALD01** (60 μ M, water), aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (255 μ L, 100 mM, pH 6.4) mixed at room temperature. After 48 h, 10 μ L of sample injected directly onto HPLC column. **A)** Blank equilibrium, **B)** mGSTM1-1 (final conc. = 20 μ M) added initially, **C)** mGSTM1-1 (final conc. = 20 μ M) added at equilibrium. Products detected by UV at $\lambda = 254$ nm. HPLC gradient A (see chapter 5, section 5.2.1). Peak 1 = **HZN01-01-01**, 2 = **HZN01-04-01**, 3 = **HZN01-03-01**, 4 = **HZN01-02-01**.

In the presence of the biological template mGSTM1-1, a clear amplification was observed over the background equilibrium, Figure 15. Comparison of the

GST-templated reaction with the blank equilibrium showed a 2000 % increase in the amplification of **HZN01-03-01**. A smaller amplification was also observed for **HZN01-04-01**. The same amplification was observed whether the GST was added initially or after the blank equilibrium had been reached, although it took a longer time for the reaction to re-equilibrate in the latter case. The fact that the same final equilibrium could be reached regardless of when the protein was added was further proof that the library was dynamic. The same amplifications were observed when the reaction was monitored using LC/MS in SIM mode detecting for bis acylhydrazone product mass, see Appendix Figure 6.

2.2.2 Isoform specificity in the initial dynamic combinatorial library

With the conditions for the protein-templated DCL established, four different isoforms of GST were investigated; mGSTM1-1, hGSTP1-1, mGSTA4-4 and SjGST.

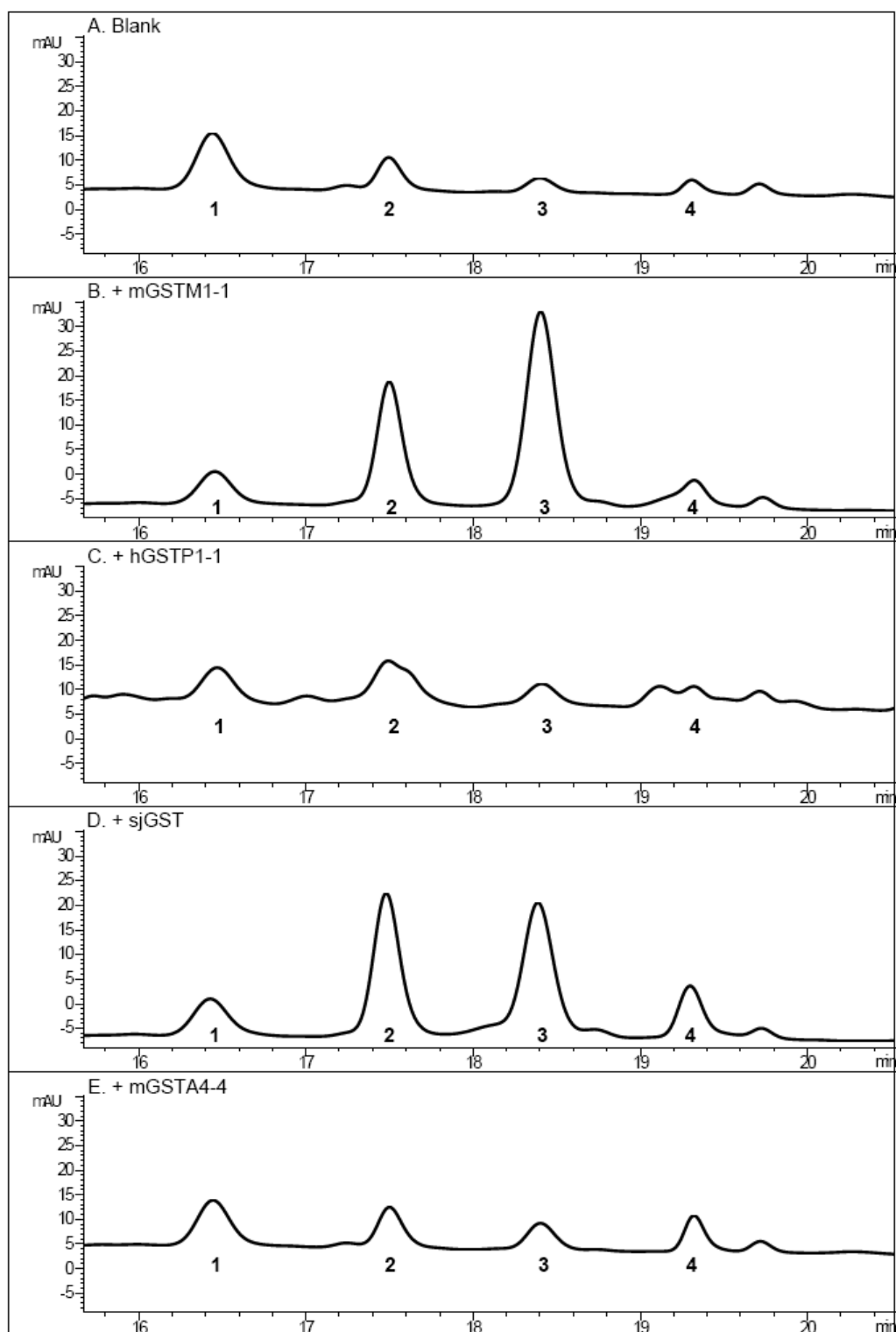


Figure 16: HPLC traces showing four hydrazide with one aldehyde DCL in presence of GST isoforms. Experimental conditions: **HZD01** to **HZD04** (final conc. = 60 μ M each, DMSO), **ALD01** (60 μ M, water), aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (255 μ L, 100 mM, pH 6.4) mixed at room temperature. After 48 h, 10 μ L of sample injected directly onto HPLC column. **A)** blank equilibrium, **B)** mGSTM1-1 (final

conc. = 20 μ M), **C**) hGSTP1-1 (final conc. = 20 μ M), **D**) SjGST (final conc. = 20 μ M), **E**) mGSTA4-4 (final conc. = 20 μ M). Products detected by UV at λ = 254 nm. HPLC gradient A (see chapter 5, section 5.2.1). Peak 1 = **HZN01-01-01**, 2 = **HZN01-04-01**, 3 = **HZN01-03-01**, 4 = **HZN01-02-01**.

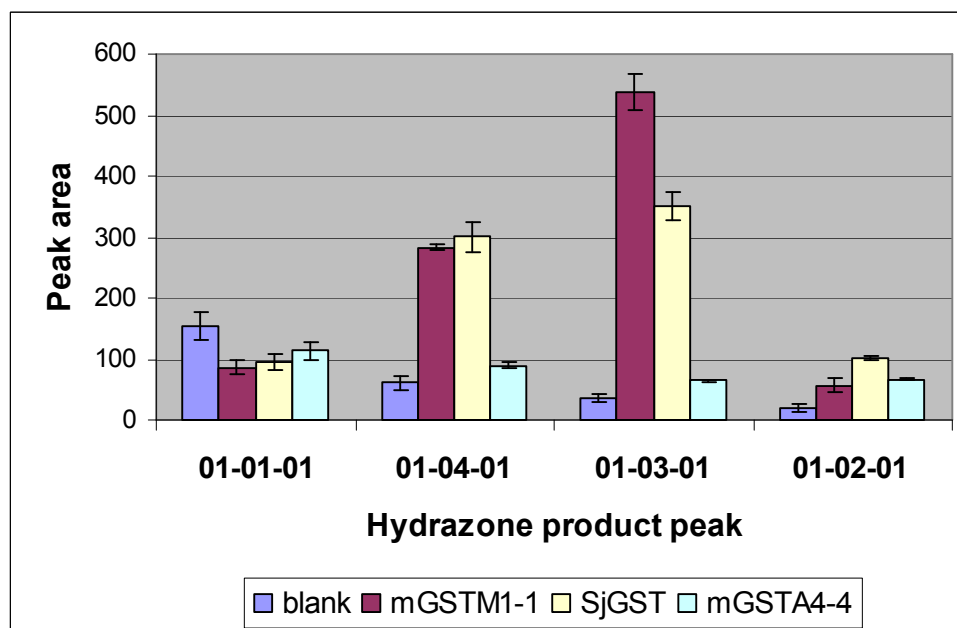


Figure 17: Graphical representation of hydrazone product distribution in the presence of different GST isoforms. Error bars represent standard deviation from two experiments.

The amplifications observed in the presence of SjGST followed the same pattern as for mGSTM1-1 with acylhydrazone **HZN01-03-01** giving the largest amplification, although the amplification was less than in the presence of mGSTM1-1. When mGSTA4-4 was added there appeared to be little observed change to the equilibrium. The only significant amplification over the blank equilibrium was for **HZN01-02-01**, but this was a much smaller amplification than those observed with mGSTM1-1 and SjGST. The addition of hGSTP1-1 appeared to disrupt the equilibrium without amplifying any of the bis acylhydrazone product peaks. The baseline became much less well defined suggesting that it was actually the monoacylhydrazone products that were favoured by the enzyme. The disruption to the baseline made it very difficult to accurately measure the bis acylhydrazone peak area, so no comparison with the blank equilibrium could be made. The monoacylhydrazone products proved very difficult to observe under HPLC conditions as they were often seen as broad peaks. The peaks marked with * on

Figure 18 show the monoacylhydrazone peaks. Figure 18 also shows the appearance of several new peaks marked with **. These peaks could not be identified as anything in the reaction mixture and could indicate that the acylhydrazone spacers were decomposing. It is unclear how the hGSTP1-1 could have done this. To further investigate, these samples were run on LC/MS. Samples were monitored for the four monoacylhydrazone mass peaks and also the four bis acylhydrazone peaks in SIM mode, see Appendix Figure 7. This data confirmed that the monoacylhydrazones were being amplified to a greater extent than any of the bis acylhydrazones by hGSTP1-1 suggesting that the 3D structure of the spacer did not lie in the correct orientation to bridge the space between the two active sites of the dimer. The new peaks observed in the HPLC trace of the hGSTP1-1-templated DCL could not be identified in the full mass scan of the same mixture as the DMSO masked the rest of the spectrum.

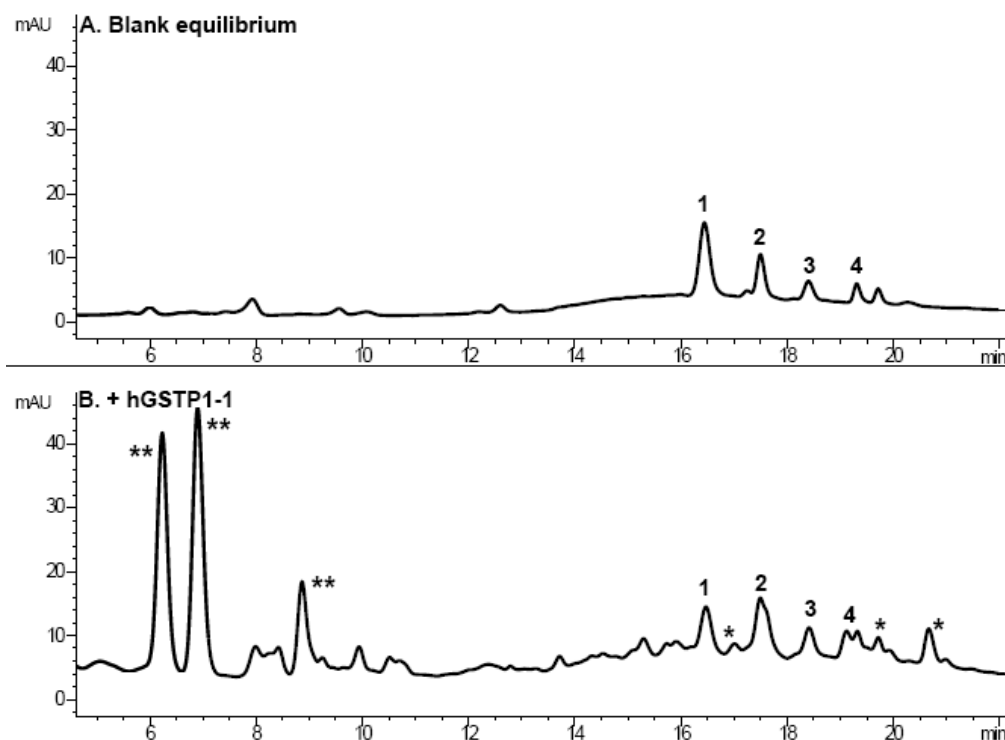
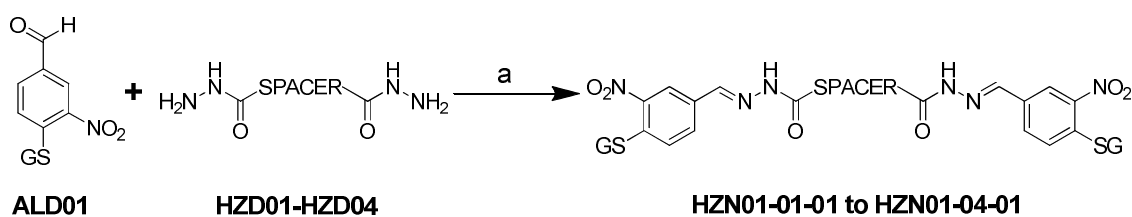


Figure 18: HPLC trace of **A)** non-templated equilibrium of **DCL01** and **B)** hGSTP1-1 templated **DCL01**. Products detected by UV at $\lambda = 254$ nm. HPLC gradient A (see chapter 5, section 5.2.1). Peak 1 = **HZN01-01-01**, 2 = **HZN01-04-01**, 3 = **HZN01-03-01**, 4 = **HZN01-02-01**. * = monoacylhydrazone, ** = unknown peak.

2.2.3 Synthesis of DCL01 products

The four bis acylhydrazone products were synthesised by the reaction of aldehyde **ALD01** with hydrazide spacer **HZD01** to **HZD04** under acidic conditions. The reaction was very straightforward, with the yellow product precipitating from the solution. Characterisation of the products proved to be much more challenging. The products were run through LC/MS which showed that they were pure and had the expected mass for the bis acylhydrazone. The products were confirmed by high resolution mass spectrometry as NMR data could not always be fully determined due to problems with solubility of the acylhydrazones.



Scheme 3: Synthesis of acylhydrazone products from **DCL01**. a; glacial acetic acid (few drops), water, rt, 30 min.

2.2.4 Binding data for DCL01 products

To confirm that the amplifications observed corresponded to the best inhibitors the IC_{50} values for the four acylhydrazone products were calculated using the CDNB assay.^[7]

Table 1: Table of IC₅₀ data of DCL01 products with different GST isoforms measured using CDNB assay. Values are calculated from three separate assays. Raw data described in experimental chapter.

Acyldiazone	IC ₅₀ mGSTM1-1 (μ M)	IC ₅₀ hGSTP1-1 (μ M)	IC ₅₀ SjGST (μ M)	IC ₅₀ mGSTA4-4 (μ M)
01-01-01	1.207	126.5	3.471	> 100
01-02-01	0.337	11.81	0.252	> 100
01-03-01	0.050	13.45	0.989	> 100
01-04-01	0.413	0.356	1.800	> 100
Aldehyde 01	341.7	\geq 500	265.6	> 500
Hydrazide 03	> 500	> 500	-	-

The data correlated well with the amplifications observed in the DCL. In the case of mGSTM1-1, the product with the greatest amplification was **HZN01-03-01** and this gave almost 10-fold greater inhibition of GST activity over the other acyldiazone products. Gratifyingly, excellent isoform specificity was observed for **HZN01-03-01** for mGSTM1-1. **HZN01-03-01** for mGSTM1-1 gave approximately twenty-fold greater inhibitory activity over SjGST and 270-fold greater inhibitory activity over hGSTP1-1. Interestingly, **HZN01-04-01** also showed some amplification in the DCL, however it appeared to have no greater inhibition of the enzyme than **HZN01-02-01**. Additionally, when looking at the data for SjGST, the inhibition data suggested that **HZN01-02-01** was the best inhibitor, however in the DCL the longer spacers **HZN01-03-01** and **HZN01-04-01** had a greater amplification. The best IC₅₀ value for hGSTP1-1 was for **HZN01-04-01** with a value similar to that for mGSTM1-1. The fact that **HZN01-04-01** was the longest of the spacers and only a small amplification of this product could be observed in the HPLC trace might suggest that the spacers being used in the library were not long enough to span the two active sites in hGSTP1-1. The bivalent spacers described by the Atkins group provide further evidence that spacers of a longer length display greater binding affinity for hGSTP1-1.^[32, 34] This would explain why the mono acyldiazones were favoured, as the second aldehyde might not fit well into the

solvent accessible cleft. For all GST isoforms there was a large increase in inhibitory activity when the two units of the aldehyde **ALD01** were connected by a spacer, regardless of the length or shape of the spacer compared to the single aldehyde unit. The inhibition data for mGSTA4-4 shows that the acylhydrazone products had little effect on the enzyme activity, explaining the minimal effect the enzyme had on the equilibrium of the DCL. This made the mGSTA4-4 isoform an excellent negative control, showing that the change in the distribution of acylhydrazone products was due to a specific binding site within the mGSTM1-1 isoform and not due to any non-specific effects from the introduction of the hydrophobic protein to the DCL.

2.2.5 Controls for protein-templated DCL01

As the most successful of the isoforms used, further controls were set up to confirm the amplification observed with mGSTM1-1. In the presence of bovine serum albumin (BSA) a protein without the binding pockets found in GSTs, no amplification was observed, Figure 19B.

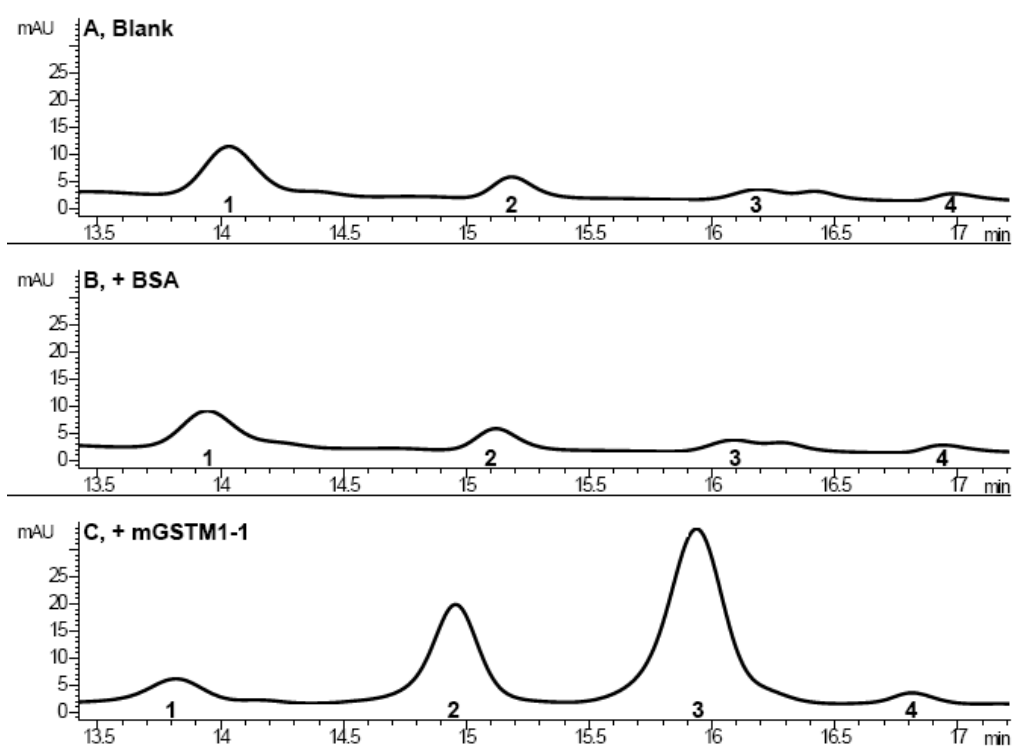


Figure 19: HPLC traces of **DCL01** in the presence of control protein BSA. Experimental conditions: **HZD01** to **HZD04** (final conc. = 60 μ M each, DMSO), **ALD01** (60 μ M, water), aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (255 μ L, 100 mM, pH 6.4) mixed at room temperature. After 48 h, 10 μ L of sample injected directly

onto HPLC column. **A)** blank equilibrium, **B)** BSA (final conc. = 20 μ M), **C)** mGSTM1-1 (final conc. = 20 μ M). Products detected by UV at $\lambda = 254$ nm. HPLC gradient A (see chapter 5, section 5.2.1). Peak 1 = **HZN01-01-01**, 2 = **HZN01-04-01**, 3 = **HZN01-03-01**, 4 = **HZN01-02-01**.

To determine where exactly in the binding pocket of mGSTM1-1 the acylhydrazone product was binding, the DCLs were set up in the presence of a variety of known GST inhibitors which bind in different areas of the hydrophobic binding pocket.

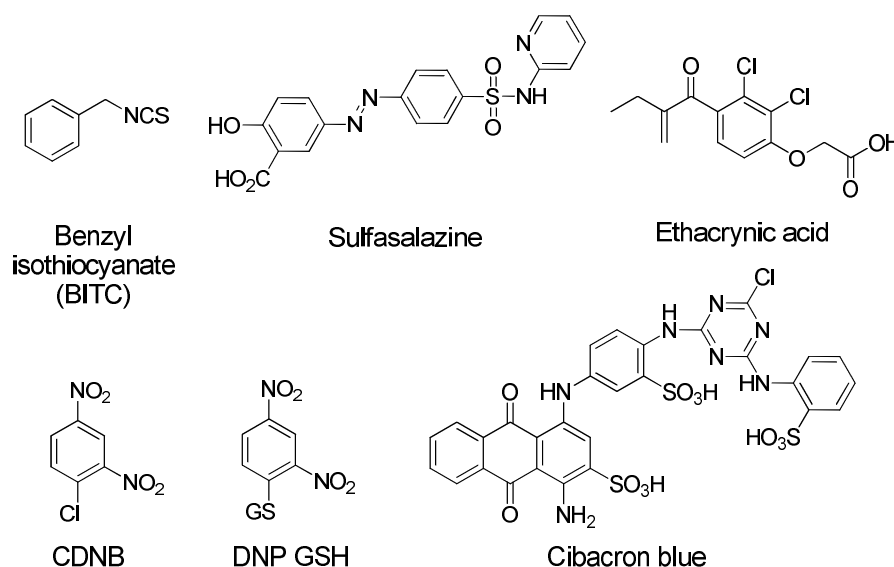


Figure 20: Known inhibitors and substrates of GST used to block the binding pocket in **DCL01**.

Benzyl isothiocyanate (BITC) binds in the H-site of GSTs and is also known to covalently modify the active site. It has a different, distinct binding site to the substrate CDNB.^[36] Sulfasalazine also binds in the H-site and is a competitive inhibitor of CDNB.^[19] The crystal structure of sulfasalazine bound to hGSTP1-1 also shows that the pyridinyl ring binds in a shallow pocket on the surface of the protein. Ethacrynic acid is a non-specific inhibitor of GSTs, which binds in the H-site. There is evidence that ethacrynic acid can bind in more than one orientation in the binding pocket.^[37, 38] The final two compounds selected were the universal substrate, CDNB and the glutathione-conjugated product of CDNB, DNP GSH. DNP GSH is believed to bind in a different position to CDNB within the H-pocket.^[19] Cibacron blue, a dye, which also binds to the ligandin site of GSTs,^[19] was found to disrupt the background equilibrium and could not be used as a control.

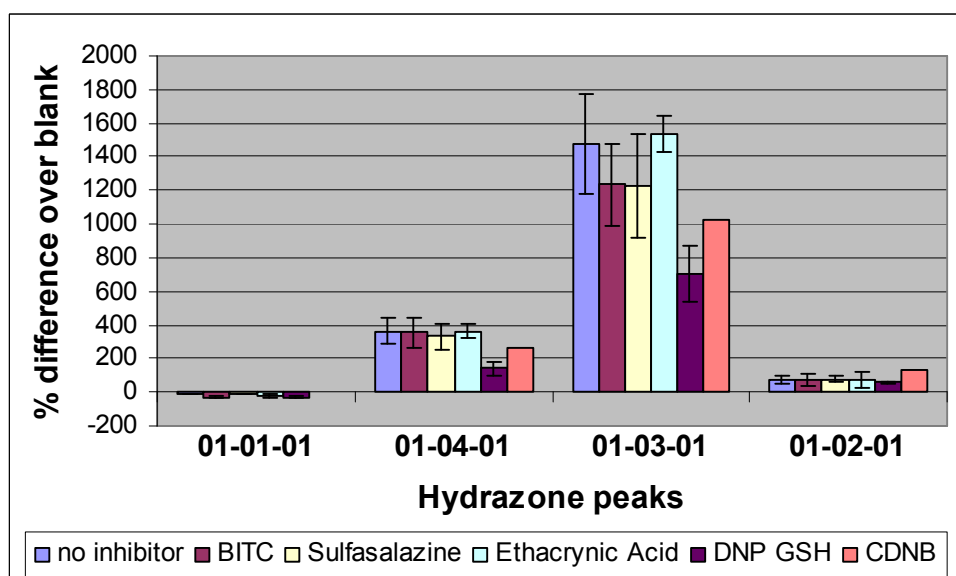


Figure 21: Graphical representation of the effect on **DCL01** of added inhibitors of GST. Bars represent the % difference in the peak area over the blank equilibrium. Error bars represent standard deviation from two experiments. The data for the reaction with the addition of CDNB is the result of only one experiment.

Although none of the inhibitors completely reduced the amplification, the best effects were seen with DNP GSH. Adding increasing amounts of DNP GSH showed a dose-dependent response on the reduction of the amplification.

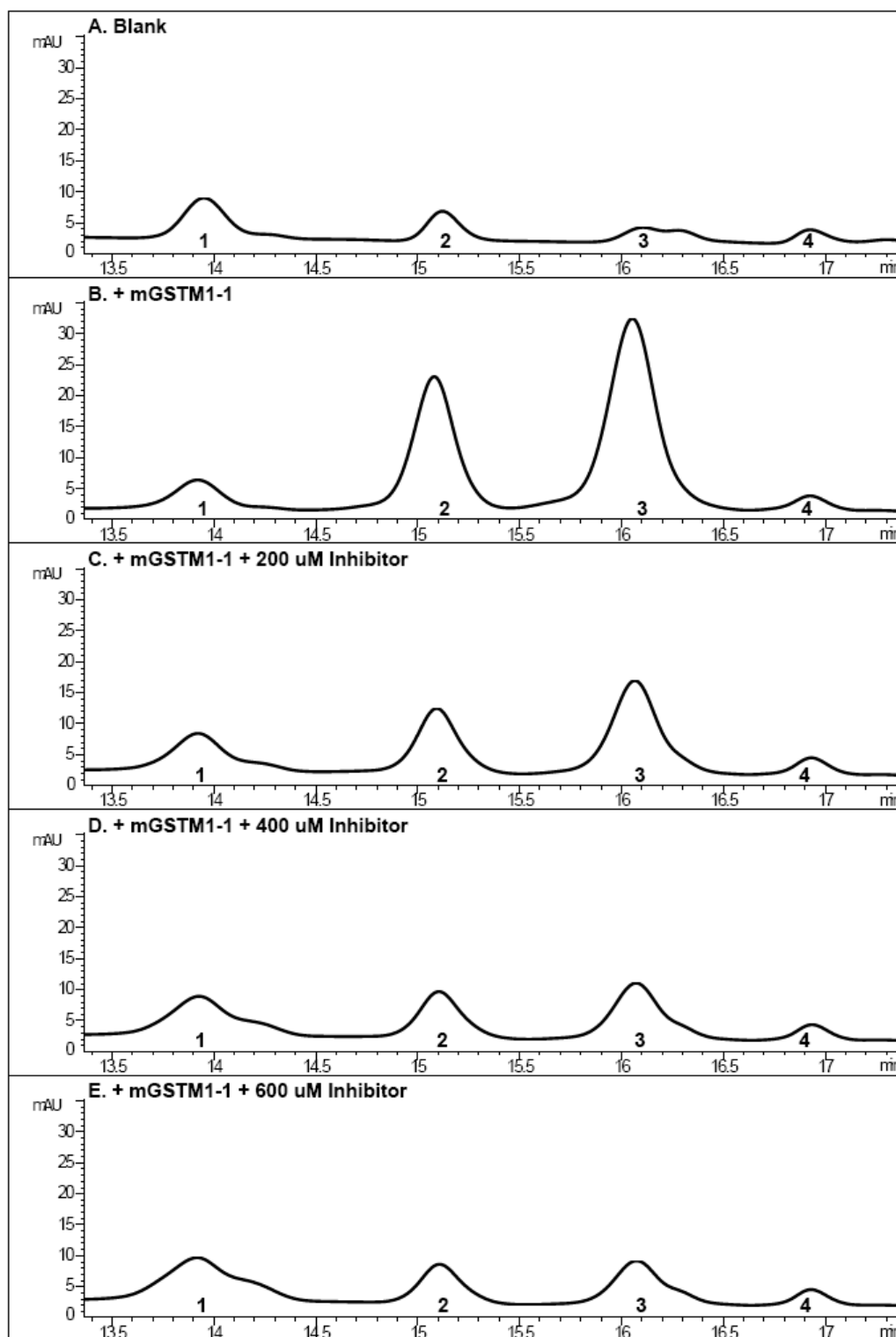


Figure 22: HPLC traces of **DCL01** in the presence of increasing concentrations of GST conjugation product DNP GSH. Experimental conditions: **HZD01** to **HZD04** (final conc. = 60 μ M each, DMSO), **ALD01** (60 μ M, water), aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (255 μ L, 100 mM, pH 6.4) mixed at room temperature. After 48 h, 10 μ L of sample injected directly onto HPLC column. **A)** blank equilibrium,

B) mGSTM1-1 (final conc. = 20 μ M), **C)** mGSTM1-1 (final conc. = 20 μ M) + 200 μ M DNP GSH, **D)** mGSTM1-1 (final conc. = 20 μ M) + 400 μ M DNP GSH, **E)** mGSTM1-1 (final conc. = 20 μ M) + 600 μ M DNP GSH. Products are detected by UV at λ = 254 nm. HPLC gradient A (see chapter 5, section 5.2.1). Peak 1 = **HZN01-01-01**, 2 = **HZN01-04-01**, 3 = **HZN01-03-01**, 4 = **HZN01-02-01**.

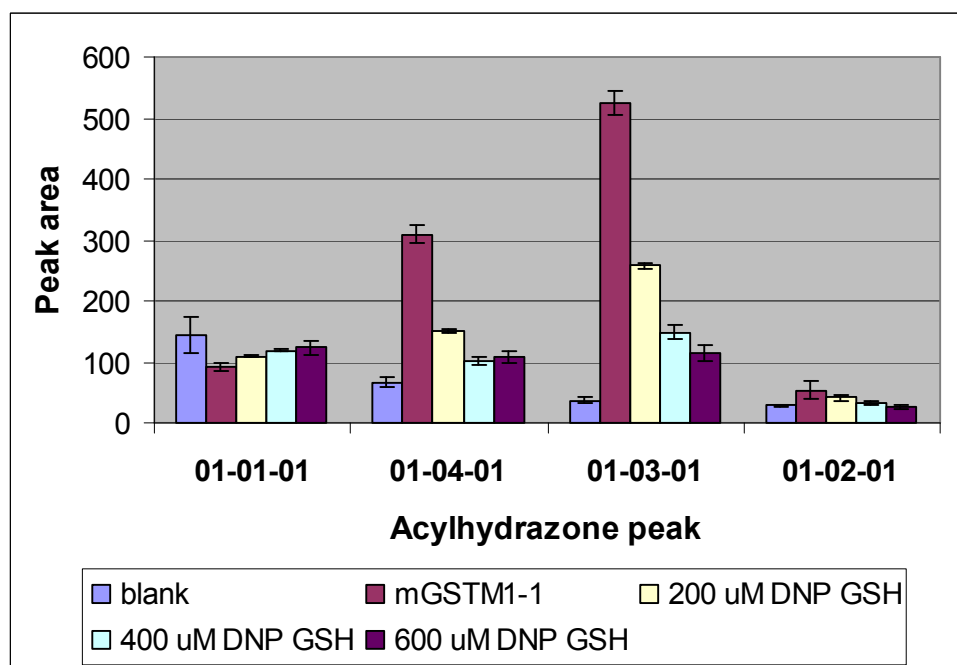


Figure 23: Graphical representation of the effect on **DCL01** of increasing concentrations of GSDNB. Error bars represent standard deviation from two experiments.

The result of using these inhibitors suggested that the acylhydrazone products were binding in the active site of the GST. There must be a very specific binding orientation for the bivalent acylhydrazones within the H-site as the addition of CDNB, sulfasalazine or ethacrynic acid did not prevent amplification of the DCL. The H-site is a very large, promiscuous pocket. Work studying the binding of CDNB during the conjugation reaction has shown that in the transition state the phenyl ring of CDNB binds with two ring stacking interactions in the H-site from the side-chains of Phe8 and Tyr108. This is known as the “in” binding mode. When glutathione is conjugated to CDNB the phenyl ring undergoes a twist and loses one of these interactions, leaving just the interaction with Tyr108. The phenyl ring is then positioned along the edge of the hydrophobic pocket in the “out” binding mode. This change in binding mode has been observed for both the rat and human form of

GSTM1-1 and the human form of GSTP1-1 and is demonstrated for the hGSTM1-1 isoform in Figure 24.^[17, 19, 39, 40]

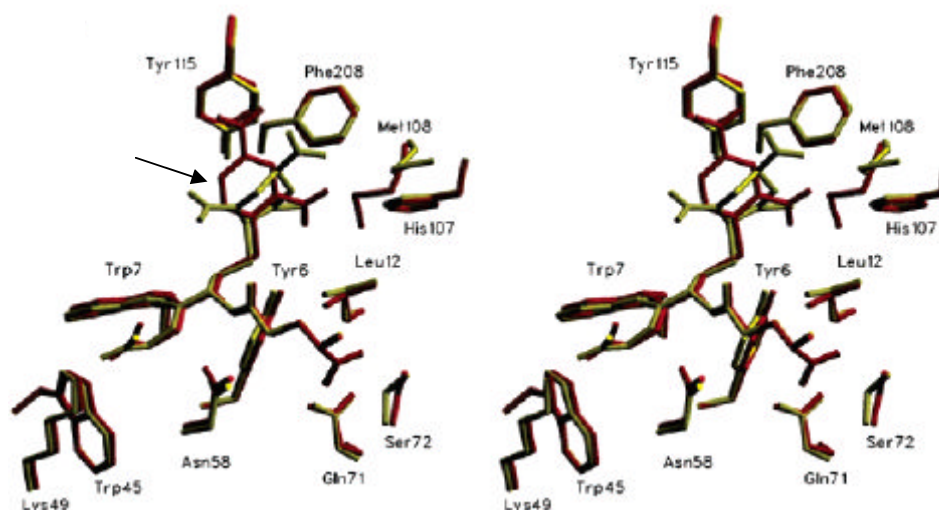


Figure 24: Overlay of the crystal structures of hGSTM1-1 with the transition state analogue 1-(S-glutathionyl)-2,4,6-trinitrocyclohexadienate anion (Yellow, PDB: 2F3M) and with the conjugate product DNP GSH (Red, PDB: 1XWK). Figure taken from Patskovsky *et al.*^[17] The twist of the phenyl ring is highlighted with an arrow.

The bivalent acylhydrazone compounds appear to be binding in the “out” mode, as it was only DNP GSH that reduced amplification in the DCL. The results from the DCL suggest that there is space within the H-site for a new molecule of CDNB to bind before the conjugated product leaves the enzyme, as CDNB does not block the amplification in the DCL. There is some precedence for inhibitors binding in the H-site, without being competitive inhibitors of CDNB.^[19, 28, 41]

2.2.6 Further binding studies for DCL01 products - competition assay and isothermal titration calorimetry

A competition assay between CDNB and the three best binding acylhydrazones **HZN01-02-01**, **HZN01-03-01** and **HZN01-04-01** was set up and the results showed that the acylhydrazones were noncompetitive partial inhibitors of CDNB, Figure 25. This result provided further evidence that the acylhydrazone product and CDNB were not binding in the same place and both can bind to GST at the same time.

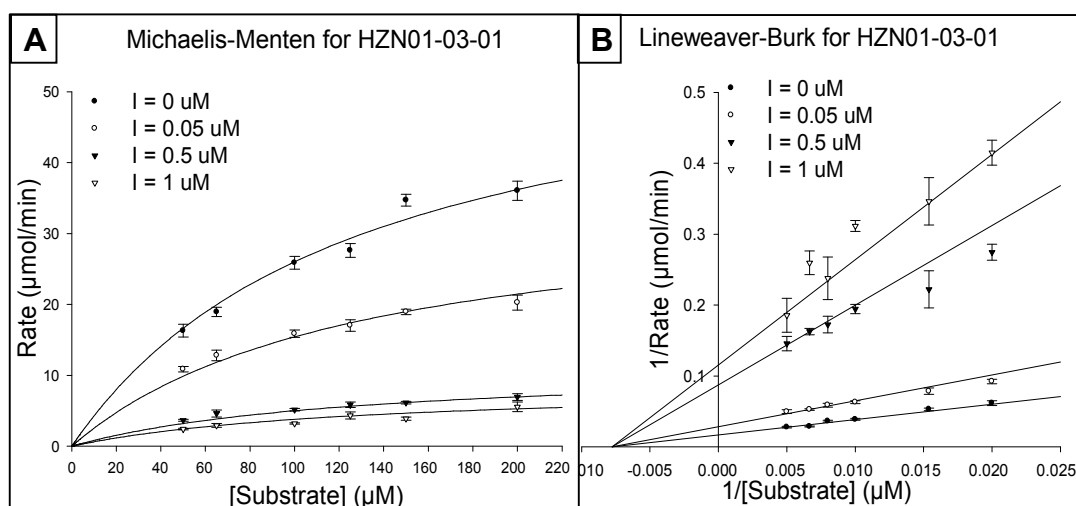


Figure 25: Michaelis-Menten plot and Lineweaver-Burk plot for **HZN01-03-01** with mGSTM1-1. Data best fit to noncompetitive (partial) equation. Standard error as a result of 3 repeat experiments.

Isothermal titration calorimetry (ITC) was used to determine the binding stoichiometry of the best acylhydrazone **HZN01-03-01**, as shown in Figure 26. The results gave a binding stoichiometry of 0.4 with respect to the GSTM1 monomer concentration, confirming that the spacer was binding across the dimer GST.

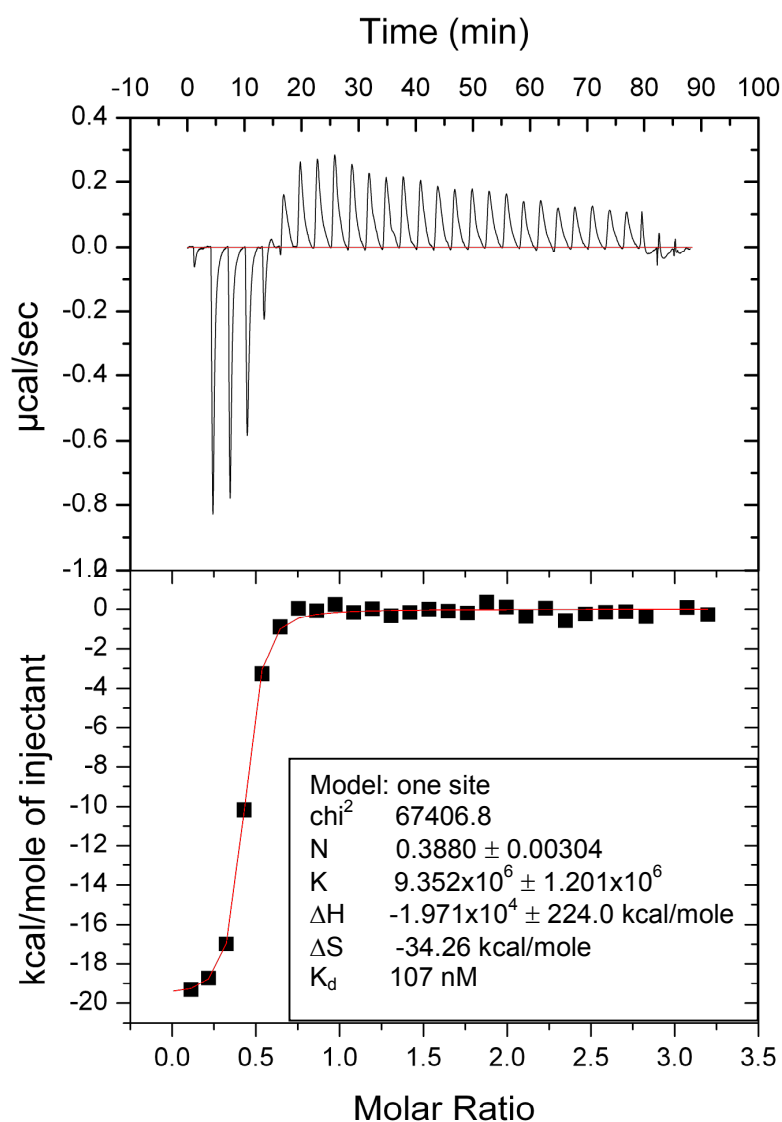


Figure 26: ITC data for **HZN01-03-01** with mGSTM1-1. Top panels show the heat effects associated with the injection of ligand into the cell containing GST. Bottom panels show the binding isotherm, generated from the integrated heats in the top panel, and the best fitted curves.

In an interesting aside, the ITC data also showed that the dilution of ligand into buffer was an endothermic process. This indicated that there was a degree of hydrogen bonding between the acylhydrazone ligands, which was disrupted upon dilution. This did not seem to affect the inhibition capability of the ligand. An overlay of the trace for ligand into buffer over ligand into protein clearly showed that the endothermic reaction of the ligand dilution was a separate event to ligand binding to protein, Figure 27. When the energy changes associated with the ligand binding to the protein have ceased the trace overlays perfectly with the background energy

changes associated with the ligand dilution. This phenomenon was also observed for the bivalent compounds with glutathione at the terminal described by the Atkins group.^[32]

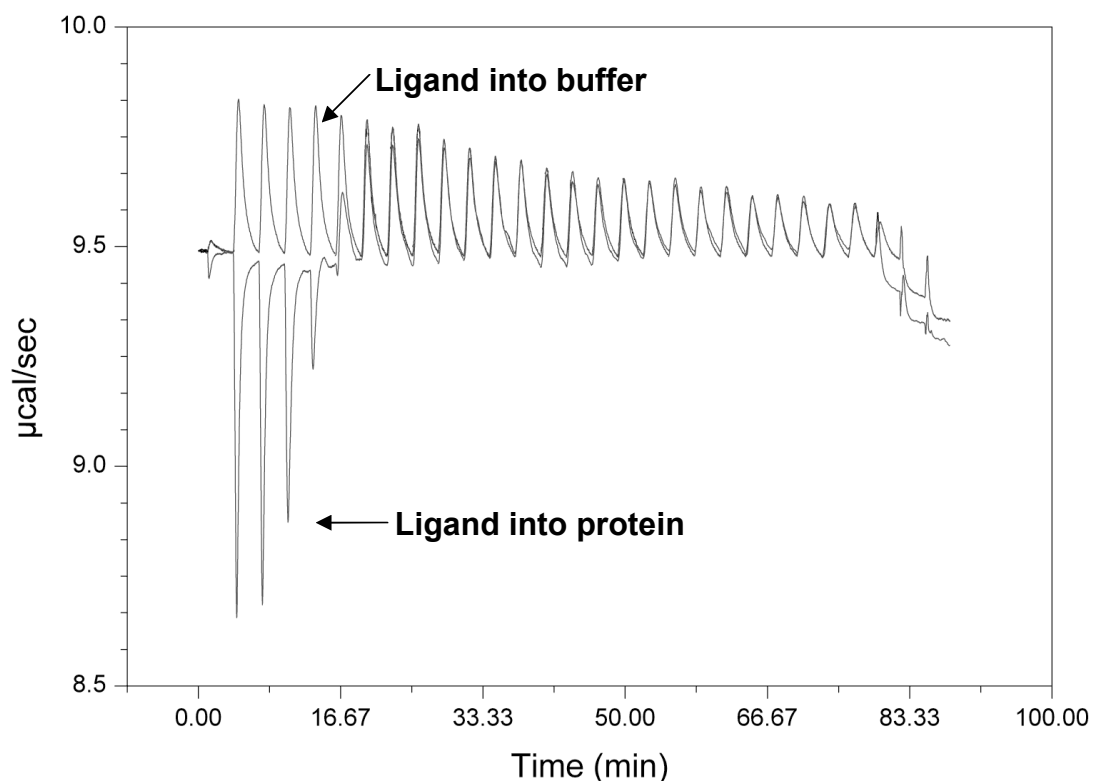


Figure 27: Overlay of the raw transpose from the ITC data for **HZN01-03-01** with mGSTM1-1 showing the heat effects associated with the injection of ligand into buffer and the heat effects associated with the injection of ligand into GST (labelled).

The Cheng-Prusoff equation, Equation 1, describes the relationship between the inhibition constant, (K_i) and the IC_{50} value for a substrate.^[42]

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

Equation 1: Cheng-Prusoff equation. [S] = concentration of substrate.

According to the Cheng-Prusoff equation, in the case of noncompetitive inhibitors, $IC_{50} = K_i$.^[42] The data for **HZN01-03-01** and **HZN01-04-01** in Table 2 fit this equation, confirming the observation that the acylhydrazone compounds were noncompetitive inhibitors with respect to the substrate CDNB.

Table 2: Binding data for the products of **DCL01** with mGSTM1-1. Standard error shown for K_i data taken from 3 replicates.

Acylhydrazone	IC ₅₀ (nM)	K _i (nM)	K _d (nM)
01-01-01	1207	-	-
01-02-01	337	646 ± 80.7	525
01-03-01	50	61 ± 4.1	107
01-04-01	413	634 ± 23.8	-

2.3 Problems with HZN01-02-01

Perhaps the most surprising result was the binding data for the shortest acylhydrazone **HZN01-02-01**. This spacer should not be long enough to span the distance between the two active sites within the GST dimer, yet the IC₅₀ value was very favourable when compared with other monovalent GST inhibitors based on a glutathione conjugated aromatic scaffold. A DCL has previously been set up in the Greaney lab, based on **ALD01** with a library of aromatic hydrazides designed to explore the hydrophobic binding pocket.^[31] The best of the compounds from the library **4** and **5** had IC₅₀ values of 22 μM for SjGST and 57 μM for hGSTP1-1 respectively, as shown in Figure 28. This is a five-fold reduction in inhibition for hGSTP1-1 and a ten-fold reduction for SjGST when compared to bivalent inhibitor **HZN01-02-01**. In the case of mGSTM1-1 the compounds in the library showed little inhibitory effect with the IC₅₀ for all compounds being greater than 100 μM. The monovalent acylhydrazones were competitive inhibitors of CDNB, however the bis acylhydrazone **HZN01-02-01** was a noncompetitive inhibitor of CDNB. Molecular docking studies of the monovalent compounds suggest that they bind into the hydrophobic pocket, explaining why they are competitive inhibitors. The fact that **HZN01-02-01** was noncompetitive indicated that the compound was pointing out into the solvent accessible cleft, in a similar manner to the longer length bivalent compounds.

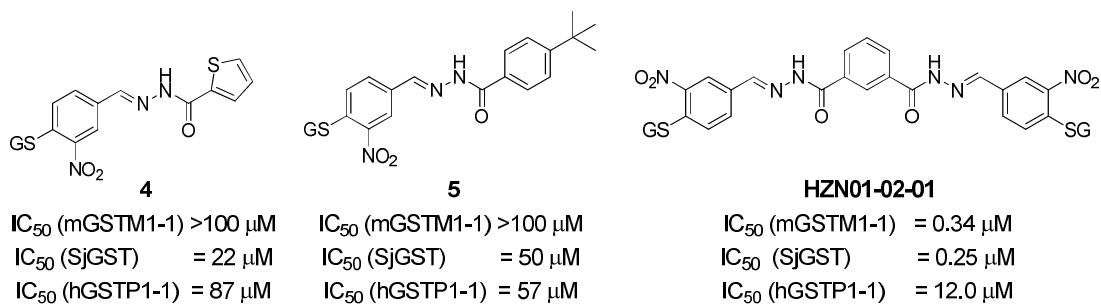


Figure 28: Comparison of binding data for competitive mono acylhydrazone GST inhibitors **4** and **5** and bis acylhydrazone **HZN01-02-01** for three GST isoforms.

ITC was used to determine whether, due to the short spacer length, two molecules were in fact binding to the GST dimer; one in each active site, with the rest of the molecule found in the solvent accessible cleft.

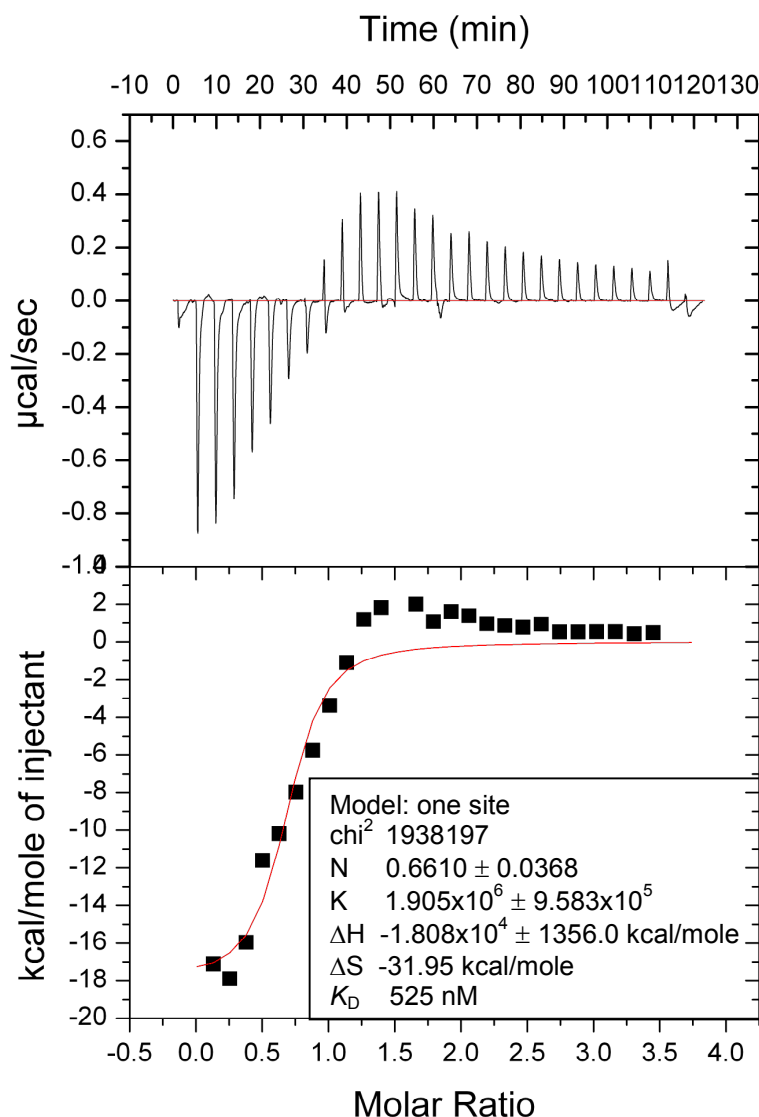


Figure 29: ITC data for **HZN01-02-01** with mGSTM1-1. Top panels show the heat effects associated with the injection of ligand into the cell containing GST. Bottom panels show the binding isotherm, generated from the integrated heats in the top panel, and the best fitted curves.

The stoichiometry of binding with respect to the monomer gave a value of 0.7, Figure 29. This did not give a clear indication of how the compound was binding to GST. The ITC data again showed that the mixing of the ligand into buffer was an endothermic process. A greater energy change was observed with this process than when **HZN01-03-01** was diluted. An interesting observation for **HZN01-02-01** and also **HZN01-01-01** at high concentrations in DMSO was that a gel formed. It seems most likely that this was due to hydrogen bonding interactions between the terminal

glutathione species, causing the formation of a polymeric species. This happened much more readily in the case of the flexible linear spacer **HZN01-01-01** and the short rigid spacer **HZN01-02-01**. In the same compounds without glutathione present at the terminal, gels no longer formed. This observation could explain some of the anomalous results associated with **HZN01-02-01**. This could explain why the binding data for **HZN01-02-01** did not always complement the amplifications observed in the protein-templated DCL. It also explains why it took a longer time period for the DCL to re-equilibrate when one or more of the acylhydrazone products were present from the beginning.

In the ITC data for **HZN01-02-01**, an extra endothermic event occurred once the protein was saturated with ligand and the event cannot be explained by ligand dilution into buffer. This can be seen at a molar ratio of approximately 1.5 in Figure 29 or at approximately 50 minutes in Figure 30. An explanation for this observation is yet to be established.

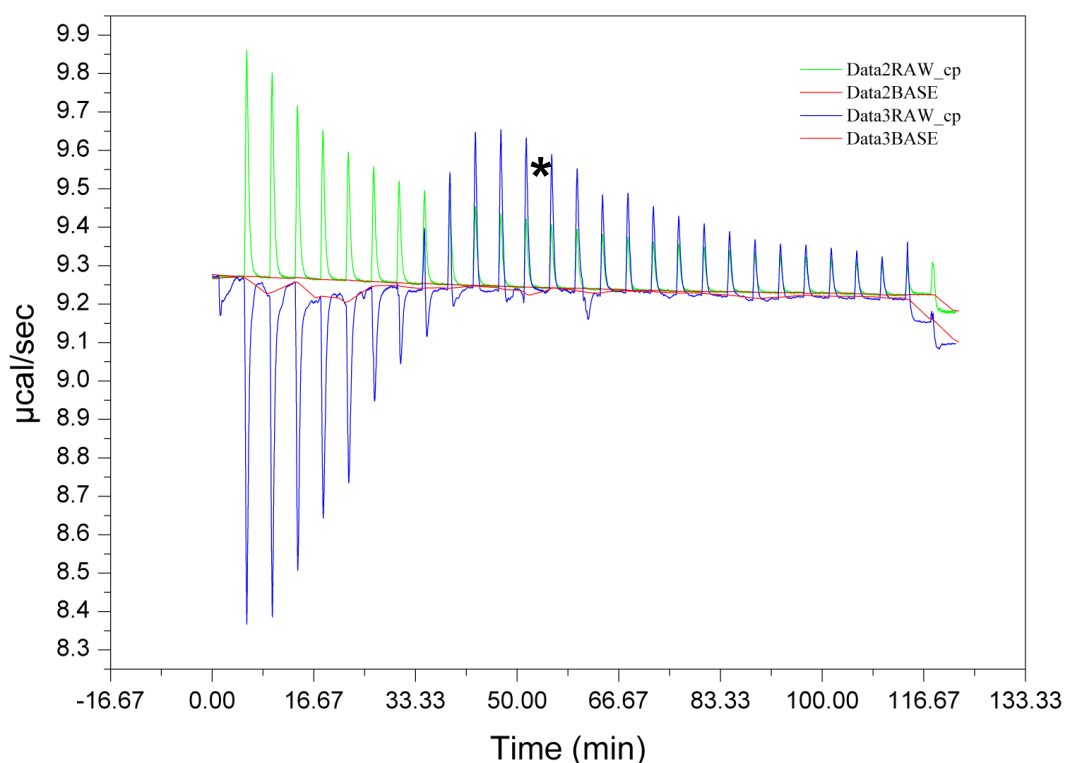
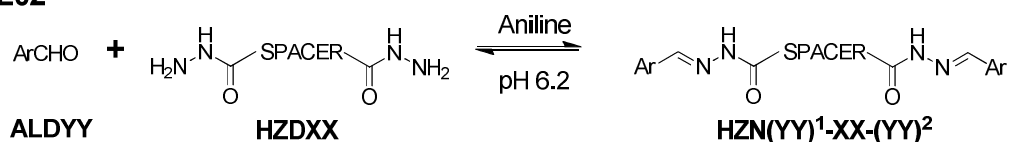


Figure 30: Overlay of the raw transpose from the ITC data for **HZN01-02-01** with mGSTM1-1. Green trace shows the heat effects associated with the injection of ligand into buffer. Blue line shows the heat effects associated with the injection of ligand into GST. Asterisk highlights the endothermic event associated with ligand added to mGSTM1-1.

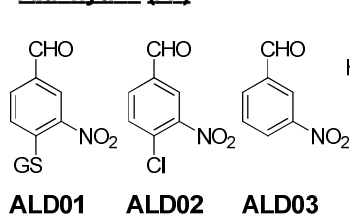
2.4 Increasing complexity within the dynamic combinatorial library

With the initial DCL set up and best inhibitor identified, the complexity of the library was increased. As the design of the library was based on bivalent spacers, with two reactive hydrazide groups per molecule, the introduction of multiple aldehydes to the group would greatly increase the complexity of the library with the possible formation of both homo-spacers and hetero-spacers. The addition of one more aldehyde to the initial library would increase the number of bis acylhydrazone products from four to twelve. The addition of a third aldehyde would increase the number of bis acylhydrazone products to twentyfour from only seven starting fragments. A second dynamic combinatorial library **DCL02** was set up using the same four hydrazide spacers with three aldehydes. The aldehyde analogue of CDNB **ALD02**, with a chlorine para to the aldehyde and the aldehyde **ALD03** with only the meta nitro group present and hydrogen in the para position were added to **ALD01**. These aldehydes were selected to study the binding effect of the glutathione species on the anchoring fragment. Again, the acylhydrazone products have been labelled in the format **HZN(YY)¹-XX-(YY)²**, where **XX** represents the code for the hydrazide spacer and **YY** represents the code for the aldehyde at each terminal position. For example the product resulting from the reaction of hydrazide **HZD02** with **ALD01** at one end and **ALD03** at the other end would be labelled **HZN01-02-03**.

DCL02



Aldehydes (Ar)



Spacers

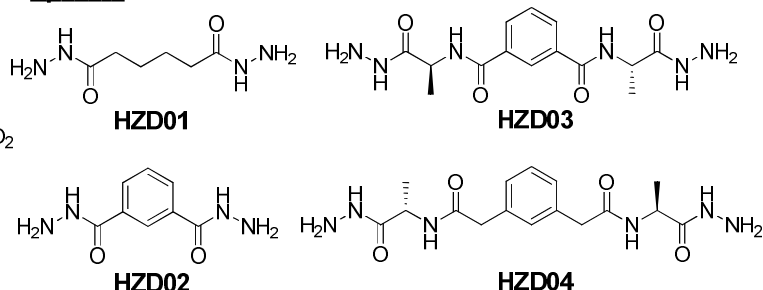


Figure 31: DCL02 with an increased aldehyde library.

The new library was set up under the same conditions as the previous **DCL01**. The introduction of the two more hydrophobic aldehydes **ALD02** and **ALD03** led to insoluble acylhydrazone products. The concentrations of the library members had to be reduced to overcome these solubility issues. Furthermore with the added complexity in the DCL, better chromatographic separation was required. The use of a second column on the HPLC led to much better separation.

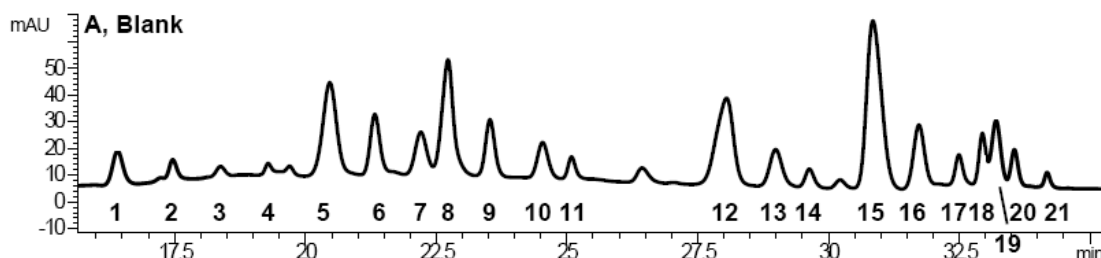


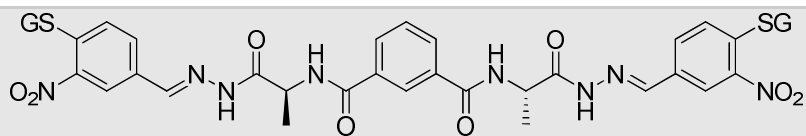
Figure 32: HPLC trace for **DCL02**. Peak numbers **1-21** are described in **Table 3**. Experimental conditions: **HZD01** to **HZD04** (final conc. = 80 μ M each, DMSO), **ALD01** (80 μ M, water), **ALD02** and **ALD03** (final conc. = 80 μ M each, DMSO) aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (255 μ L, 100 mM, pH 6.4) mixed at room temperature. After 48 h, 10 μ L of sample injected directly onto HPLC column. Products are detected by UV at $\lambda = 254$ nm. HPLC gradient A (see chapter 5, section 5.2.1).

The peaks in **DCL02**, Figure 32, were identified by a deconvolution of the library, using mixtures of one hydrazide with one, two or three aldehydes and one aldehyde with one, two, three or four hydrazides. The HPLC traces can be seen in the Appendix, Section 6.1.2, Figures 8-13. The peaks have been described in Table 3.

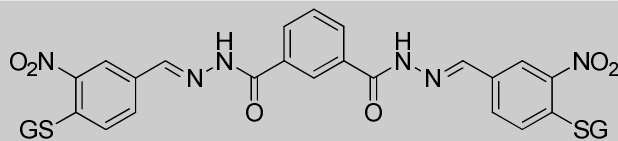
Table 3: A table showing the assignment of peaks observed in HPLC trace for **DCL02**.

Peak number	Acylhydrazone product
1	 HZN01-01-01
2	 HZN01-04-01

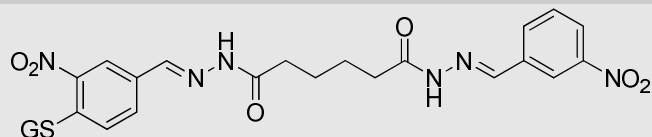
3

**HZN01-03-01**

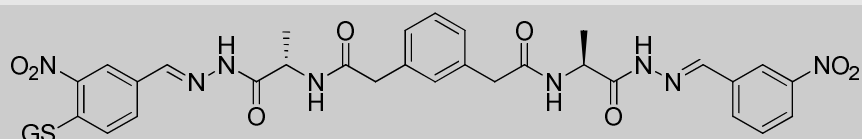
4

**HZN01-02-01**

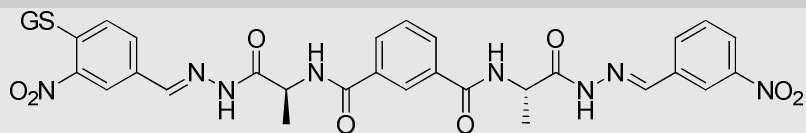
5

**HZN01-01-03**

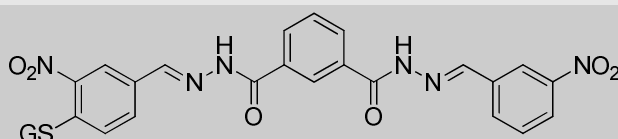
6

**HZN01-04-03**

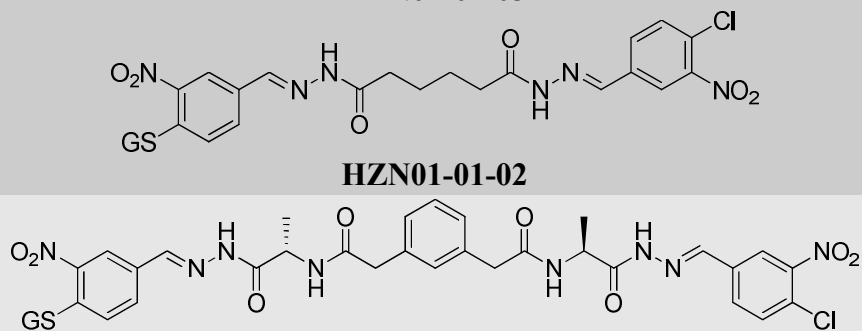
7

**HZN01-03-03**

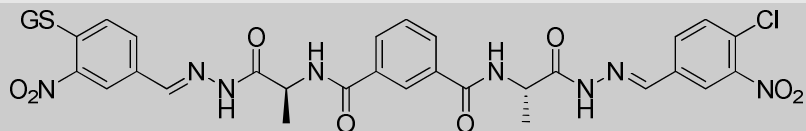
8

**HZN01-02-03**

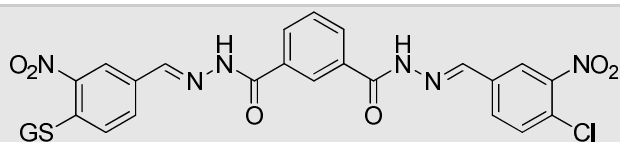
9

**HZN01-04-02**

10

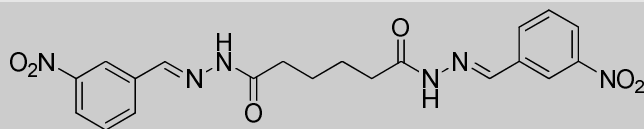
**HZN01-03-02**

11

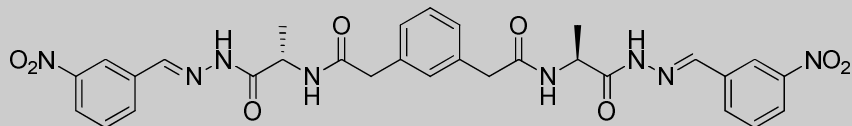


HZN01-02-02

12

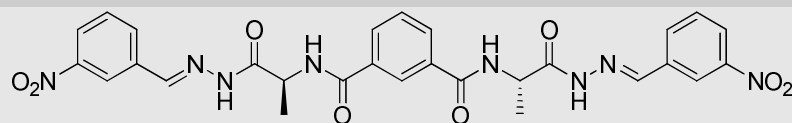


HZN03-01-03



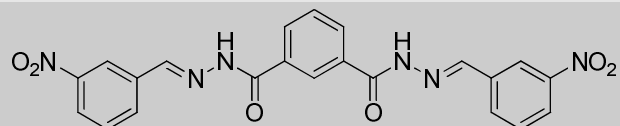
HZN03-04-03

13



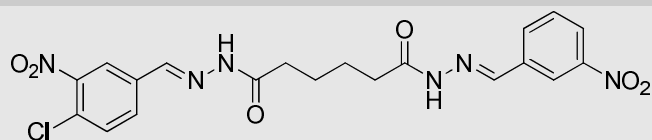
HZN03-03-03

14

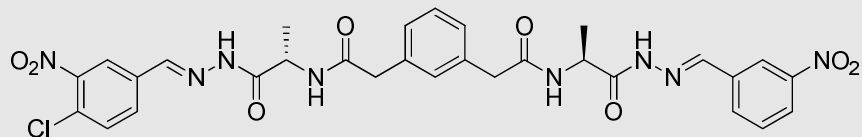


HZN03-02-03

15

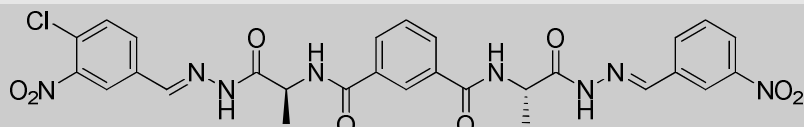


HZN02-01-03



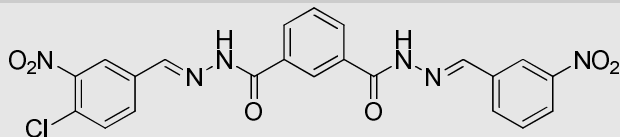
HZN02-04-03

16



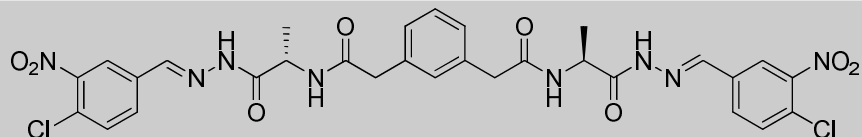
HZN02-03-03

17



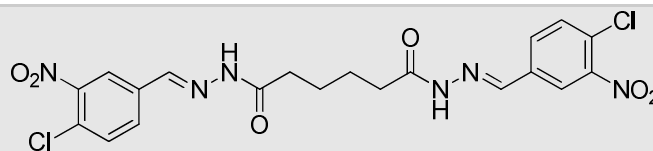
HZN02-02-03

18



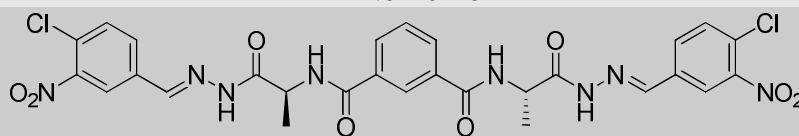
HZN02-04-02

19



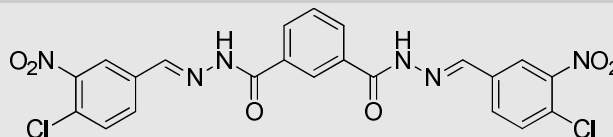
HZN02-01-02

20



HZN02-03-02

21



HZN02-02-02

2.4.1 Exploring isoform specificity with DCL02

The background library was set up along with mixtures containing one each of the four GST isoforms.

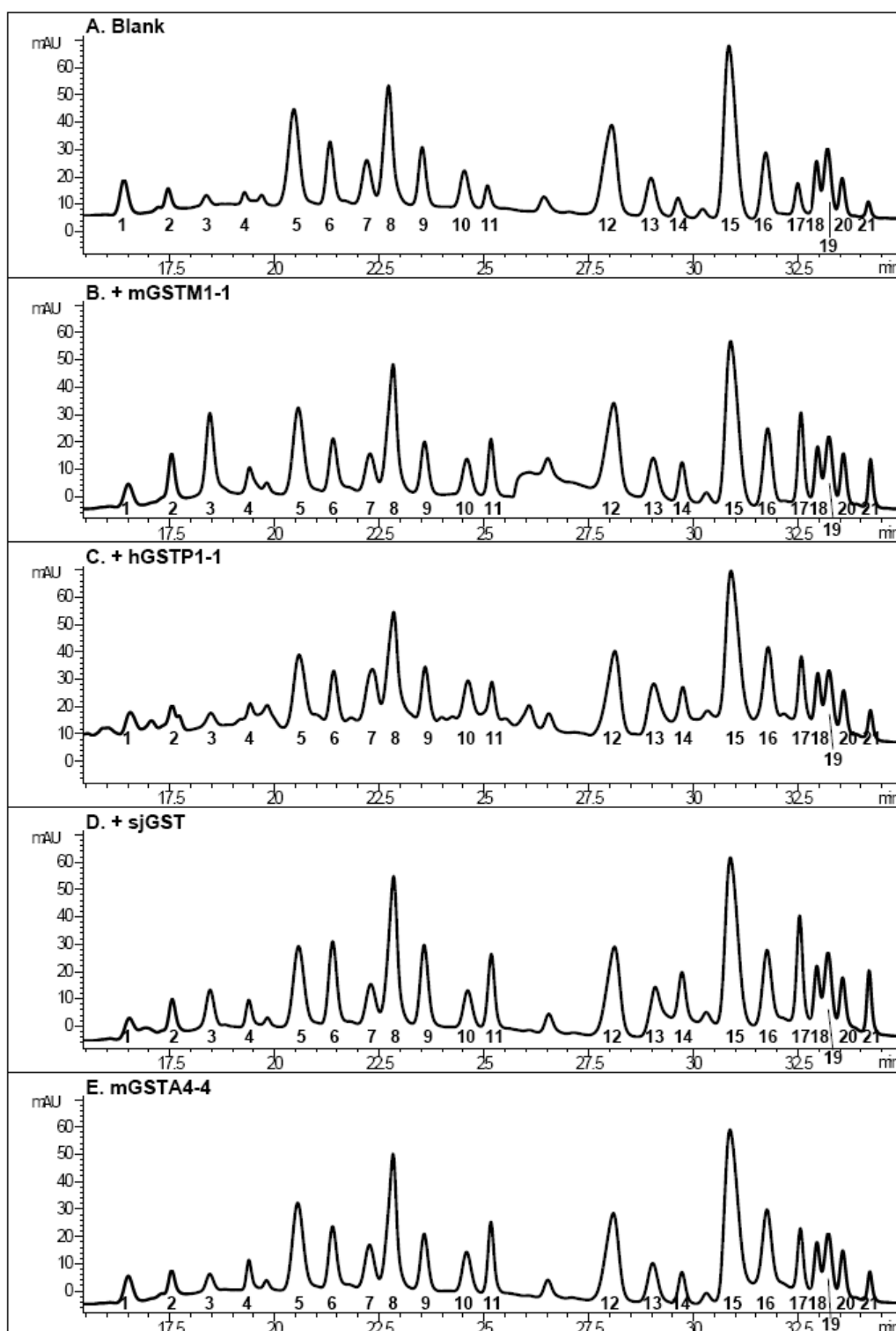


Figure 33: HPLC traces of **DCL02** reaction of **HZD01-HZD04** with **ALD01-AL03** in the presence of different GST isoforms. Experimental conditions: **HZD01** to **HZD04** (final conc. = 80 μ M each, DMSO), **ALD01** (80 μ M, water), **ALD02** and **ALD03** (final conc. = 80 μ M each, DMSO) aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (255 μ L mL, 100 mM, pH 6.4) mixed at room temperature. After 48 h, 10 μ L of

sample injected directly onto HPLC column. **A**) blank equilibrium, **B**) mGSTM1-1 (final conc. = 20 μ M), **C**) hGSTP1-1 (final conc. = 20 μ M), **D**) SjGST (final conc. = 20 μ M), **E**) mGSTA4-4 (final conc. = 20 μ M). Peak numbers **1-21** are described in **Table 3**. Products are detected by UV at $\lambda = 254$ nm. HPLC gradient A (see chapter 5, section 5.2.1).

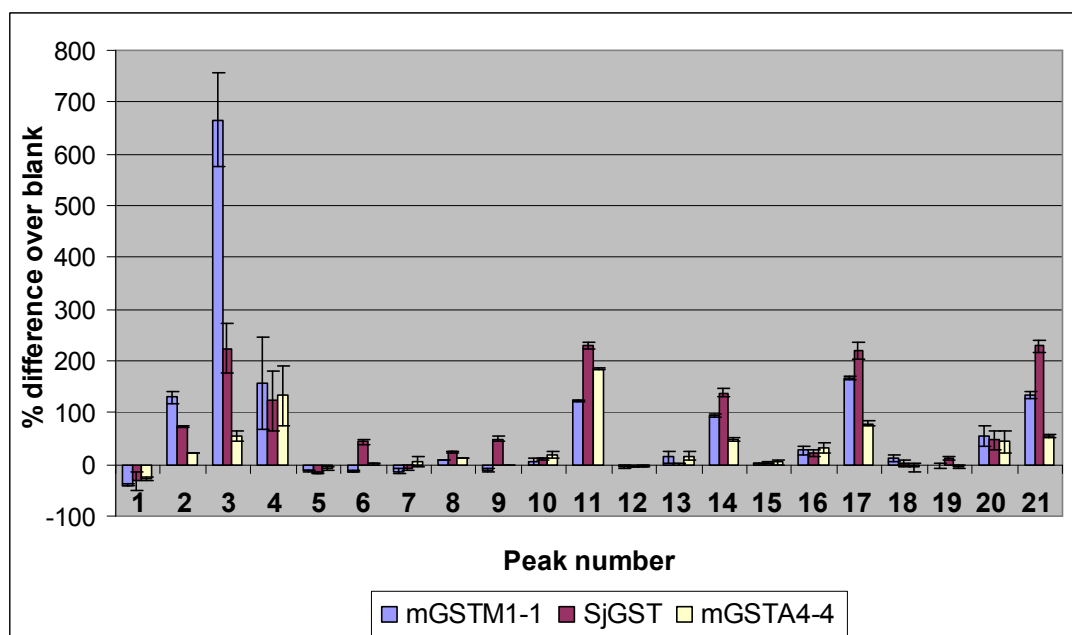


Figure 34: Graphical representation of **DCL02** in the presence of different GST isoforms. Bars represent % difference of peak area over blank equilibrium. Peak numbers **1-21** are described in **Table 3**. Error bars represent standard deviation over two experiments.

Again the peak showing the largest amplification was the acylhydrazone **HZN01-03-01**. Several smaller amplifications were also observed corresponding to acylhydrazones containing the hydrazide spacer **HZD02**, the shortest of the spacers. For the SjGST isoform several peaks were amplified to a similar extent over the blank equilibrium. With the exception of **HZN01-03-01**, all of the other amplifications were due to acylhydrazones containing the hydrazide spacer **HZD02**. As with the smaller DCL library **DCL01**, adding hGSTP1-1 resulted in the destruction of the products containing both terminal units from **ALD01**. The other components in the library appeared to remain unaffected by the presence of hGSTP1-1, with no change in the equilibrium observed. The only amplifications observed in the presence of mGSTA4-4 were for acylhydrazones containing the hydrazide spacer **HZD02**, the shortest of the hydrazide spacers.

2.4.2 Control experiments for DCL02

Control experiments were again set up to confirm that the amplifications that were observed were due to a templative effect from the GST active site. The addition of bovine serum albumin had little effect on the equilibrium of the DCL, Figure 35.

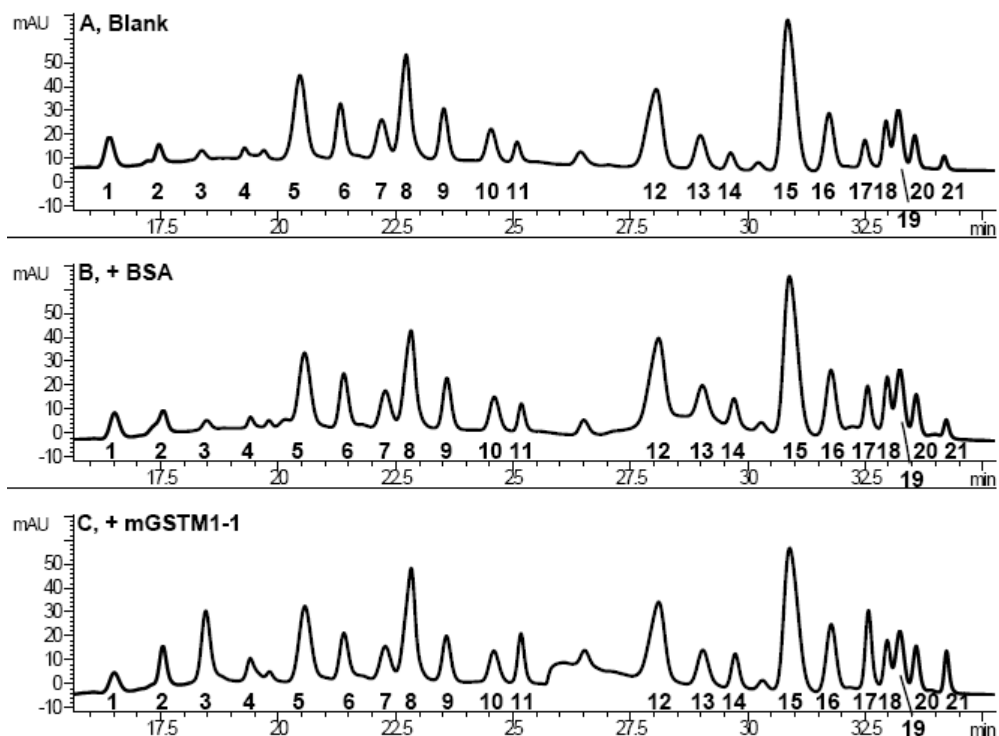


Figure 35: HPLC traces of **DCL02** in the presence of control protein BSA. Experimental conditions: **HZD01** to **HZD04** (final conc. = 80 μ M each, DMSO), **ALD01** (80 μ M, water), **ALD02** and **ALD03** (final conc. = 80 μ M each, DMSO) aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (255 μ L, 100 mM, pH 6.4) mixed at room temperature. After 48 h, 10 μ L of sample injected directly onto HPLC column. **A)** blank equilibrium, **B)** BSA (final conc. = 20 μ M), **C)** mGSTM1-1 (final conc. = 20 μ M). Products are detected by UV at λ = 254 nm. HPLC gradient A (see chapter 5, section 5.2.1).

The glutathione-conjugated CDNB substrate, DNP GSH was again used to block the binding pocket, to prevent amplification of the DCL.

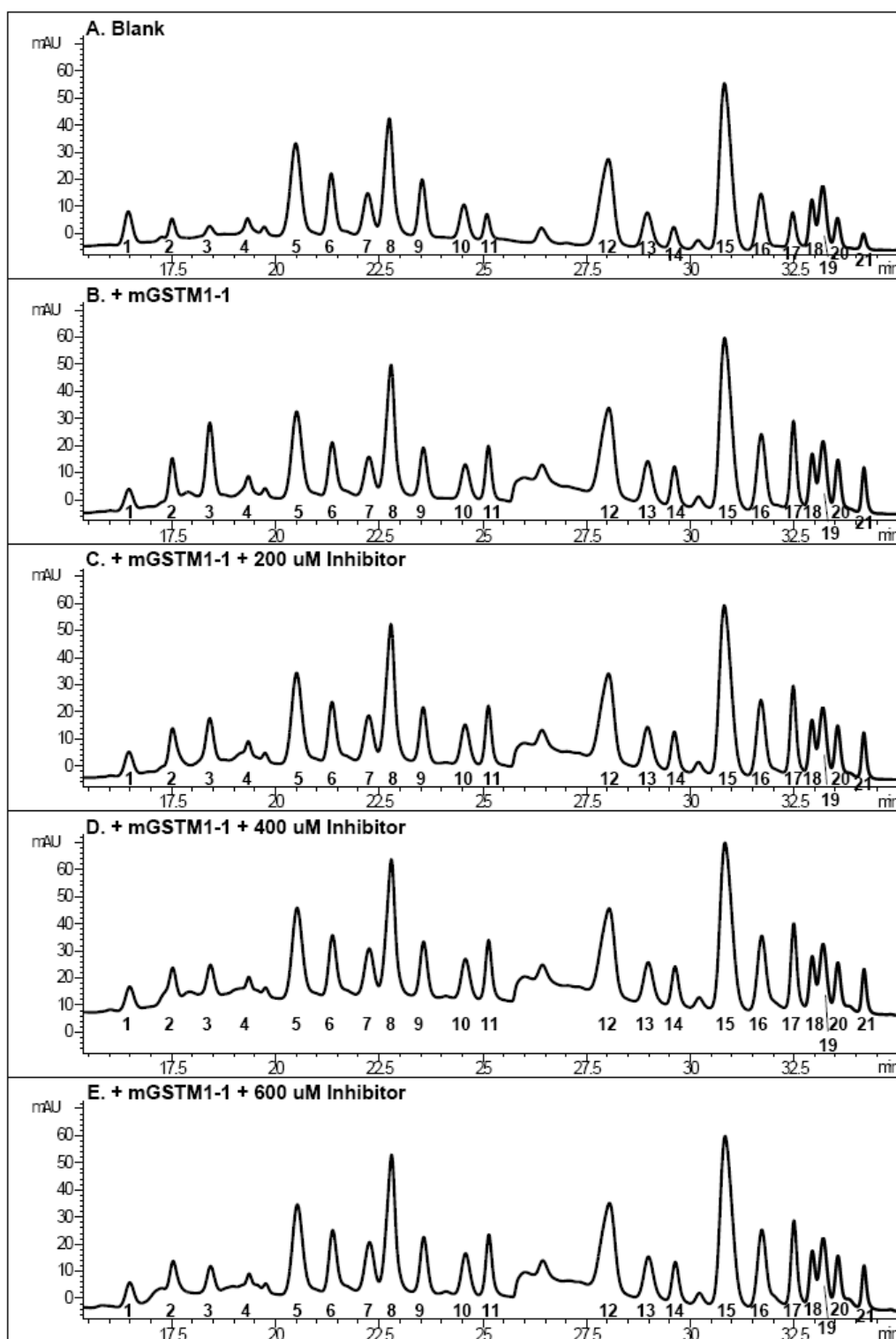


Figure 36: HPLC traces of **DCL02** in the presence of increasing concentrations of DNP GSH. Experimental conditions: **HZD01** to **HZD04** (final conc. = 80 μ M each, DMSO), **ALD01** (80 μ M, water), **ALD02** and **ALD03** (final conc. = 80 μ M each, DMSO) aniline (5 mM, DMSO), DMSO (15 % by vol.) and ammonium acetate buffer (255 μ L, 100 mM, pH 6.4) mixed at room temperature. After 48 h, 10 μ L of sample injected directly onto HPLC column.

A) blank equilibrium, **B)** mGSTM1-1 (final conc. = 20 μ M) + 0 μ M DNP GSH, **C)** mGSTM1-1 (final conc. = 20 μ M) + 200 μ M DNP GSH, **D)** mGSTM1-1 (final conc. = 20 μ M) + 400 μ M DNP GSH, **E)** mGSTM1-1 (final conc. = 20 μ M) + 600 μ M DNP GSH. Products are detected by UV at λ = 254 nm. HPLC gradient A (see chapter 5, section 5.2.1).

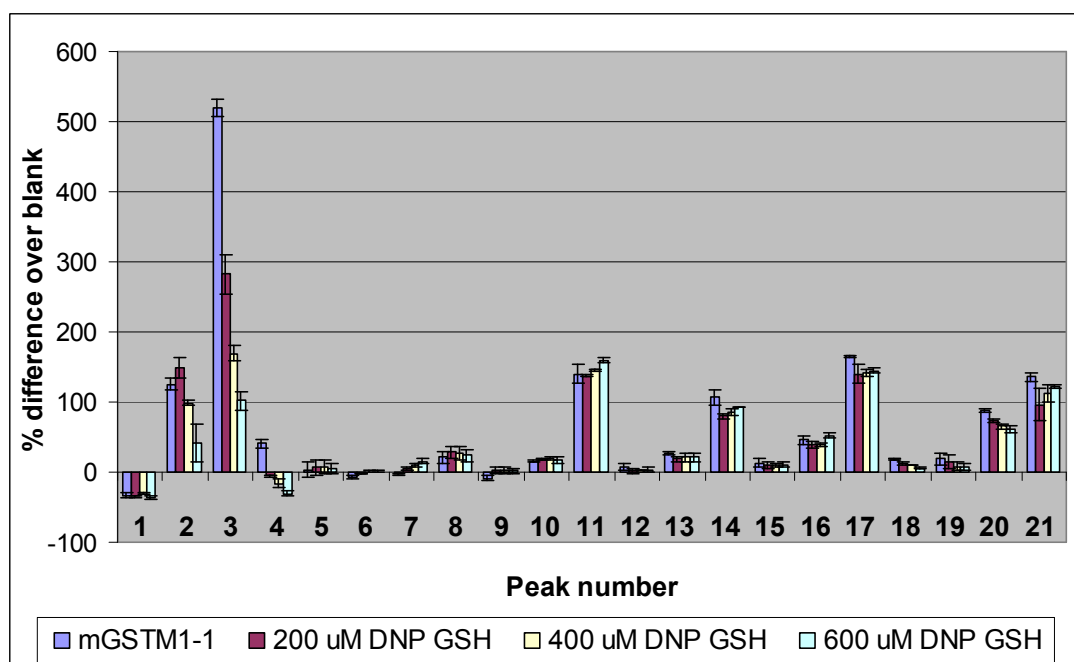


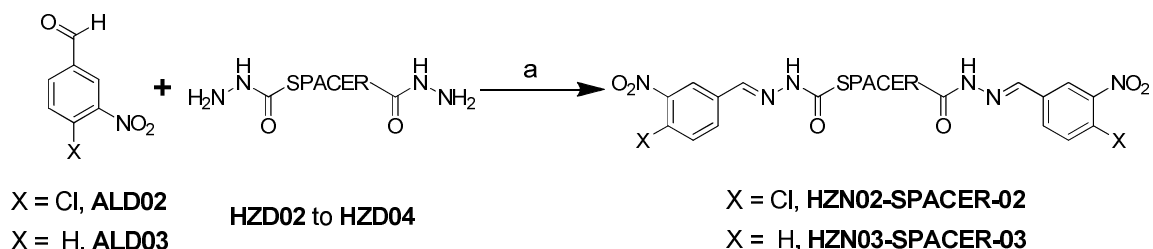
Figure 37: Graphical representation of **DCL02** in the presence of increasing concentrations of DNP GSH. Bars represent % difference of peak area over blank equilibrium. Peak numbers 1-21 are described in **Table 3**. Error bars represent standard deviation over two experiments.

The addition of DNP GSH to block the binding pocket produced a dose-dependent decrease in the amplification of **HZN01-03-01**, peak number 3 and **HZN01-04-01**, peak number 2. Surprisingly the amplification of peak numbers 11, 14, 17 and 21 corresponding to acylhydrazones not containing terminal glutathione remained unaffected by the addition of DNP GSH. It could indicate that the important interaction that was being blocked for the cases containing **ALD01** was the binding of the glutathione species. To see if those compounds not containing glutathione were binding in a different location within the enzyme, other inhibitors, such as ethacrynic acid and CDNB, were also added to the DCL. Unfortunately the inhibitors eluted in the middle of the library of bis acylhydrazones making it very difficult to compare the GST-templated library with the blank. As was seen in **DCL01**, these inhibitors did not appear to prevent the amplifications observed in the

protein-templated library for those products that were not masked in the HPLC trace, suggesting that these compounds might not be binding in the H-site at all.

2.4.3 Synthesis and binding data for homodimers from DCL02

The homodimers made from **ALD02** and **ALD03** were synthesised using the same chemistry as in the synthesis of **HZN01-01-01** to **HZN01-04-01**, Scheme 4.



Scheme 4: Synthesis of homo acylhydrazone products from **DCL02**. a; glacial acetic acid (few drops), water/ethanol 1:1, rt, 30 min.

The IC_{50} values for these compounds were determined using the CDNB assay.^[7] As shown in Table 4, the removal of glutathione from the products resulted in a 100-fold decrease in inhibition for the longer spacers **HZD03** and **HZD04**, demonstrating the importance of glutathione for binding affinity. The loss of glutathione for the shortest spacer **HZD02** was much better tolerated, with only a two-fold reduction in inhibition. This also suggests that these compounds were binding in a different way to the longer spacers.

Table 4: IC_{50} data for homo acylhydrazone spacers with mGSTM1-1. Data acquired using the CDNB assay. Values are calculated from three separate assays. Raw data described in experimental chapter.

Hydrazone spacer	HZN01-XX-01 IC_{50} mGSTM1 (μM)	HZN02-XX-02 IC_{50} mGSTM1 (μM)	HZN03-XX-03 IC_{50} mGSTM1 (μM)
Aldehyde (no spacer)	341.7	-	> 500
HZD01	1.21	-	-
HZD02	0.337	2.02	3.38
HZD03	0.050	9.32	106.4
HZD04	0.413	49.89	> 100

HZN02-02-02, the most potent bivalent inhibitor without glutathione, was run in the CDNB competition assay. This compound was also a noncompetitive partial inhibitor with respect to CDNB. The K_i value was $1.98 \mu\text{M} \pm 0.47$. This again fit with the Cheng-Prusoff equation $\text{IC}_{50} = K_i$ in the case of noncompetitive inhibitors.^[42]

2.5 Conclusion and future work

The possibilities of DCC have clearly been demonstrated with the identification of a novel isoform specific bivalent inhibitor of mGSTM1-1. The power of DCC has been shown, not only in its primary goal of identifying novel inhibitors from weakly binding fragments, but also with the insight offered into the mode of binding by the careful use of “inhibitors” to block the amplification of library members. It has also been shown that DCC is capable of coping with the increased complexity of bivalent inhibitors.

Bis acylhydrazone **HZN01-03-01** has been identified as a novel inhibitor for mGSTM1-1, with an IC_{50} value of 50 nM. The design of GST inhibitors described in the literature has primarily focussed on ligands for the GSTP1-1 isoform because of the associations between GSTP1-1 and cancer. The dye cibacron blue is the most potent inhibitor of GSTM1-1 known to date and has an IC_{50} of 50 nM for the human isoform and 700 nM for the mouse isoform, which is used in this study.^[2] Bis acylhydrazone **HZN01-03-01** exhibits a 14-fold increase in inhibitory activity over the best literature example, cibacron blue. **HZN01-03-01** also shows excellent isoform specificity for the Mu isoform of GST over several other isoforms studied, a 20-fold increase in inhibitory activity is observed for mGSTM1-1 over SjGST. The best example of isoform specificity towards GSTM1-1 described in the literature is from Kunze and Heps.^[43] They report an analogue of glutathione with a phosphono group replacing the cysteine side chain, which has an IC_{50} value of $4.7 \mu\text{M}$ for hGSTM1-1 and only a 3-fold increase in potency over hGSTP1-1 and hGSTA1-1.

The isoform selectivity demonstrated by the bis acylhydrazone **HZN01-03-01** may be explained by looking at the 3D structural differences between the GST isoforms, which have been described in detail in section 2.1. The Mu isoform

contains a mu loop, which creates a deeper solvent-accessible cleft between the binding pockets of each monomer unit.^[8] This deeper cavity could present more space for the acylhydrazone spacer to bind. The Alpha class contains an additional alpha helix, which fits partially into the solvent-accessible cleft and makes the hydrophobic pocket smaller.^[9] This reduction in space could explain why the acylhydrazone compounds had no inhibitory activity with mGSTA4-4. The Pi class of GST does not contain the Mu loop or the extra alpha helix so the space available lies somewhere between the Mu isoform and the Alpha isoform. The binding affinity of the acylhydrazones for hGSTP1-1 was also between the Mu isoform and the Alpha isoform (13.45 μ M for hGSTP1-1 compared with 50 nM for mGSTM1-1 and > 100 μ M for mGSTA4-4 for **HZN01-03-01**), providing evidence that the available space within the H-site and into the solvent-accessible cleft is playing an important role in recognition for the acylhydrazones. The SjGST isoform most closely resembles the Mu isoform and contains a shortened version of the Mu loop.^[10-12] The acylhydrazones also show reasonable, although reduced, inhibition for SjGST compared with inhibition for mGSTM1-1. Several other factors may also play an important role in the spacer selection, including the distance between binding pockets and the polarity of the residues lining the solvent-accessible cleft. The isoform specificity gained by using the bivalent spacers could be further probed by increasing the size of the hydrazide library. The bivalent compounds described by the Atkins group suggested that increasing spacer length resulted in an increase in affinity for hGSTP1-1, so adding longer hydrazide spacers to the library may result in the identification of hGSTP1-1 isoform specific ligands.^[32, 34]

The careful selection of inhibitor compounds used to block the GST binding pockets during the DCC reaction led to the conclusion that the bivalent acylhydrazones were binding in the “out” mode within the hydrophobic binding pocket. This was further confirmed by the observation that the bivalent compounds were noncompetitive inhibitors with respect to CDNB. The presence of glutathione at the terminal of each spacer was of particular importance for binding affinity for the longer spacers **HZN01-03-01** and **HZN01-04-01**. The glutathione was less important for products made from **HZD02**. All products made from **HZD02** with any of the three aldehydes at the end gave amplifications in the GST-templated reaction.

These results may indicate that these smaller compounds are binding in a different location in the GST protein, perhaps in the solvent-accessible cleft as observed with the binding of praziquantel to SjGST.^[11] A possible way to determine this would be to repeat the inhibitor-blocked DCC reactions using a larger library of ligands, such as bromosulfophthalein or praziquantel that bind in the ligandin site or the solvent-accessible cleft, and observing the effects on the library amplification. A crystal structure of the bivalent ligand bound to GST would be most desirable to confirm the binding mode of the acylhydrazones and particularly to confirm if those products made from **HZD02** are binding in a different way. Molecular modelling is currently being employed, in collaboration with Iain McNae of the Walkinshaw group, to determine the possible binding modes of the bivalent compounds within the GST dimer.

2.6 References

- [1] J. Booth, E. Boyland, P. Sims, *Biochem. J.* **1961**, *79*, 516.
- [2] B. Mannervik, U. Helena Danielson, *Crit. Rev. Biochem. Mol. Biol.* **1988**, *23*, 283.
- [3] T. H. Rushmore, C. B. Pickett, *J. Biol. Chem.* **1993**, *268*, 11475.
- [4] R. N. Armstrong, *Chem. Res. Toxicol.* **1997**, *10*, 2.
- [5] D. Sheehan, G. Meade, V. M. Foley, C. A. Dowd, *Biochem. J.* **2001**, *360*, 1.
- [6] J. D. Hayes, J. U. Flanagan, I. R. Jowsey, *Annual Review of Pharmacology & Toxicology* **2005**, *45*, 51.
- [7] W. H. Habig, M. J. Pabst, W. B. Jakoby, *J. Biol. Chem.* **1974**, *249*, 7130.
- [8] X. Ji, P. Zhang, R. N. Armstrong, G. L. Gilliland, *Biochemistry* **1992**, *31*, 10169.
- [9] I. Sinning, G. J. Kleywegt, S. W. Cowan, P. Reinemer, H. W. Dirr, R. Huber, G. L. Gilliland, R. N. Armstrong, X. Ji, P. G. Board, B. Olin, B. Mannervik, T. A. Jones, *J. Mol. Biol.* **1993**, *232*, 192.
- [10] K. Lim, J. X. Ho, K. Keeling, G. L. Gilliland, X. Ji, F. Rüker, D. C. Carter, *Protein Science* **1994**, *3*, 2233.
- [11] M. A. McTigue, D. R. Williams, J. A. Tainer, *J. Mol. Biol.* **1995**, *246*, 21.

- [12] R. M. F. Cardoso, D. S. Daniels, C. M. Bruns, J. A. Tainer, *Proteins: Structure, Function, and Bioinformatics* **2003**, *51*, 137.
- [13] A. C. Rufer, L. Thiebach, K. Baer, H. W. Klein, M. Hennig, *Acta Crystallographica Section F* **2005**, *61*, 263.
- [14] J. Rossjohn, W. J. McKinstry, A. J. Oakley, M. W. Parker, G. Stenberg, B. Mannervik, B. Dragani, R. Cocco, A. Aceto, *J. Mol. Biol.* **2000**, *302*, 295.
- [15] A. M. Caniard, PhD thesis, University of Edinburgh (Edinburgh), **2010**.
- [16] U. Krengel, K.-H. Schröter, H. Hoier, A. Arkema, K. H. Kalk, P. Zimniak, B. W. Dijkstra, *FEBS Lett.* **1998**, *422*, 285.
- [17] Y. Patskovsky, L. Patskovska, S. C. Almo, I. Listowsky, *Biochemistry* **2006**, *45*, 3852.
- [18] R. N. Armstrong, *Chem. Res. Toxicol.* **1991**, *4*, 131.
- [19] A. J. Oakley, M. Lo Bello, M. Nuccetelli, A. P. Mazzetti, M. W. Parker, *J. Mol. Biol.* **1999**, *291*, 913.
- [20] D. M. Townsend, K. D. Tew, *Oncogene* **2003**, *22*, 7369.
- [21] C. C. McIlwain, D. M. Townsend, K. D. Tew, *Oncogene* **2006**, *25*, 1639.
- [22] P. Ruzza, A. Rosato, C. R. Rossi, M. Floreani, L. Quintieri, *Anti-Cancer Agents Med. Chem.* **2009**, *9*, 763.
- [23] K. D. Tew, *Cancer Res.* **1994**, *54*, 4313.
- [24] E. Laborde, *Cell Death Differ* **2010**, *17*, 1373.
- [25] S.-G. Cho, Y. H. Lee, H.-S. Park, K. Ryoo, K. W. Kang, J. Park, S.-J. Eom, M. J. Kim, T.-S. Chang, S.-Y. Choi, J. Shim, Y. Kim, M.-S. Dong, M.-J. Lee, S. G. Kim, H. Ichijo, E.-J. Choi, *J. Biol. Chem.* **2001**, *276*, 12749.
- [26] S. Dorion, H. Lambert, J. Landry, *J. Biol. Chem.* **2002**, *277*, 30792.
- [27] K. Ryoo, S.-H. Huh, Y. H. Lee, K. W. Yoon, S.-G. Cho, E.-J. Choi, *J. Biol. Chem.* **2004**, *279*, 43589.
- [28] S. Mahajan, W. M. Atkins, *Cell. Mol. Life Sci.* **2005**, *62*, 1221.
- [29] M. H. Lyttle, M. D. Hocker, H. C. Hui, C. G. Caldwell, D. T. Aaron, A. Engqvist-Goldstein, J. E. Flatgaard, K. E. Bauer, *J. Med. Chem.* **1994**, *37*, 189.
- [30] B. Shi, R. Stevenson, D. J. Campopiano, M. F. Greaney, *J. Am. Chem. Soc.* **2006**, *128*, 8459.

- [31] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat Chem* **2010**, *2*, 490.
- [32] R. P. Lyon, J. J. Hill, W. M. Atkins, *Biochemistry* **2003**, *42*, 10418.
- [33] M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2754.
- [34] D. Y. Maeda, S. S. Mahajan, W. M. Atkins, J. A. Zebala, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3780.
- [35] S. S. Mahajan, L. Hou, C. Doneanu, R. Paranj, D. Maeda, J. Zebala, W. M. Atkins, *J. Am. Chem. Soc.* **2006**, *128*, 8615.
- [36] L. A. Ralat, R. F. Colman, *J. Biol. Chem.* **2004**, *279*, 50204.
- [37] A. J. Oakley, J. Rossjohn, M. Lo Bello, A. M. Caccuri, G. Federici, M. W. Parker, *Biochemistry* **1997**, *36*, 576.
- [38] A. J. Oakley, M. Lo Bello, A. P. Mazzetti, G. Federici, M. W. Parker, *FEBS Lett.* **1997**, *419*, 32.
- [39] X. Ji, R. N. Armstrong, G. L. Gilliland, *Biochemistry* **1993**, *32*, 12949.
- [40] L. Prade, R. Huber, T. H. Manoharan, W. E. Fahl, W. Reuter, *Structure* **1997**, *5*, 1287.
- [41] D. Kolobe, Y. Sayed, H. W. Dirr, *Biochem. J.* **2004**, *382*, 703.
- [42] Y.-C. Cheng, W. H. Prusoff, *Biochem. Pharmacol.* **1973**, *22*, 3099.
- [43] T. Kunze, S. Heps, *Biochem. Pharmacol.* **2000**, *59*, 973.

Chapter 3: An *in situ* click chemistry approach to identify novel bivalent inhibitors of glutathione transferases

3.1 Introduction

A second approach to target-guided synthesis (TGS) is to use reactions that function under kinetic control. This approach makes use of irreversible reactions that will be selectively promoted by the introduction of a biological template, which can bring reactive species in close proximity allowing the reaction to occur (Figure 1).^[1-3]

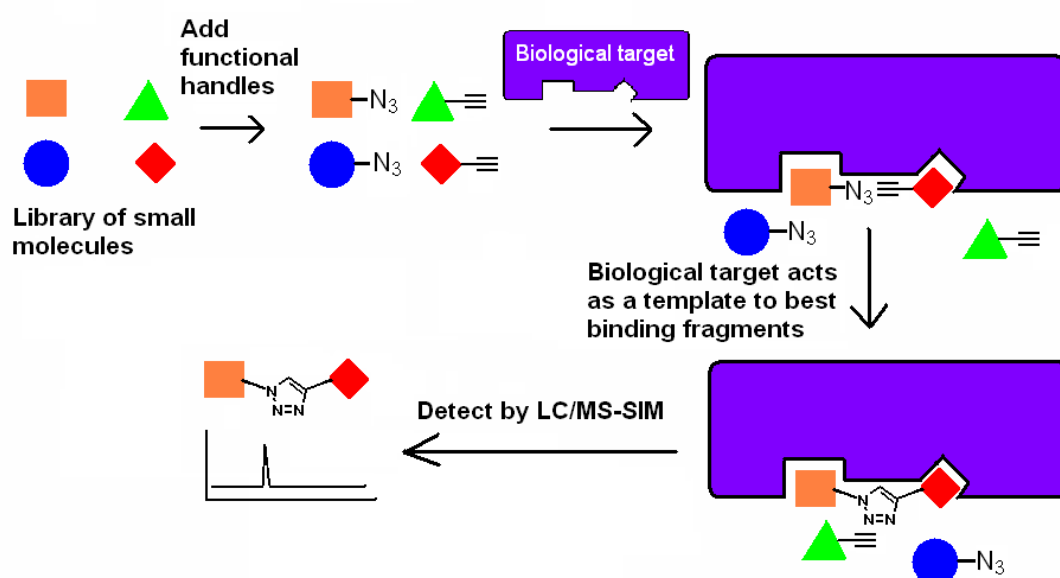


Figure 1: A cartoon diagram outlining the kinetic approach to target-guided synthesis using the Huisgen 1, 3 dipolar cycloaddition reaction.

A number of reactions used in this approach have been described in the introduction of this thesis. Of the reactions discussed, by far the most successful is the Huisgen 1, 3 dipolar cycloaddition reaction or as it is more commonly known the “click” reaction. A click reaction is a reaction that is reliable, selective and efficient. Sharpless *et al.* have outlined a series of rules to define a click reaction as a reaction that “*must be modular, wide in scope, give very high yields, generate only inoffensive*

byproducts that can be removed by nonchromatographic methods, and be stereospecific.”^[4] The 1, 3 dipolar cycloaddition reaction between an azide and a terminal alkyne is atom efficient, producing no by-products, and bio-orthogonal, preventing side reactions with the biological template. The copper (I) catalysed reaction to yield the *anti* (1,4-substituted) triazole was introduced by the Meldal group in the solid phase synthesis of peptidotriazoles and further developed by Sharpless and Fokin to be utilized with a broad spectrum of reactants.^[5, 6] While the copper (I) catalysed reaction is commonly used in a wide range of chemistries,^[7-10] the uncatalysed reaction is typically extremely slow. As described in the introduction, Sharpless and co-workers have shown that in the presence of a biological template the click reaction can be accelerated if the reactants have a good fit within the template.^[11] This acceleration in rate allows the products to be detected over any background reaction.

3.2 Designing the *in situ* click chemistry library – a triazole as a peptide mimic

In order to compare the thermodynamic approach to TGS with the kinetic approach, an *in situ* click chemistry (ISCC) library was designed to develop novel inhibitors of glutathione transferase (GST). Inspiration was again taken from the bivalent inhibitors described by Atkins *et al.*, which included inhibitors designed with a peptidic spacer **41**, Figure 2.^[12-14]

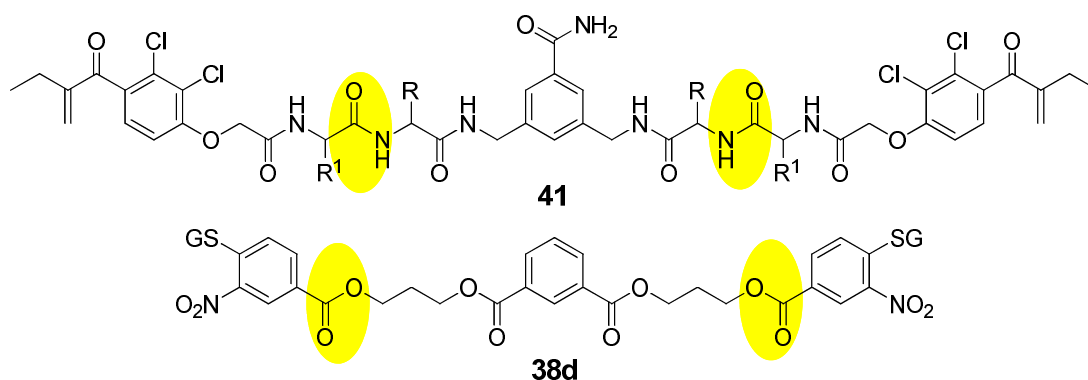


Figure 2: Bivalent inhibitors of GST designed by Atkins and co-workers.

There is precedent in the literature for replacing an amide bond with a disubstituted triazole ring.^[15, 16] The 1,4-substituted triazole mimics the *Z*-amide

bond, with similar atom placement and electronic properties, as detailed in Figure 3. The lone pair found on the 3-nitrogen mimics the carbonyl oxygen of the amide bond, the polarised C(5)-H bond is able to act as a hydrogen bond donor in the same way as the amide N-H bond, and finally the carbonyl carbon is mimicked by the electrophilic and polarised 4-carbon in the triazole. Structurally, the triazole unit contains an extra atom between the substituents giving an overall increased distance of approximately 1.1 Å.

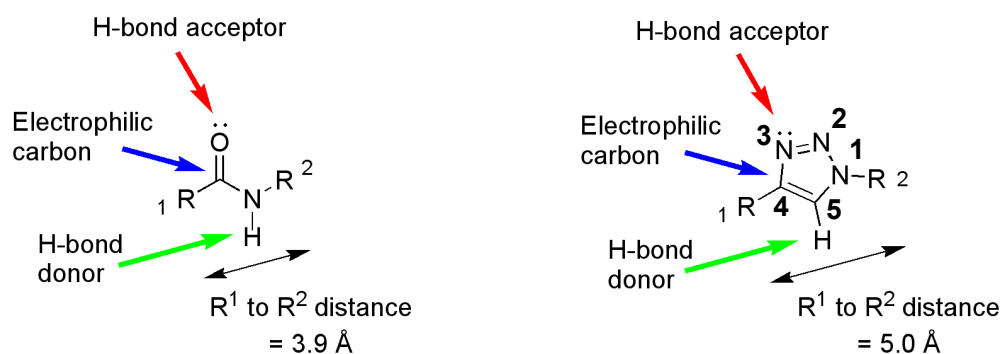


Figure 3: A diagram to show the similarities between a peptide bond and a 1,4 -disubstituted triazole.

Brik and co-workers have developed novel HIV-1 protease inhibitors containing a triazole species.^[15] Crystal structures of these inhibitors showed that the triazole was interacting with the protease, *via* hydrogen bonding, in the same way as the amide bond found in known inhibitor Amprenavir, as shown in Figure 4.

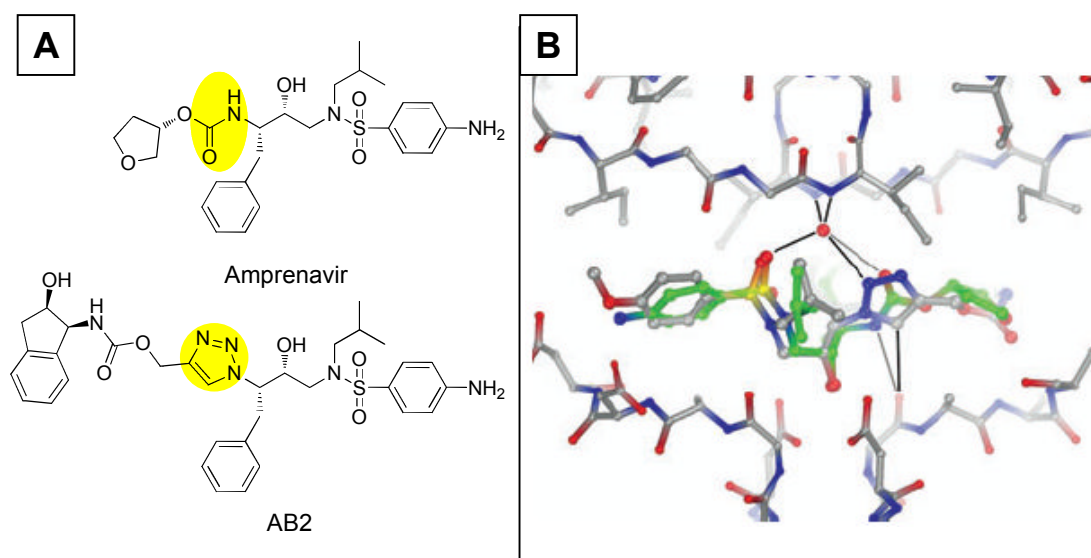
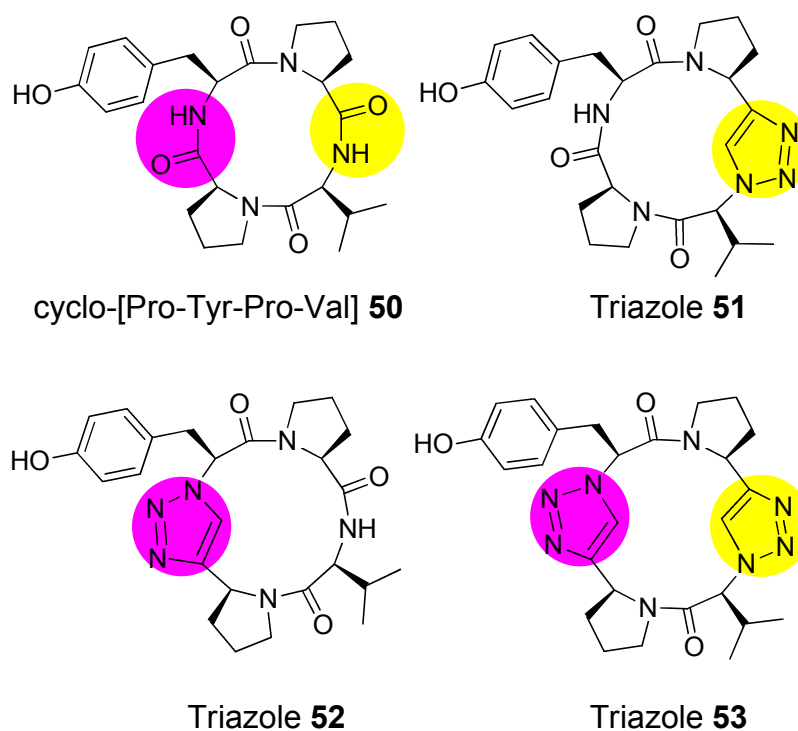


Figure 4: Peptide mimics of HIV-1 protease inhibitors. **A)** HIV-1 protease inhibitor Amprenavir and the triazole analogue AB2, **B)** Crystal structure of AB2 bound into the active site of HIV-1 protease. The position of Amprenavir is shown with the carbon atoms in green. Key hydrogen bonds to a structural water molecule and the main chain of the protease are shown with black lines. Crystal structure taken from paper by Brik *et al.*^[15]

Van Maarseveen and co-workers successfully replaced amide bonds in the cyclic peptide tyrosinase inhibitor cyclo-[Pro-Val-Pro-Tyr] **50** with a 1,4-substituted triazole, as demonstrated in Figure 5.^[16] These compounds, with a triazole as a peptide bond mimic, not only match the inhibitory activity of the original cyclic peptide, but in two examples, **51** and **52**, actually show a three-fold increase in activity with IC_{50} values of 0.5 mM and 0.6 mM compared with an IC_{50} value of 1.5 mM for the naturally occurring peptide **50**.^[17]



Compound	Tyrosinase activity IC ₅₀ (mM)
cyclo-[Pro-Tyr-Pro-Val] 50	1.5
Triazole 51	0.6
Triazole 52	0.5
Triazole 53	1.6

Figure 5: The naturally occurring cyclotetrapeptide cyclo-[Pro-Tyr-Pro-Val] a potent tyrosinase inhibitor and three triazole analogues with one or two amide bonds replaced by 1,4 substituted triazoles. Tyrosinase activity for the cyclopeptide and three triazole analogues is included.^[16, 17]

Bivalent compound **41**, designed by Atkins *et al.*, contains two amide bonds on either side of a central ring that could be potentially replaced by a triazole.^[14] The amide bonds, highlighted in yellow, in Figure 6 were selected as target-bonds to be modified. This would provide a connection between the central spacer unit, which bridges the space between the active sites of each monomer unit of GST, and the terminal unit that would anchor in the hydrophobic binding pocket (H-site) of GST.

Disconnection of the triazole to the starting azide and alkyne led to a bisazide spacer and an alkynyl-functionalised hydrophobic compound.

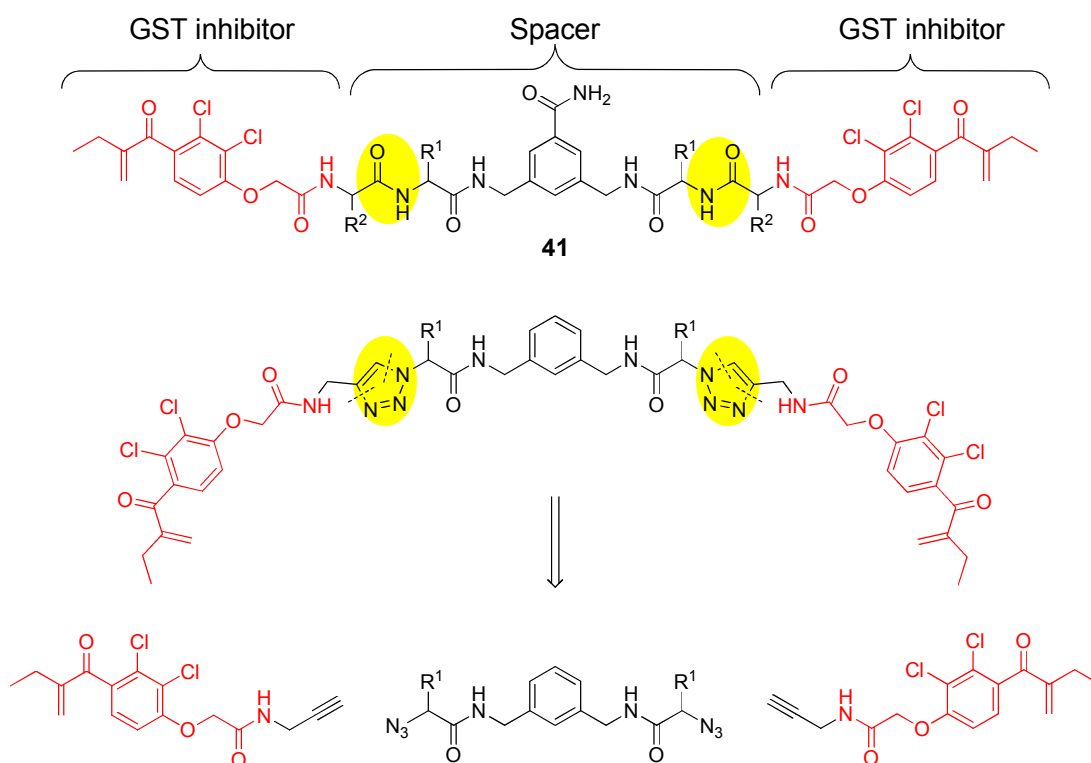
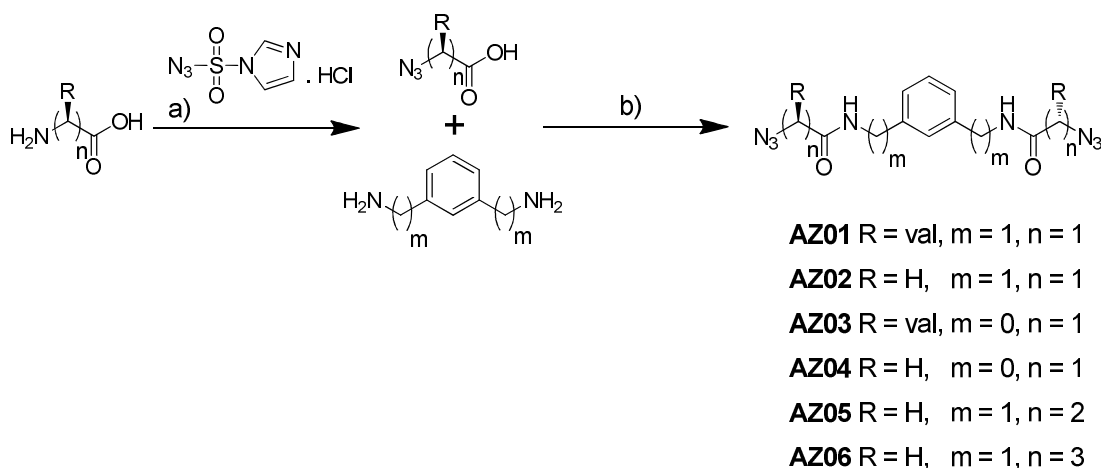


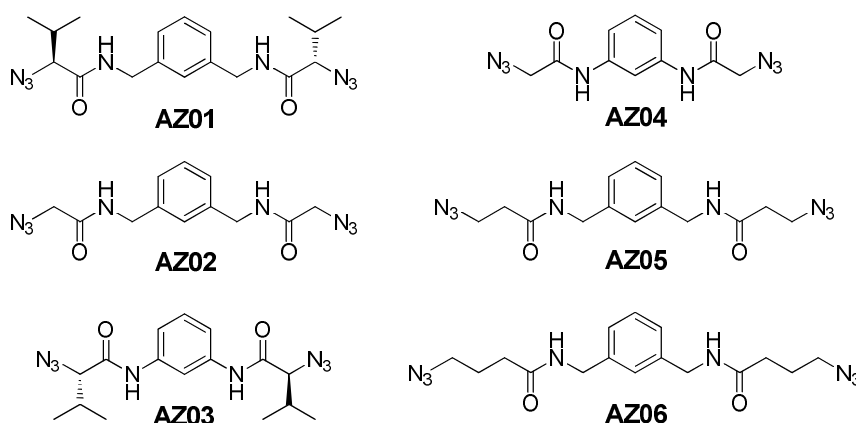
Figure 6: Design of ISCC fragments based on known bivalent inhibitors.^[14]

3.2.1 Synthesis of the azide spacer library

A small library of six azide spacers was synthesised, including a range of four different lengths (4-8 carbons) and two different side chains (glycine or valine). The valine side chain was selected as this was present in one of the most potent inhibitors for hGSTP1-1 identified by Atkins *et al.* with a K_d of $204 \text{ nM} \pm 21$.^[14] The synthesis of the spacers was carried out in two uncomplicated steps. In the initial step azide formation was realised by the reaction of imidazole sulfonyl azide in a diazotransfer reaction, as described by Goddard-Berger.^[18] This was followed by a straightforward EDC-coupling to a central amine unit, Scheme 1.



Azide spacer library

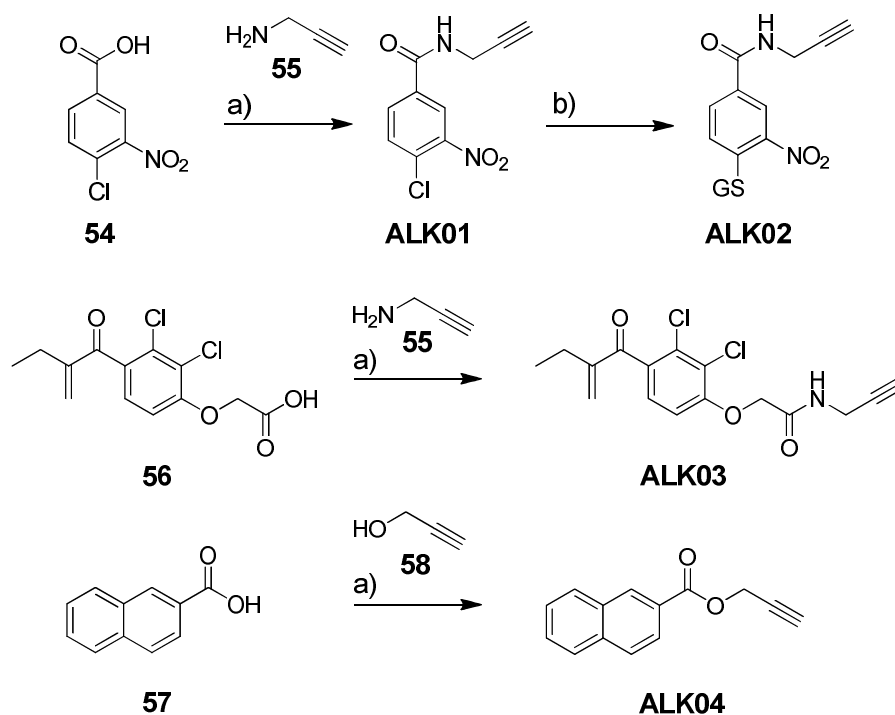


Scheme 1: Synthesis of azide spacers. a) 0.01 eq $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 2.2 eq K_2CO_3 , methanol, rt, 18 h. b) 1.2 eq EDC.HCl, 0.2 eq DMAP, dichloromethane, rt, 18 h.

3.2.2 Design and synthesis of a small alkyne library

For the preliminary research into the ISCC conditions required for this system, a small library of four alkyne fragments was designed based on known inhibitors or substrates of GST. The first two alkynes were based on 1-chloro-2,4-dinitrobenzene, CDNB the universal substrate for GST.^[19, 20] The nitro group *para* to the chloride was replaced with carboxylic acid **54** and this group coupled with propargylamine **55** using EDC.HCl to generate the first alkyne **ALK01**. The second alkyne in the library, alkyne **ALK02**, was generated by conjugation of glutathione to alkyne **ALK01**. These alkynes closely resemble the aldehyde compounds used in the DCL described in the previous chapter. A third alkyne **ALK03** was synthesised from known GST inhibitor ethacrynic acid **56**.^[20] This was

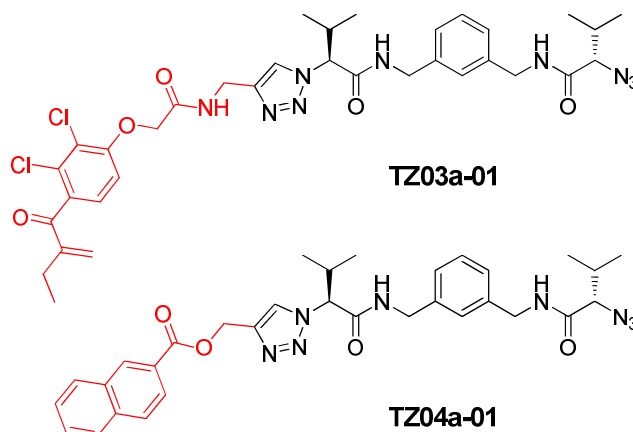
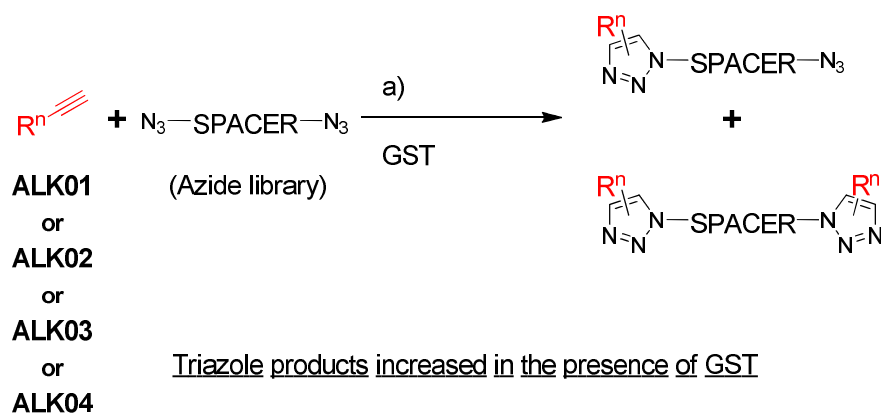
again achieved *via* EDC-coupling of the carboxylic acid group of ethacrynic acid with propargylamine **55**. A final alkyne **ALK04**, based on naphthalene, was added to the library and was selected as a more general group to bind in the hydrophobic pocket of the active site. Naphthalene is also known to be metabolised by GSTs.^[21] Naphthalene-2-carboxylic acid **57** was coupled to propargyl alcohol **58** to generate alkyne **ALK04**.



Scheme 2: Synthesis of the initial alkyne library **ALK01-ALK04**. a) 1.2 eq EDC.HCl, 0.2 eq DMAP, dichloromethane, room temperature, 18 h. b) 1 eq reduced glutathione, 1 eq sodium carbonate, 1:1 water/acetonitrile, room temperature, 48 h.

3.3 The initial *in situ* click chemistry reaction

The preliminary ISCC reaction was set up using the six member azide spacer library with each of the four alkynes.



Scheme 3: ISCC reaction using azide spacer library and alkynes **ALK01** to **ALK04**. Triazoles detected by LC/MS-SIM in the ISCC reaction are also shown, **TZ03a-01** and **TZ04a-01**. a) AZ (250 μM), ALK (250 μM), hGSTP1-1 or mGSTM1-1 or BSA (15 μM), 15 % DMSO, 10 mM ammonium acetate buffer, pH 7.0, 7 d, 25 $^\circ\text{C}$.

Each reaction contained one azide spacer with one alkyne. Two isoforms of GST, hGSTP1-1 and mGSTM1-1, were initially chosen to be investigated. hGSTP1-1 was one of the isoforms used by Atkins and co-workers in their search for bivalent inhibitors and mGSTM1-1 was the most successful isomer used in the DCL work described in the previous chapter.^[12-14] Control reactions containing bovine serum albumin, BSA, a control protein, were set up in tandem with the GST-catalysed reactions in order to measure any background triazole formation or any non-specific protein effects.

Fragments were added in large excess of the enzyme concentration (*ca.* 250 μM fragment / 15 μM enzyme), as is common in literature examples for fragments with low or unknown binding affinity for the target.^[22] An enzyme concentration of 15 μM was chosen to maximise the formation of triazole product to

ensure detection of the product. Reactions were set up and incubated at 25 °C; for seven days, after which time the reactions were quenched with methanol in order to remove the enzyme. Samples were then run on liquid chromatography mass spectrometry (LC/MS) in selected ion monitoring (SIM) mode. SIM mode was necessary to provide the required sensitivity to identify the product peaks as they would be at such small concentrations within the mixture. For each sample, the monotriazole and bistriazole product masses were scanned. Small scale thermal reactions between each alkyne and azide pairing were also set up to provide a reference sample for analysis.

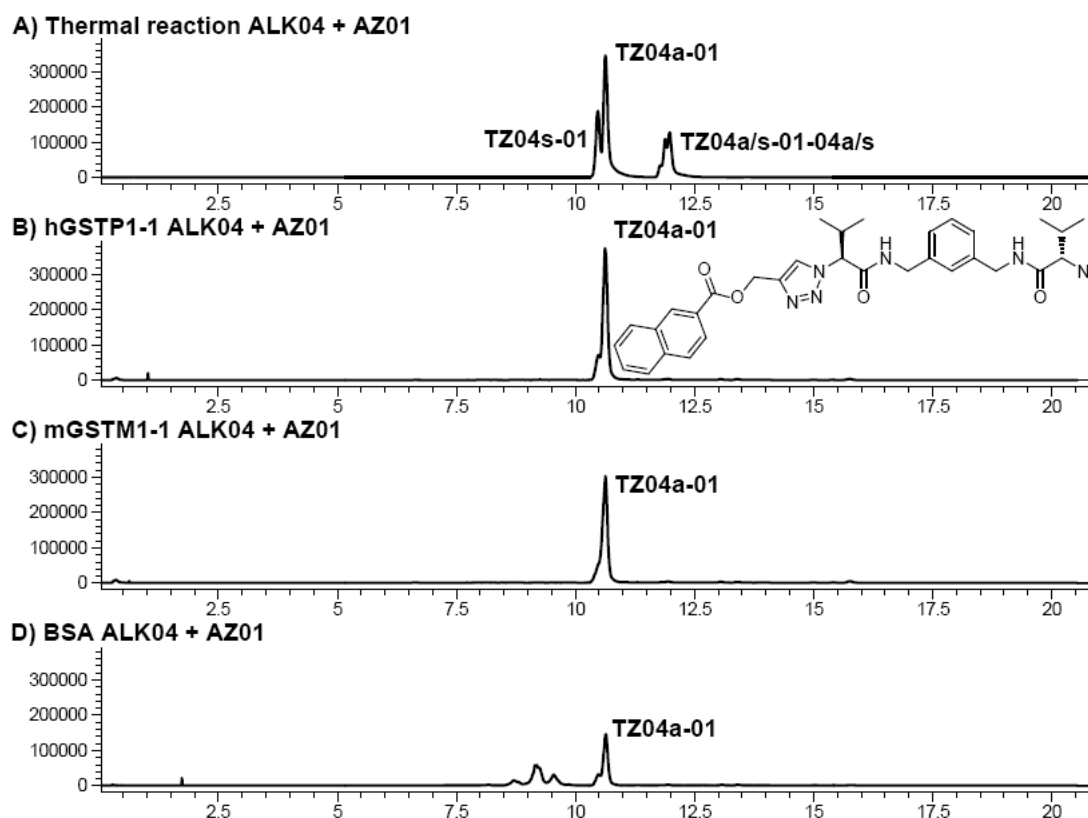


Figure 7: LC/MS-SIM traces for the ISCC reaction between **ALK04** and **AZ01** and reference traces. **A)** reference from the thermal reaction between **ALK04** and **AZ01**, **B)** ISCC reaction between **ALK04** and **AZ01** in the presence of hGSTP1-1, **C)** ISCC reaction between **ALK04** and **AZ01** in the presence of mGSTM1-1, **D)** ISCC reaction between **ALK04** and **AZ01** in the presence of BSA. ISCC reaction conditions: Azide (250 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST/BSA (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 300 μ L, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS. Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for

monotriazole and bistriazole products. Plots show TIC for selected ions versus retention time using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1).

An example of the LC/MS-SIM traces obtained for the reactions between **ALK04** and **AZ01** are shown in Figure 7. The increase in triazole formation in the presence of GST over the control reaction was calculated by dividing the peak area for the product in the presence of GST by the peak area for the product formed in the control. It is this ratio that is reported in subsequent ISCC reactions throughout this chapter.

Table 1: Results of the first ISCC screen between the azide spacer library and **ALK03** or **ALK 04** in the presence of hGSTP1-1 or mGSTM1-1. Data shows the ratio of the amount of monotriazole formed in the presence of GST over the amount of monotriazole formed in the presence of BSA, therefore a value of 1 represents no difference in triazole formation between the two templates. NT = no triazole observed

Azide spacer	ALK03 hGSTP1-1	ALK03 mGSTM1-1	ALK04 hGSTP1-1	ALK04 mGSTM1-1
AZ01	1.4	1.5	2.7	2.6
AZ02	NT	NT	1	1
AZ03	1	1	1	1
AZ04	1	1	1	1
AZ05	1	1	1	1
AZ06	NT	NT	1	1

For both of the alkyne fragments based on CDNB, **ALK01** and **ALK02**, no product peaks were observed. For the ethacrynic acid based alkyne, **ALK03** and the naphthalene based alkyne, **ALK04** peaks corresponding to the monotriazole products were observed in all but two examples, **AZ02** or **AZ06** with **ALK03**. However in the

case of only one azide spacer was the amount of triazole observed in the GST-catalysed reaction higher than that of the control reaction. For both **ALK03** and **ALK04**, this spacer was **AZ01**, the spacer with the valine side-chain. It was interesting to note that despite spacer **AZ02** being the same length as **AZ01**, no increase in triazole formation was observed over the control BSA reaction. This suggested that the additional hydrophobic bulk of the valine side-chain was required to gain enough binding affinity for GST. The formation of the bistriazole product was not observed in any of the reactions.

The observation that the monotriazole product formed in the presence of BSA was unexpected and could be due to product forming in the background reaction. In most examples of ISCC, background reactions are not observed. An exception is Sharpless' report on novel HIV protease inhibitors using ISCC where background reaction could be observed.^[22] The concentration of the alkyne and azide reacting fragments in Sharpless' report was much larger than had previously been used in ISCC as the reacting fragments had much weaker binding affinity for the biological target than in previous examples. It was proposed that the use of higher concentrations was the reason for the observation of product in the background reaction. The same high concentrations were used in the ISCC reactions described in this chapter and this could be why the product was seen in the control. A second possibility is that the product was not formed in the background, but was formed due to a non-specific effect from the presence of the hydrophobic environment of the protein in the reaction. This is discussed later in the chapter. Although the amplification in monotriazole formation was not very large in the presence of GST (1.4 – 2.7 times greater), the reactions between **AZ01** with **ALK03** and **AZ01** with **ALK04** were investigated further. There were several conditions in the ISCC reaction which needed optimizing and these are discussed below. For the optimization of the reaction only one GST isoform was used; hGSTP1-1 was selected as similar peptidic bivalent inhibitors **41** were known to bind with this isoform.^[14] It was also important to determine the cause of the triazole formation in the presence of BSA, in particular, the background reaction needed to be studied in the absence of any protein.

3.3.1 Optimising the ISCC reaction - the volume of DMSO.

A volume of 15 % DMSO in buffer was initially chosen for the ISCC reaction as this value was known to be suitable to maintain the activity of the GST enzyme.^[23] This volume had been previously used in DCC experiments within the Greaney group and the activity of the enzyme was unaffected for 48 hours. The ISCC reaction was incubated for up to seven days so the activity of hGSTP1-1, stored in 15% DMSO under the ISCC reaction conditions of 25 °C in ammonium acetate buffer at pH 7.2, was measured over a seven day time period.

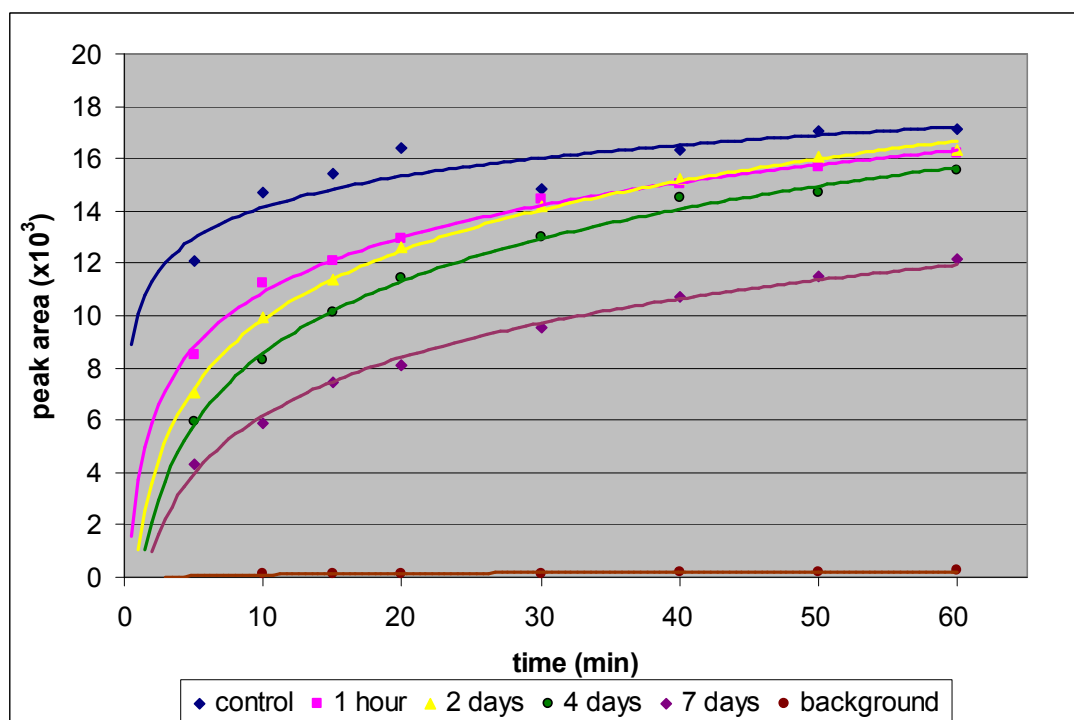


Figure 8: A graph to show the activity of hGSTP1-1 stored in 15 % DMSO over time. Activity is measured by monitoring the rate of formation of the glutathione conjugation product DNP GSH. Plot shows the peak area of GSH DNP, as measured using LC/MS-SIM (monitoring for the mass of DNP GSH), against the time of the reaction.

There was little effect on the activity of the enzyme up to four days after incubation of hGSTP1-1 in 15 % DMSO. After day four the activity started to decline, however after seven days the enzyme still had some activity, as evidenced in Figure 8.

The use of relatively large hydrophobic ligands resulted in some solubility issues at the higher concentrations (> 250 μ M) desired for the initial ISCC screens.

The effects of increasing the percentage of DMSO on the enzyme activity were measured to determine whether an increase in DMSO could be tolerated to aid ligand solubility.

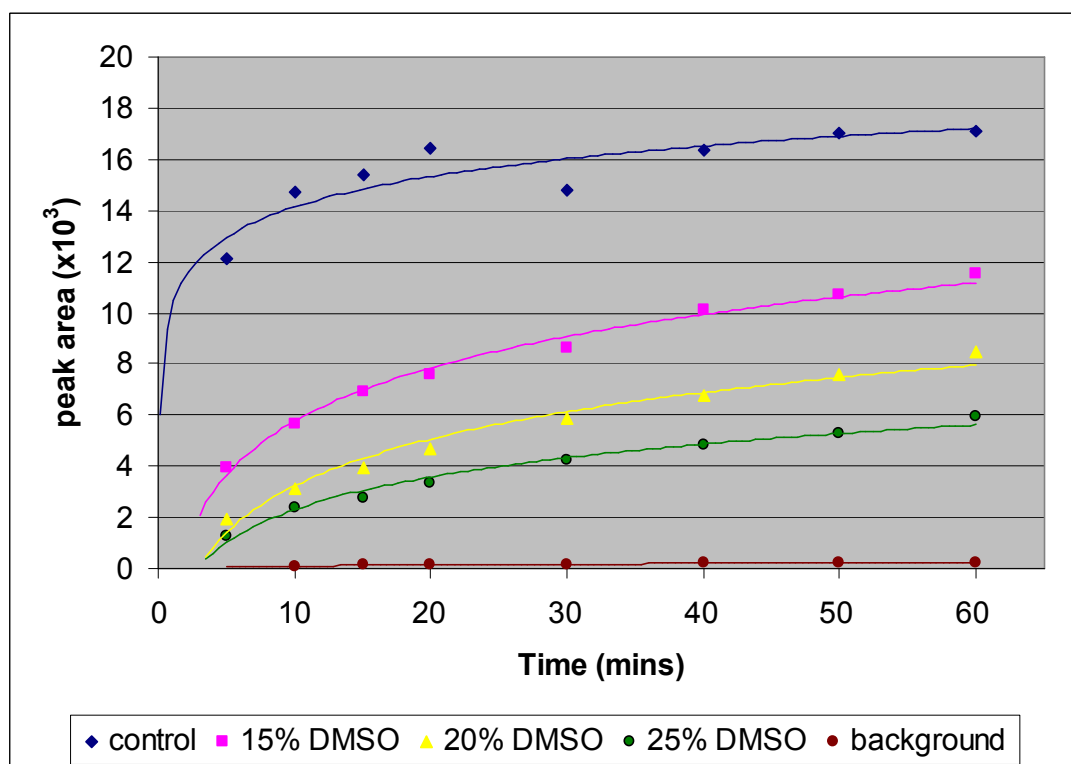


Figure 9: A graph to show the activity of hGSTP1-1 stored for seven days in increasing volumes of DMSO. Activity is measured by monitoring the rate of formation of the glutathione conjugation product DNP GSH. Plot shows the peak area of GSH DNP, as measured using LC/MS-SIM (monitoring for the mass of DNP GSH), against the time of the reaction.

The increase in the % volume of DMSO used resulted in a loss of activity of the enzyme, as shown in Figure 9. These increases in DMSO did not yield a notable improvement to the solubility of the azide and alkyne fragments, so the % volume of DMSO used was kept at 15 %.

3.3.2 Optimising the ISCC reaction – the time of the reaction

The initial assay was incubated for seven days to ensure that for any hits there would be enough triazole formed to be detected. The results of the hGSTP1-1 activity over time experiments would suggest that activity begins to decline after four days. Time course experiments were set up to monitor both the enzyme-templated reaction and the background reaction (no enzyme present) over the seven day period.

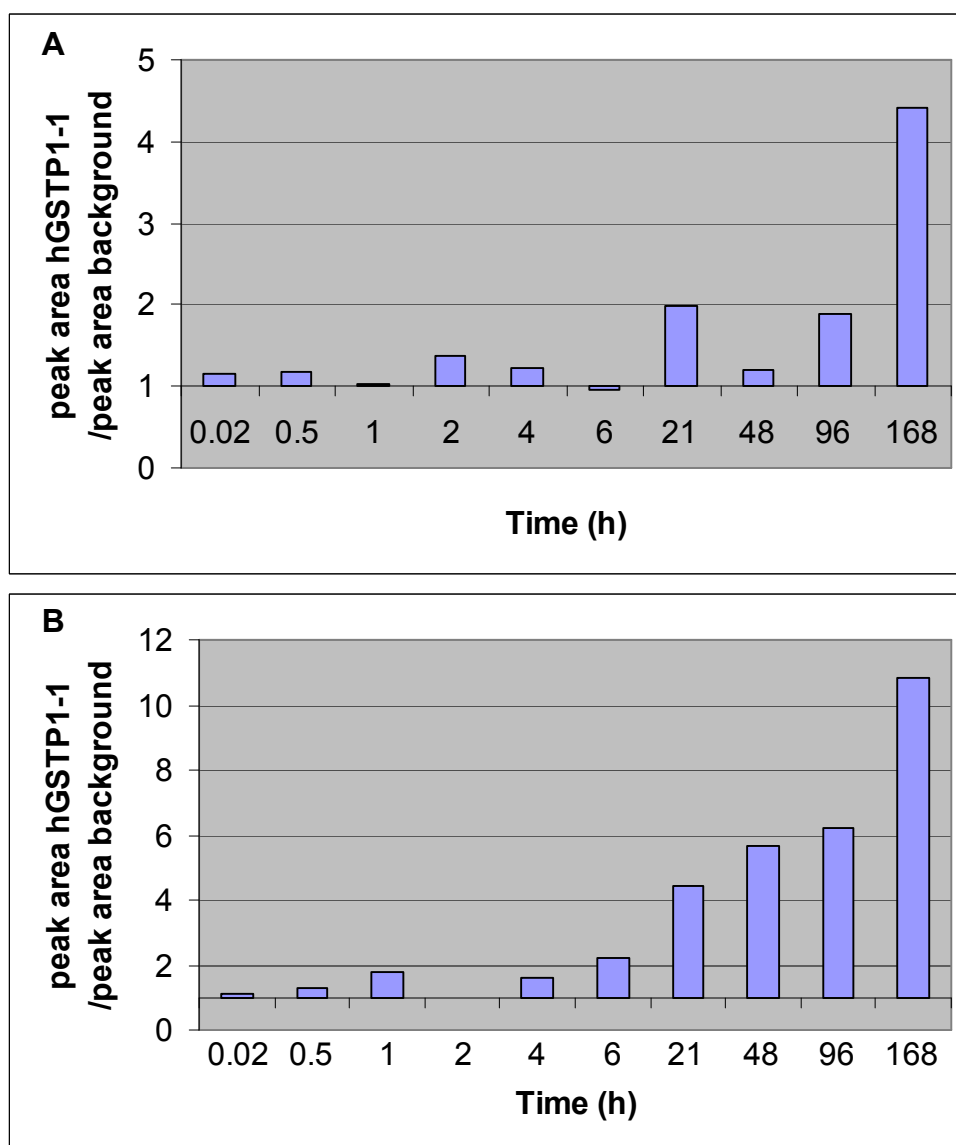


Figure 10: ISCC reaction monitored over time. Graphs show the ratio of triazole observed in the presence of hGSTP1-1 over the background reaction. Graph **A)** **ALK03 + AZ01**, Graph **B)** **ALK04 + AZ01**. ISCC reaction conditions: Azide (250 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 1.5 mL, 10 mM, pH 7.0) mixed at room temperature. Over several time points, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

Both the reactions of **ALK03** with **AZ01** and **ALK04** with **AZ01** were monitored over a seven day period with the results shown in Figure 10. The reaction between **ALK04** and **AZ01** in the presence of hGSTP1-1 resulted in a large

amplification of the monotriazole product after one day (over four times greater), with the ratio over the background increasing to 11:1 over the next six days. In the example using **ALK03** with **AZ01** the ratio between the protein-templated reaction and the background did not give a significant increase until seven days, when a ratio of 4:1 was achieved. This result indicated that despite a reduction in the activity of the enzyme, the ISCC reaction should be incubated for seven days to ensure any hits are observed.

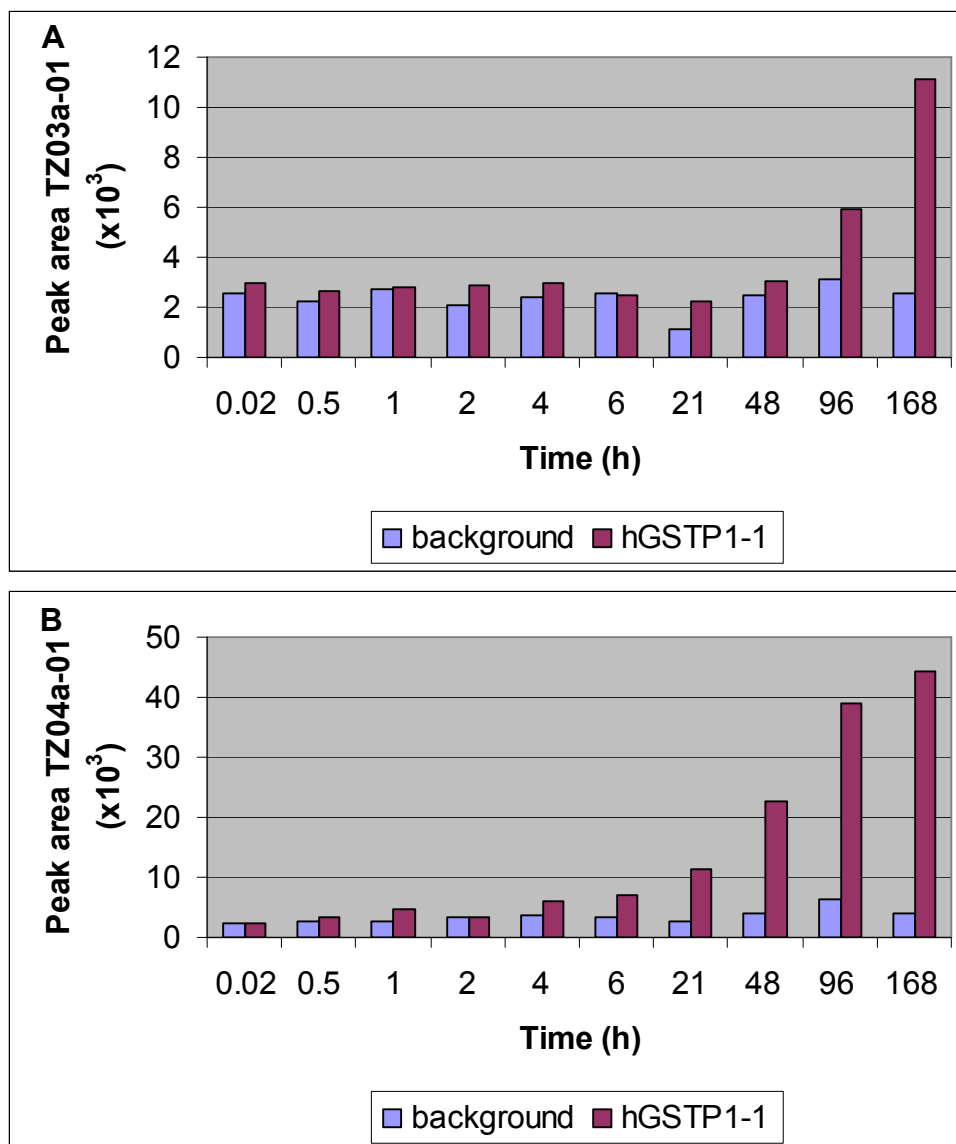


Figure 11: ISCC reaction monitored over time. Graphs show the peak area of the monotriazole products observed in the presence of hGSTP1-1 and in the background reaction. Graph **A**) **ALK03** + **AZ01**. Graph **B**) **ALK 04** + **AZ01**. ISCC reaction conditions: Azide (250 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST (15 μ M

final conc.) and ammonium acetate buffer (to total reaction volume of 1.5 mL, 10 mM, pH 7.0) mixed at room temperature. Over several time points, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

Another key point taken from the time course experiment was the observation of triazole in the background reaction, with no protein present. Figure 11 shows the peak area for the monotriazole products, **TZ03a-01** and **TZ04a-01** for both the background reaction and the reaction with hGSTP1-1. Interestingly the triazole was observed in the background reaction after only one minute though the concentration did not increase with time after this. This might suggest that there was an impurity in the reaction with the same retention time as the product, although it would also have to have the same mass as the product as the reaction was being monitored in the SIM mode. The individual components of the reaction mixture were all run through the LC/MS on the same gradient looking for the mass of the monotriazole product to try to determine the cause of any impurity present. The peak corresponding to the monotriazole was not observed in any of these separate components. It is unlikely for a different impurity to be present in each reaction, especially if each of those impurities has the same mass and retention time as the corresponding monotriazole. This means that the most likely explanation was that the peak was due to the presence of the monotriazole. The monotriazole must be forming upon mixing of the reaction components or alternatively whilst loading the reaction mixture onto the LC/MS column. Either way this small amount of background reaction seems to be unavoidable.

3.3.3 Optimising the ISCC reaction – fragment concentration

One possible way to reduce the background reaction was to reduce the concentration of the reacting fragments; the azide and the alkyne species. It was hoped that by making the reaction more dilute, the background reaction would fall below the detection limits of the LC/MS. Of course the reduction of the reacting fragments may also reduce the amount of triazole formed in the presence of the GST protein to below detectable limits.

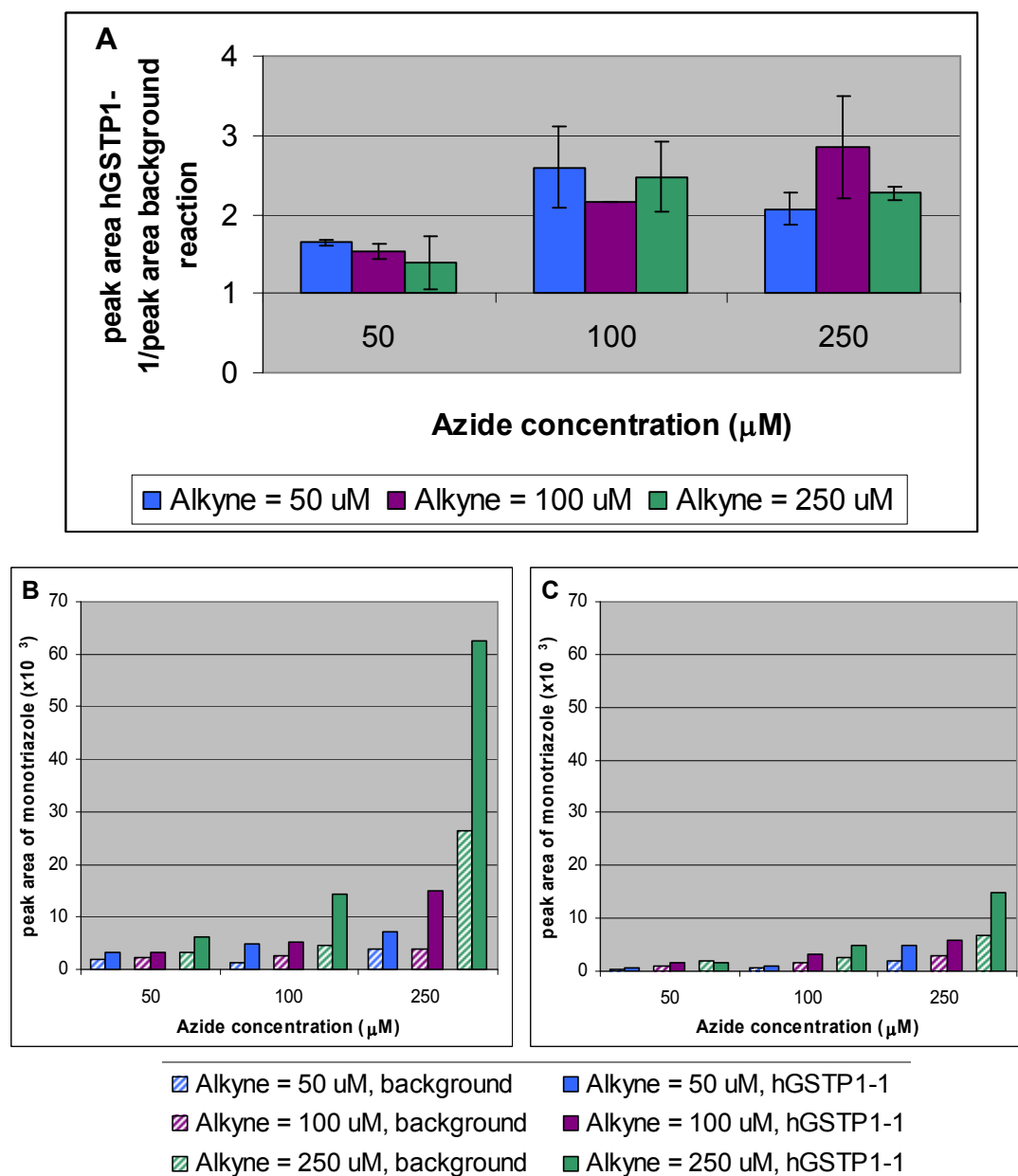


Figure 12: ISCC reaction monitored with different concentrations of azide **AZ01** and alkyne **ALK04**. **Graph A)** the ratio of triazole observed in the presence of hGSTP1-1 over the background reaction. Error bars represent standard deviation. **Graph B** and **C)** the peak area of the monotriazole products observed in the presence of hGSTP1-1 and in the background reaction for two separate assays with identical reaction conditions showing the variation observed in the peak area of the triazole product. ISCC reaction conditions: Azide (50/100/250 μM final conc.), Alkyne (50/100/250 μM final conc.), DMSO (15 % by vol.), GST (15 μM final conc.) and ammonium acetate buffer (to total reaction volume of 0.5 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μL of reaction quenched in 400 μL

methanol, filtered and 10 μL injected onto LC/MS using gradient C (see Chapter 5, Section 5.2.2). Products detected in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for monotriazole and bistriazole products.

The reaction between **ALK04** and **AZ01** was studied over a range of alkyne and azide concentrations, from 50 μM to 250 μM . The ratio of monotriazole **TZ04a-01** observed in the presence of hGSTP1-1 over the background reaction is shown in Figure 12A. A ratio of two or above was only observed for concentrations of azide at 100 μM and higher. The concentration of alkyne used made little difference on the ratio of product observed. The peak areas of **TZ04a-01** are shown in Figure 12B and C and suggest that the higher concentrations of the fragments would ensure more reliable product detection. The lower concentration resulted in very small peaks, which would increase the error when calculating the ratio between the peaks. It was disappointing to see that despite repeating the reaction several times, the higher ratios that had been observed in the time course experiments could not be repeated. Another issue became apparent when analysing the peak areas in Figure 12B and C, Graph B and C show the peak areas observed in repeat experiments. The conditions of the reaction were identical, with the only difference being the date that they were monitored on the LC/MS. For the reaction with 250 μM concentration for both fragments the peak area for the reaction in the presence of hGSTP1-1 is four times greater in experiment B compared to experiment C. This observation was noted many times during the course of this work and seemed to be a limitation with the ionization of the mass spectrometer. When samples with known concentrations were run on the LC/MS on different days the peak area for the product was seen to vary by large amounts (60×10^3 in experiment B and 15×10^3 in experiment C). Importantly however, the ratio between the background and the reaction containing the protein was unaffected between the two experiments even though the size of the peak area changed drastically. It became clear during the optimization reactions that while the peak areas could not be compared between experiments monitored days apart, those monitored on the same day could be reliably compared. For all subsequent reactions, any ratios that are quoted were calculated from reactions run at the same time and repeat results were calculated from taking the average of the ratio and not the average of the peak areas.

3.3.4 Optimising the ISCC reaction – enzyme concentration

The final variable investigated was the concentration of enzyme used. Both the reactions of **AZ01** with **ALK03** and **AZ01** with **ALK04** were studied and the results are shown in Figure 13.

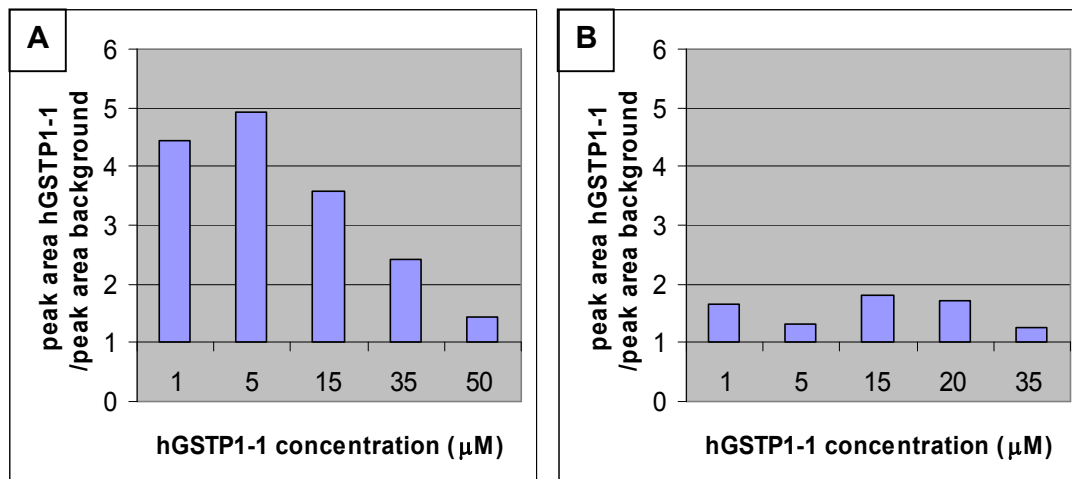


Figure 13: ISCC reaction with varying concentrations of hGSTP1-1. Graphs show the ratio of triazole observed in the presence of hGSTP1-1 over the background reaction. Graph **A**; **ALK03** + **AZ01**. Graph **B**; **ALK 04** + **AZ01**. ISCC reaction conditions: Azide (250 µM final conc.), Alkyne (250 µM final conc.), DMSO (15 % by vol.), GST (1/5/15/20/35/50 µM final conc.) and ammonium acetate buffer (to total reaction volume of 0.8 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 µL of reaction quenched in 400 µL methanol, filtered and 10 µL injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

The results for the reaction with **ALK03** suggested that an increase in protein resulted in a decrease in triazole formation. This could be because the protein becomes unstable at higher concentrations. Worryingly, in the reaction between **ALK04** and **AZ01**, there was little amplification in the formation of triazole in the presence of hGSTP1-1 regardless of the concentration used. This is unexpected as ratios up to 12:1 had been previously observed under the same conditions. This might indicate that in some reactions the enzyme is losing its activity or structure. For each ISCC reaction set up subsequently, a sample of the enzyme was set up under the same reaction conditions, at the same temperature, with the same

concentration of DMSO. The activity of the enzyme was monitored at the same time as the ISCC reaction was quenched and analysed.

3.4 The optimised ISCC reaction for AZ01 with ALK03 and ALK04

After studying the variables in the ISCC reaction, the initial conditions seemed to be suitable so no changes were made to the original conditions for subsequent reactions. A more important point arose during the optimisation experiments with the inconsistent results of the protein-templated reaction. Values for the ratio between the protein-containing reaction and the background varied from one to twelve. These reactions were repeated several times to gain a reliable value for the amplification in triazole formed in the presence of the protein. The reaction was also set up for two more isoforms of GST; mGSTM1-1 and SjGST.

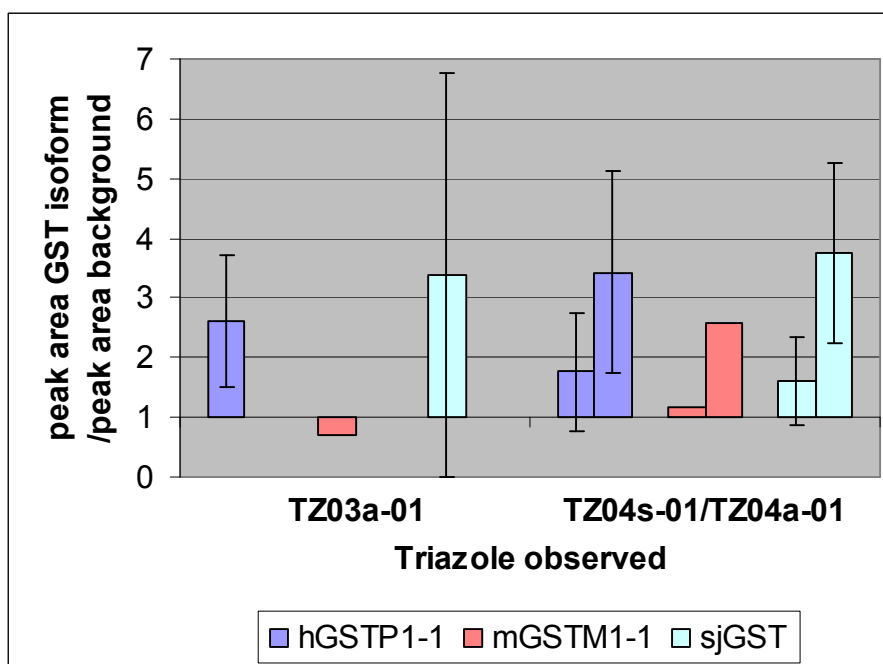


Figure 14: ISCC reaction for the reaction between **AZ01** and **ALK03** and the reaction between **AZ01** and **ALK04** in the presence of various isoforms of GST. Graph shows the ratio of triazole observed in the presence of GST over the background reaction. Error bars represent standard deviation. ISCC reaction conditions: Azide (250 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto

LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

All of the reactions, both for **ALK03** and **ALK04**, and for all three GST isoforms resulted in an increase in monotriazole in the presence of GST compared with the background reaction with the exception of the reaction between **ALK03** and **AZ01** in the presence of mGSTM1-1. The error bars showing the standard deviation for the results were very large showing the unreliability of this reaction. As has been previously discussed, there are only two examples in the literature where the background click reaction is observed. The first was in an article describing ISCC approaches to discovering HIV-1 protease inhibitors.^[22] In this article there is no statistical value stated to quantify a hit from the reaction. Any hit is described as being an amplification in the amount of triazole product observed over the background reaction. The second example is the ISCC identification of Chitinase inhibitors.^[24] In this case however, one triazole regioisomer, the *anti*-triazole, is observed in the background but it is the other regioisomer, the *syn*-isomer, which is observed in the templated reaction, with no increase in formation of the *anti*-triazole over the background reaction. Work published by Xiangdong Hu *et al.* may provide a guideline to the statistical values that constitute a protein-templated hit.^[25] In their work Hu and co-workers use the reaction between sulfonyl azides and thio acids in a target-guided approach to develop novel Bcl-XL inhibitors. Unlike the 1,4 dipolar cycloaddition reaction, a background reaction is always observed. A ratio of 4:1 increase in product formation in the presence of the protein template over a control was used to determine a hit in this example. Though this is for a different reaction, the ratio was calculated in part to eliminate the errors associated with the detection technique of LC/MS in SIM mode, the same detection method that was employed in this work. Certainly a ratio below 2:1 could not indicate a hit in this assay based on the changes in peak area observed for samples of known concentrations run on the LC/MS in SIM mode. Further controls were set up to try to build up a clearer picture of what was happening with the ISCC reaction.

3.4.1 Further controls for the ISCC reaction between AZ01 and ALK03 or ALK04

An amplification of triazole formation was being observed in the presence of enzyme over the background reaction. It now had to be proven that it was the enzyme's active site templating the click reaction rather than a non-specific hydrophobic effect from the enzyme. Two types of controls were used. The first was to replace the GST protein with a non-specific protein, in this case BSA, in a repeat of the initial ISCC reactions run to improve the reliability of the observed amplifications in triazole formation.

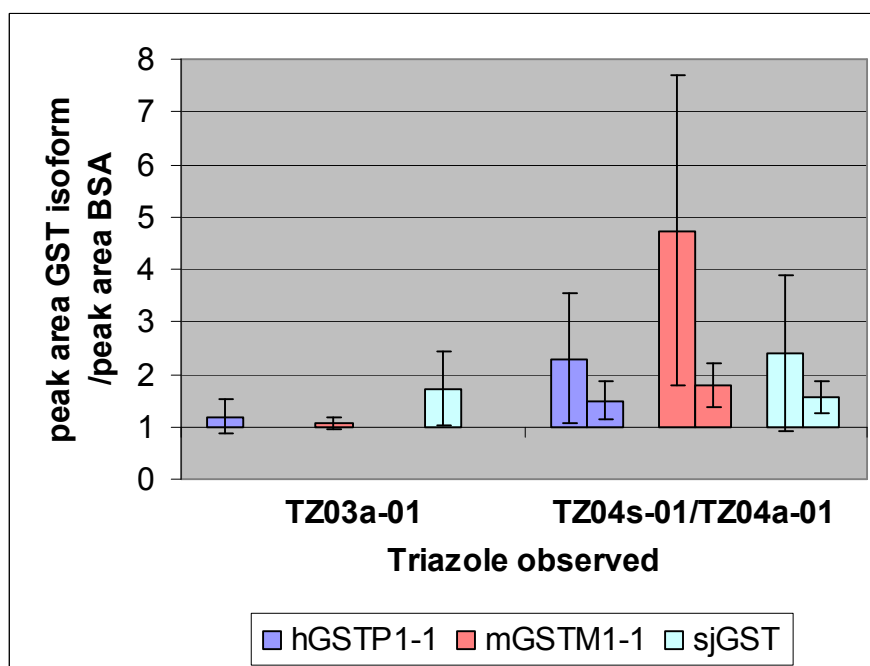


Figure 15: ISCC reaction for the reaction between **AZ01** and **ALK03** and the reaction between **AZ01** and **ALK04** in the presence of various isoforms of GST compared with the reaction in the presence of BSA. Graph shows the ratio of triazole observed in the presence of GST over the BSA reaction. Error bars represent standard deviation. ISCC reaction conditions: Azide (250 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST/BSA (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

The results of the increase in triazole observed in the presence of GST over BSA are shown in Figure 15. Disappointingly in the reaction between **AZ01** and **ALK03**, there was little amplification in the amount of triazole **TZ03a-01** observed in the presence of all three GST isoforms compared with BSA. This would suggest that the amplification that was observed over the background reaction was the result of a non-specific effect due to the presence of a protein in the reaction mixture. The results were more interesting for the reaction between **AZ01** and **ALK04**. The formation of **TZ04s-01** was more than double in the presence of either hGSTP1-1 or SjGST compared with BSA and four times greater in the presence of mGSTM1-1 than in the presence of BSA. Unlike the previous reactions it was the 1,5 triazole that showed a greater increase in amplification as opposed to the 1,4 triazole. The error bars in the data comparing the GST reaction with the BSA reaction were very large so this result could be misleading.

The second control utilised a known inhibitor of GST to block the binding pocket and thus prevent any templating effects from the binding pocket. As described previously in chapter 2, the hydrophobic pocket in GST is large and promiscuous. Not all of the inhibitors of GST were able to block the amplifications observed in the DCC reaction; therefore DNP GSH was selected as this had demonstrated the best effect on blocking the DCL described in chapter 2. Ethacrynic acid was also used as an inhibitor in the ISCC reaction as **ALK03** is synthesised from ethacrynic acid so it might be expected to bind to GST in a similar way.

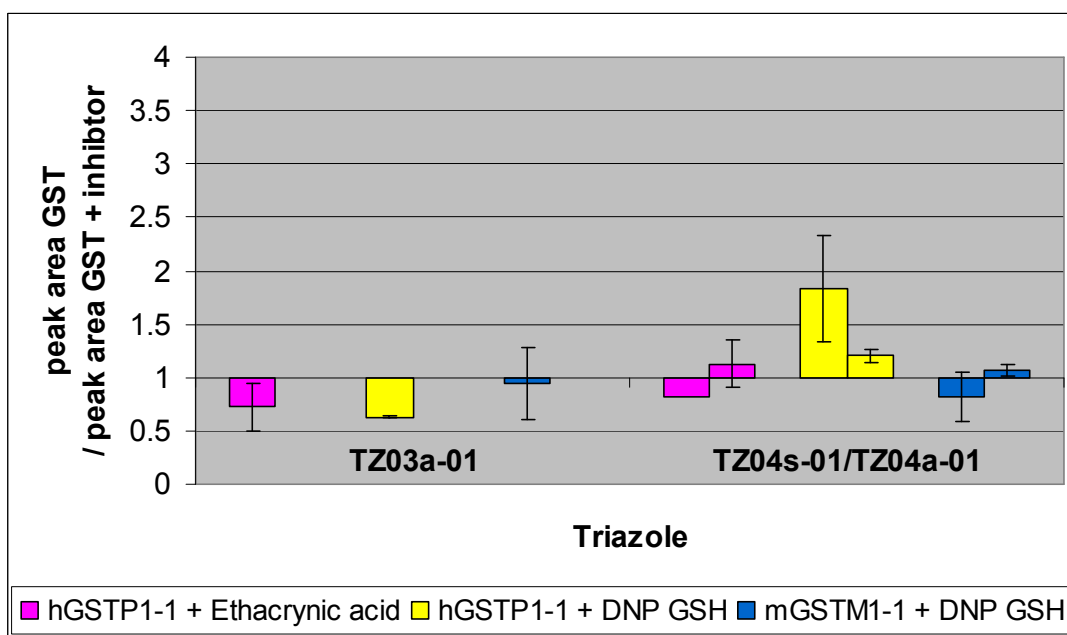
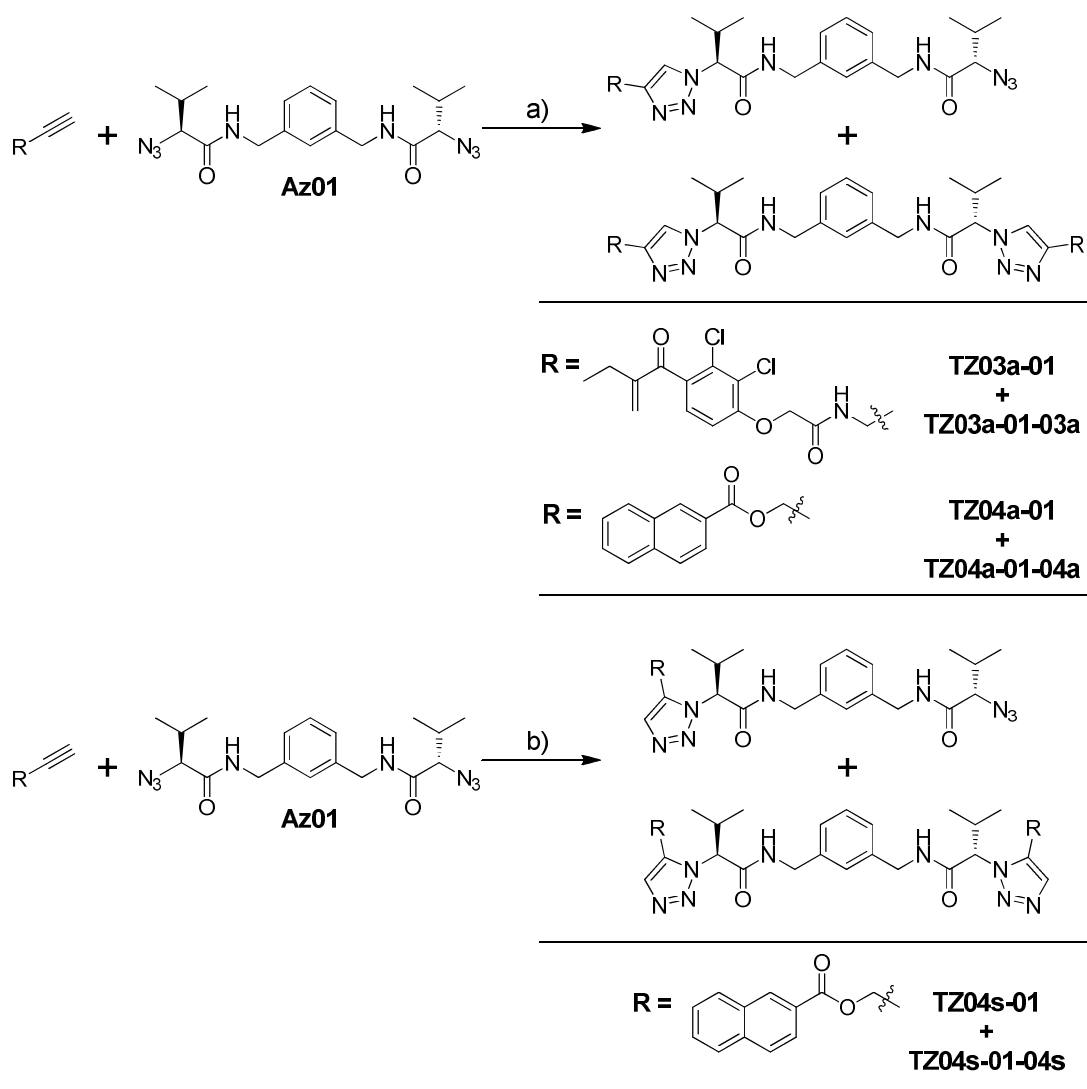


Figure 16: ISCC reaction for the reaction between **AZ01** and **ALK03** and the reaction between **AZ01** and **ALK04** in the presence of hGSTP1-1 or mGSTM1-1 with the addition of GST inhibitors ethacrynic acid and DNP GSH. Graph shows the ratio of triazole observed in the presence of GST over the reactions in the presence of GST + inhibitor. Error bars represent standard deviation. ISCC reaction conditions: Azide (250 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), inhibitor (500 μ M final conc.) GST (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

The introduction of inhibitors to block the binding pockets generally had little effect on the amount of triazole forming. The inhibitor that might be expected to block the GST-templated triazole formation is the glutathione conjugate of the substrate CDNB, labelled DNP GSH, but as Figure 16 highlights there was no amplification above 2:1. This result adds support to the argument that the amplifications in the triazole formation that were observed in the presence of GST over the background were as a result of a non-specific effect.

3.4.2 Synthesis and binding studies of the triazole products observed in the ISCC reaction

The results of the control reactions were unclear, but seemed to be suggesting that the increase of triazole production in the presence of GST was as a result of a non-specific effect and not by the GST acting as a template in the reaction. To try to gain a better understanding of the results seen in the ISCC reaction, the monotriazole products that were observed were synthesised with a view to measuring their binding affinity for GST. The reaction was carried out in one step, using copper (I) iodide to synthesise the 1,4 triazole, Scheme 4. Both the monotriazole and bistriazole products were synthesised in one pot by reacting the bisazide and alkyne in a one to one ratio, Scheme 4. The products were easily separated by column chromatography. As a result of the observation using the BSA control that the 1,5 triazole **TZ04s-01** was preferentially synthesized in the presence of GST, this triazole was also synthesized using a ruthenium catalyst $\text{Cp}^*\text{RuCl}(\text{PPh}_3)_2$.^[26]



Scheme 4: Synthesis of mono- and bistriazoles from **ALK03** and **ALK04** with **AZ01**. a) 1:1 ratio of alkyne to azide, 0.2 eq copper (I) iodide, acetonitrile, rt, 18 h. b) 1:1 ratio of alkyne to azide, 0.1 eq Cp^{*}RuCl(PPh₃)₂, dioxane, 60°C, under N₂ atmosphere, 18 h.

With the monotriazoles in hand, their binding affinity for all of the GST isoforms tested in the ISCC reaction was measured using the CDNB binding assay.^[19]

Table 2: IC₅₀ values for the alkynes monitored and the triazole products detected in the ISCC reaction. Data calculated using the CDNB binding assay. Results were calculated from three experiments. Raw data described in experimental chapter. (-) = assay not performed.

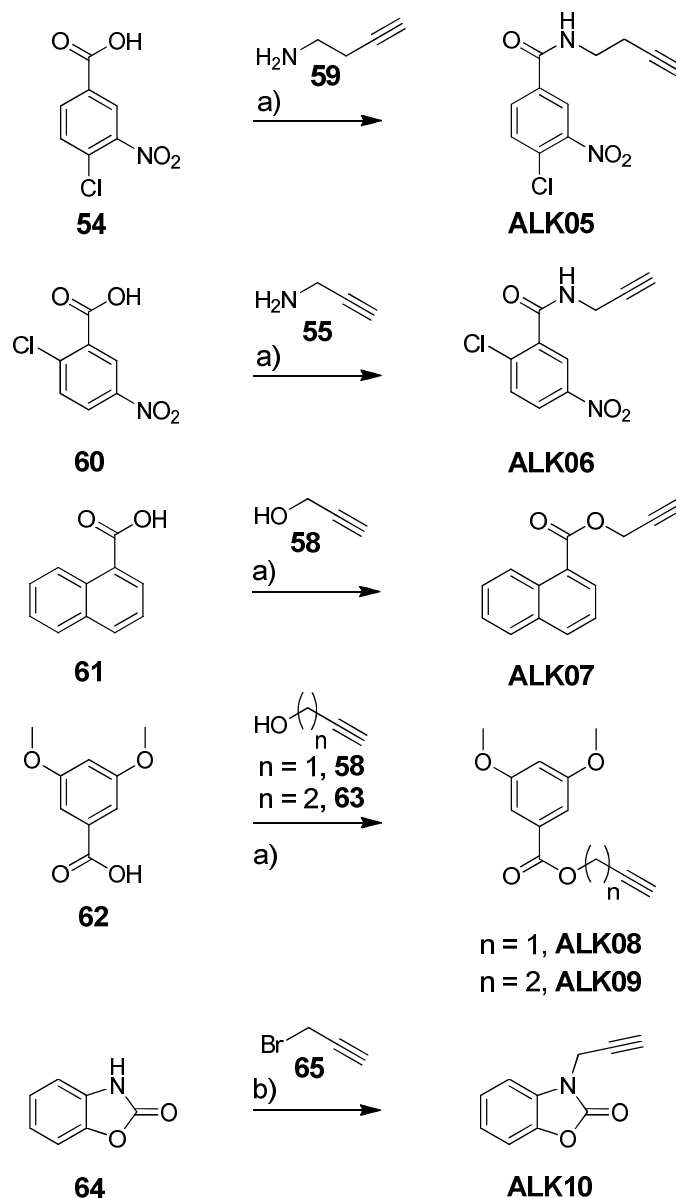
Compound	hGSTP1-1 IC ₅₀ (μM)	mGSTM1-1 IC ₅₀ (μM)	SjGST IC ₅₀ (μM)
ALK03	27.2	1.1	10.5
TZ03a-01	29.6	0.14	7.97
ALK04	≥ 100	≥ 100	-
TZ04a-01	≥ 100	129.3	≥ 100
TZ04s-01	-	≥ 50	-

TZ03a-01 showed good inhibitory activity for GST, particularly for the M1 isoform. This might be expected, as the product was based on ethacrynic acid, a known inhibitor of GSTs. There is a 10-fold increase in binding affinity for mGSTM1-1 by adding the azide spacer **AZ01** to alkyne **ALK03**. In the case of hGSTP1-1 and SjGST there was no gain in binding affinity by addition of the spacer **AZ01** to **ALK03**. This suggests that the spacer unit was accessing a binding site present in the mGSTM1-1 isoform that was absent in the hGSTP1-1 or SjGST isoforms. Both alkyne **ALK04** and the monotriazole product **TZ04a-01** have little binding affinity for any of the GST isoforms with IC₅₀ values greater than 100 μM. This failure to correlate inhibitory activity with the small amplifications observed in the ISCC reaction suggests that it was unlikely that the triazole was formed by a genuine templative effect from the hydrophobic binding pocket of GST. This provides evidence that the increase in triazole formation in the presence of GST over the controls used must be greater than 4:1 for the ISCC reaction to identify a hit.

3.5 Increasing the alkyne library

The inhibition data for the products resulting from the reaction between **AZ01** and **ALK03/ALK04** did not match the results seen in the ISCC reaction. To further explore the relationship between ISCC and inhibitory activity, the alkyne

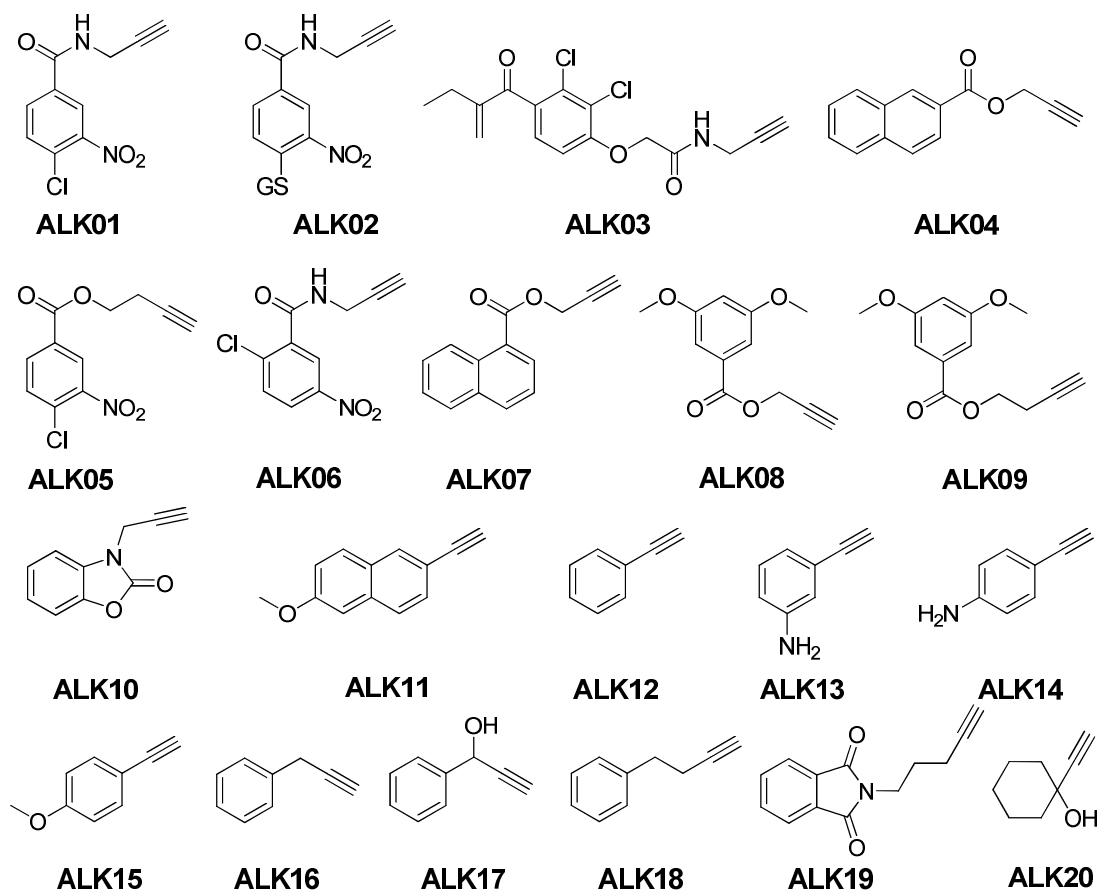
library was expanded to study the role of this fragment in the *in situ* reaction.



Scheme 5: Synthesis of extended alkyne library. a) 1.2 eq EDC.HCl, 0.2 eq DMAP, dichloromethane, rt, 18 h. b) 1.5 eq DIPEA, acetonitrile, 60 °C, 18 h.

Two more alkynes were synthesized based on CDNB, varying the length of the alkyne linker, **ALK05** and the position on the benzyl ring, **ALK06**. Another naphthyl alkyne was synthesized, this time with the alkyne added in the one position of the naphthalene ring, **ALK07**. Three more alkynes were synthesized in one step; **ALK08** and **ALK09** were again made *via* the EDC-coupling of 3,5-dimethoxybenzoic acid **62** with propargyl alcohol **58** or butyne-3-ol **63**. **ALK10** was made by an S_N2 reaction of propargyl bromide **65** to benzoxazolinone **64**. The

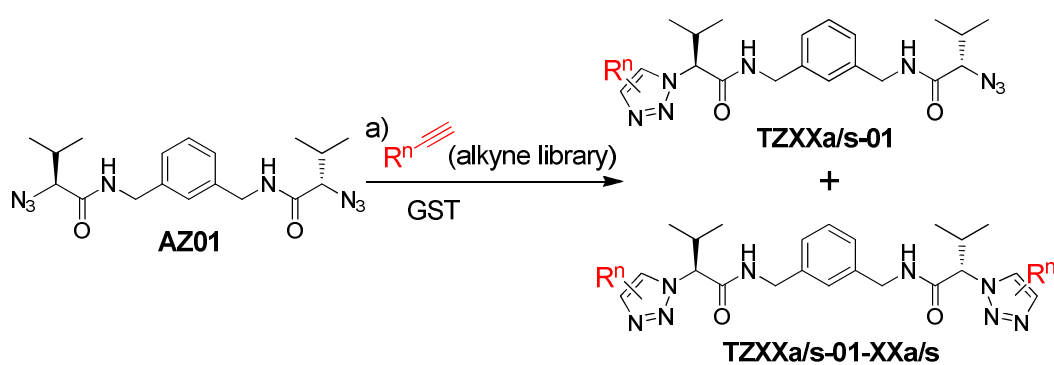
library was made up to twenty members by commercially available alkynes.



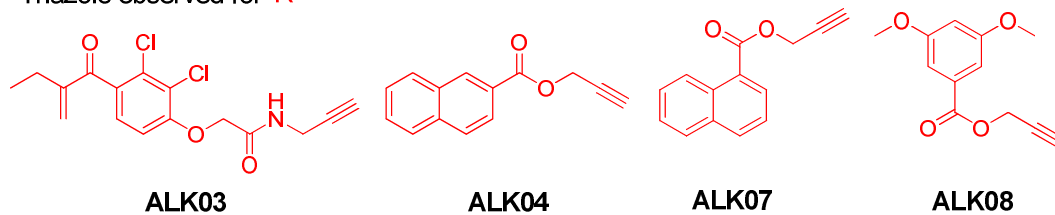
Scheme 6: Alkyne library

3.5.1 ISCC reaction with AZ01 and the 20-member alkyne library

This alkyne library was set up under the ISCC conditions with the azide spacer **AZ01**. One alkyne was added per reaction. Each combination was set up in duplicate; one reaction contained hGSTP1-1 whilst the other reaction was kept blank, with no protein added, in order to monitor the background reaction.



Triazole observed for $R^n \text{—} \text{C} \equiv \text{C—H} =$



Scheme 7: ISCC screen using **AZ01** and alkyne library. Alkynes resulting in the triazole product being observed in the reaction are highlighted in red. a) **AZ01** (250 μM), **ALK** (250 μM), hGSTP1-1 (15 μM), 15 % DMSO, 10 mM ammonium acetate buffer, pH 7.0, 7 d, 25 $^{\circ}\text{C}$.

The results indicated that triazole product was formed in the presence of protein with four alkynes, the two alkynes that had previously been studied **ALK03** and **ALK04**, and two further alkynes **ALK07** and **ALK08**. Triazole product was also detected in the background reaction, but at a much smaller quantity than in the presence of hGSTP1-1, as evidenced in Figure 17.

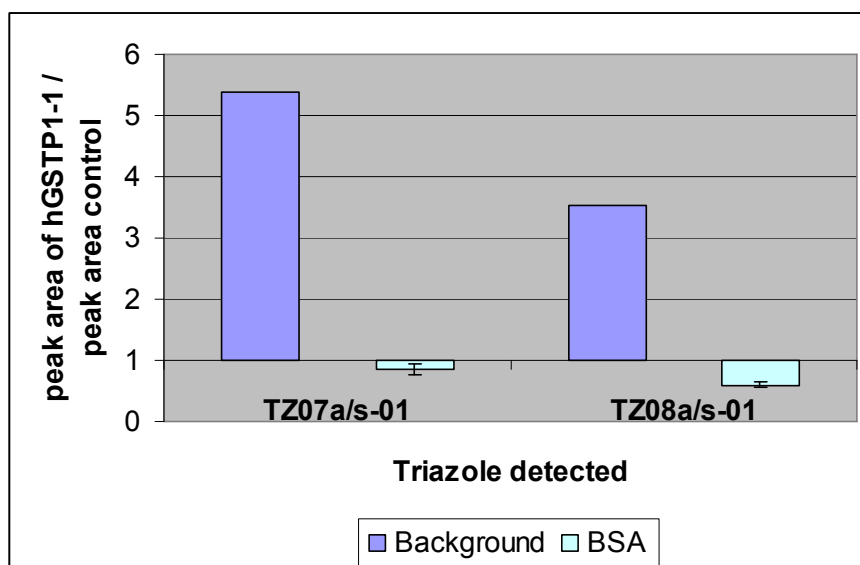
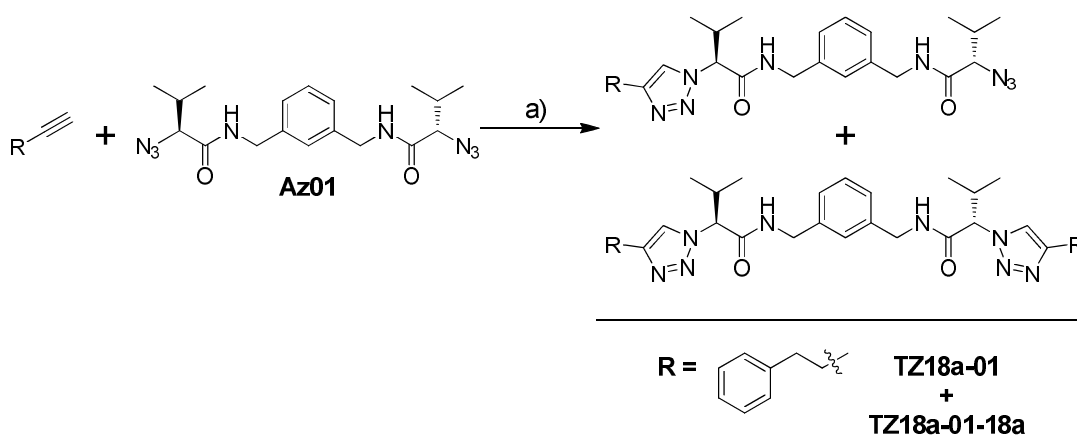


Figure 17: ISCC reaction between **AZ01** and **ALK07** and between **AZ01** and **ALK08** in the presence hGSTP1-1. Graph shows the ratio of triazole observed in the presence of hGSTP1-1 over the background reaction and over the reaction containing BSA. Error bars represent standard deviation. ISCC reaction conditions: Azide (250 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST/BSA (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

The successful combinations were repeated in the presence of BSA. In all cases, the amount of triazole formed was very similar in the presence of BSA or GST. This would indicate that the effect was down to the hydrophobic environment of the protein in the reaction solution, rather than a specific templative effect from the binding pocket of GST. What seems strange with this proposal was that only four of the alkynes reacted with the azide but for sixteen of the alkynes, no triazole formed. There appeared to be no trend within the successful alkynes to distinguish them from those that did not react. Many of the alkynes contained an aromatic ring. There were library members containing alkynes linked by an amide or ester to the aromatic ring that reacted and others that did not react. There was no observable common feature present amongst the alkynes that reacted in the presence of protein and those that did not.

3.6 Investigating the sensitivity of LC/MS-SIM

It was assumed with the LC/MS-SIM detection method that all of the triazole products would ionise to a similar extent, but it could have been that some hits were not detected because the products were not as easily ionised as others. To check this, the triazole products of **ALK18**, which were not detected in the ISCC reaction, were synthesized using copper (I) catalysis (Scheme 8).



Scheme 8: Synthesis of **TZ18a-01** and **TZ18a-01-18a**. a) 1:1 ratio of alkyne to azide, 0.2 eq copper (I) iodide, acetonitrile, rt, 18 h.

The MS detection for these compounds was then compared with those for the products of the reaction with **ALK04**, one of the compounds that was detected during the ISCC reaction, section 3.4. The ionisation for **AZ01**, **ALK04** and **ALK18**, and their respective mono-triazole products, **TZ04a-01** and **TZ18a-01**, and bistriazole products, **TZ04a-01-04a** and **TZ18a-01-18a** was measured over a range of concentrations.

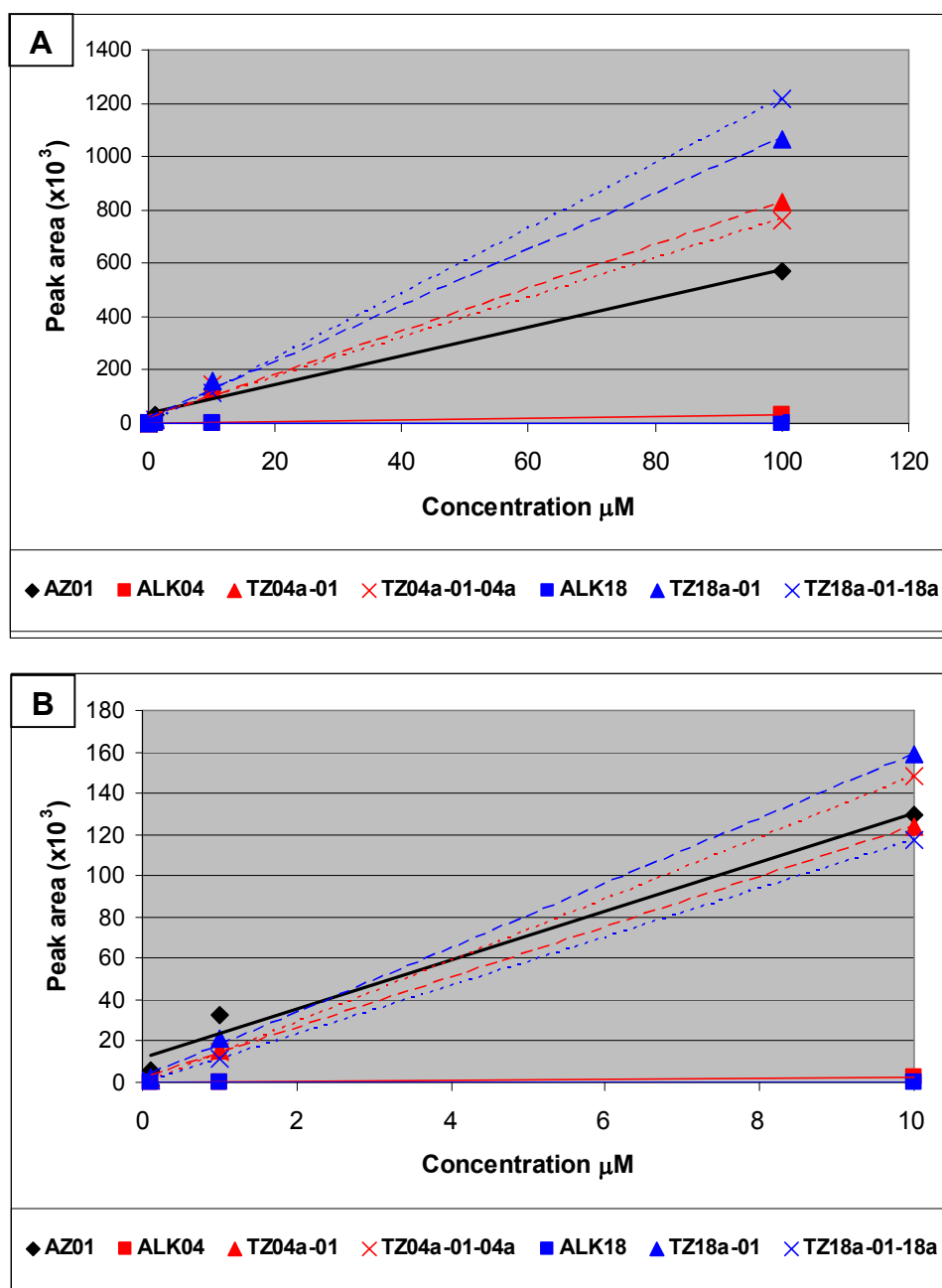


Figure 18: Graphs to show the LC/MS-SIM response to **AZ01**, **ALK04**, **ALK18** and the triazole products from these fragments. Graph B is a zoomed in version of graph A highlighting the concentration range of 100 nM to 10 μM. Concentration (along x axis) is the concentration of the analysed compound under the same conditions as the *in situ* click reactions, prior to dilution and injection onto LC/MS as per the *in situ* click reaction conditions. Peak area (along y axis) is the peak area for the total ion count of ions at the selected mass of the product being analysed.

As shown in Figure 18 neither of the alkyne fragments (blue and red squares) ionised very well, however the azide spacer **AZ01** (black diamonds) was observed at

low micromolar concentrations. All of the triazole products were more easily detected than the azide spacer. Pleasingly, the triazoles (blue triangle and cross) that were not observed in the ISCC reaction, **TZ18a-01** and **TZ18a-01-18a**, ionized to a greater extent than those triazoles (red triangle and cross) that were observed in the ISCC reaction. This confirms that should a triazole be formed, whether it was just the background reaction or as a result of the presence of protein in the reaction mixture, then it would be detected under the LC/MS-SIM conditions utilised in the reaction.

3.7 Binding data of triazole **TZ18a-01**, a triazole not formed in the ISCC reaction

Having synthesised the monotriazole **TZ18a-01**, which was not observed in the ISCC reaction, this monotriazole was run in the CDNB assay to measure its binding affinity for GST. This could then be compared with the binding data for **TZ03a-01** and **TZ04a-01**, the triazoles that were observed in the ISCC reaction. The binding affinity of **TZ18a-01** was only measured for mGSTM1-1, the GST isoform for which the monotriazoles previously tested had shown the best inhibition.

Table 3: A table showing the IC_{50} values for monotriazoles measured in the CDNB assay. Values are calculated from three separate assays. Raw data described in experimental chapter.

Triazole	IC_{50} (μ M) mGSTM1-1	Observed in ISCC reaction
TZ03a-01	0.138	✓
TZ04a-01	129.3	✓
TZ18a-01	≥ 100	x

The inhibitory activity for **TZ04a-01**, which was observed in the ISCC reaction, was very small (129.3 μ M) and showed very little difference to that for **TZ18a-01** (> 100 μ M). This provides further evidence to suggest that the triazoles observed in the ISCC reaction were forming due to a non-specific hydrophobic effect from the presence of the protein in the solution and not from a template effect from the binding pocket of the protein. The IC_{50} value of **TZ03a-01** was 1000-fold better

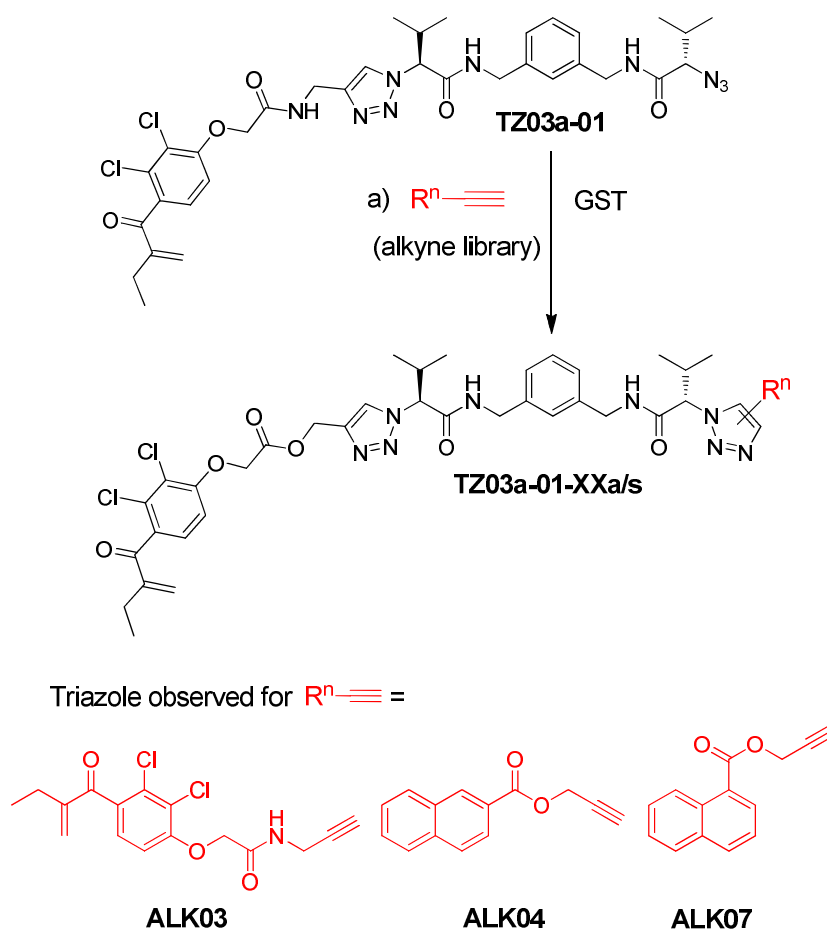
than the negative control triazole **TZ18a-01**. This result might suggest that **TZ03a-01** would prove to be a hit in the ISCC reaction, however the ratios in the amount of triazole formed in the presence of GST over BSA were very similar for both **TZ03a-01** and **TZ04a-01**. The binding data suggested that it was very unlikely that **TZ04a-01** was forming as a result of a templative effect from the GST, therefore this would suggest that **TZ03a-01** was still forming as a result of a non-specific effect.

The ISCC reaction failed to yield any positive hits that involved double cycloaddition with both azide groups on the spacer. This is perhaps not surprising as it would require three weakly binding fragments to be brought together in the protein's active site and this may be beyond the scope of this particular set of ISCC reactions, or even any ISCC approach.

In particular, the bisazide spacers were very weakly binding, as they would bind in the solvent-accessible cleft and not in a well defined binding pocket. In order to add some binding affinity to the spacers, a monoazide compound with one of the alkyne fragments already attached to the spacer was selected to trial in the ISCC reactions. The monoazide that was selected was **TZ03a-01** as this had by far the best binding affinity for GST.

3.8 *In situ* click chemistry screen of TZ03a-01 with alkyne library

An initial screen of **TZ03a-01** with the 20-member alkyne library and hGSTP1-1 as the protein template highlighted three combinations in which triazole formation was observed. These were the reactions with **ALK03**, **ALK04** and **ALK07**. Triazole was observed both in the presence of hGSTP1-1 and in the background reaction.



Scheme 9: ISCC screen using **TZ03a-01** and alkyne library. Alkynes resulting in the triazole product being observed in the reaction are highlighted in red. a) **TZ03a-01** (50 μM), ALK (250 μM), hGSTP1-1 (15 μM), 15 % DMSO, 10 mM ammonium acetate buffer, pH 7.0, 7 d, 25 $^{\circ}\text{C}$.

The reactions with **ALK03** and **ALK04** were further investigated with four different GST isoforms to see if amplification of triazole production above the background was possible for other isoforms.

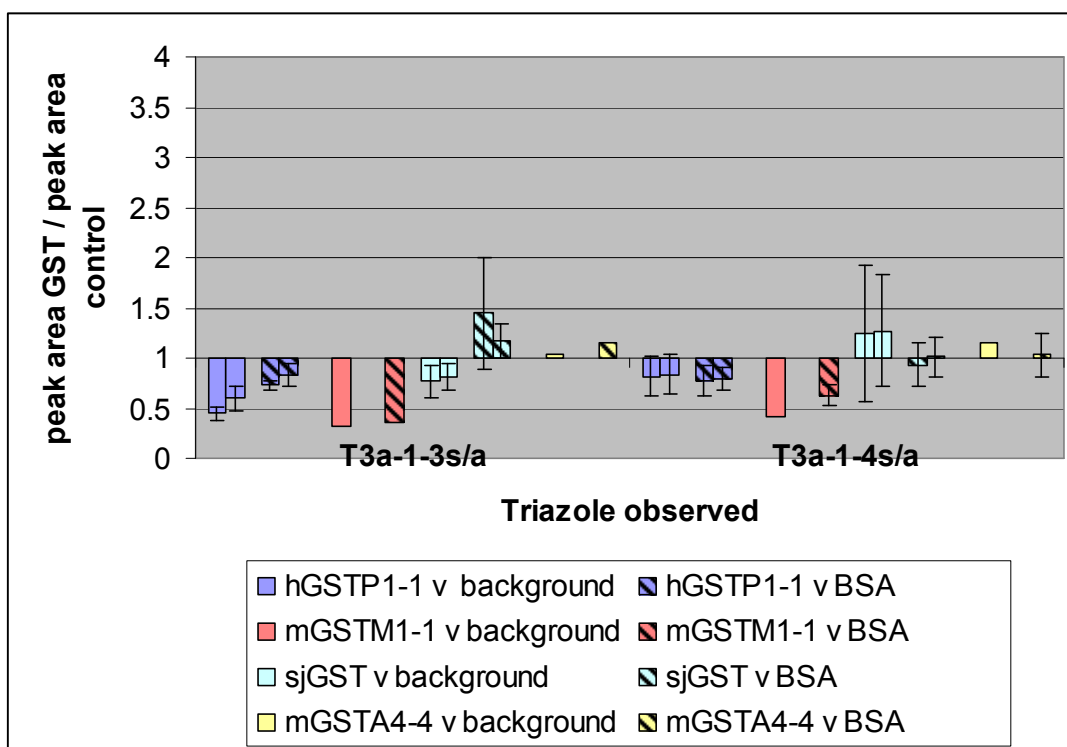


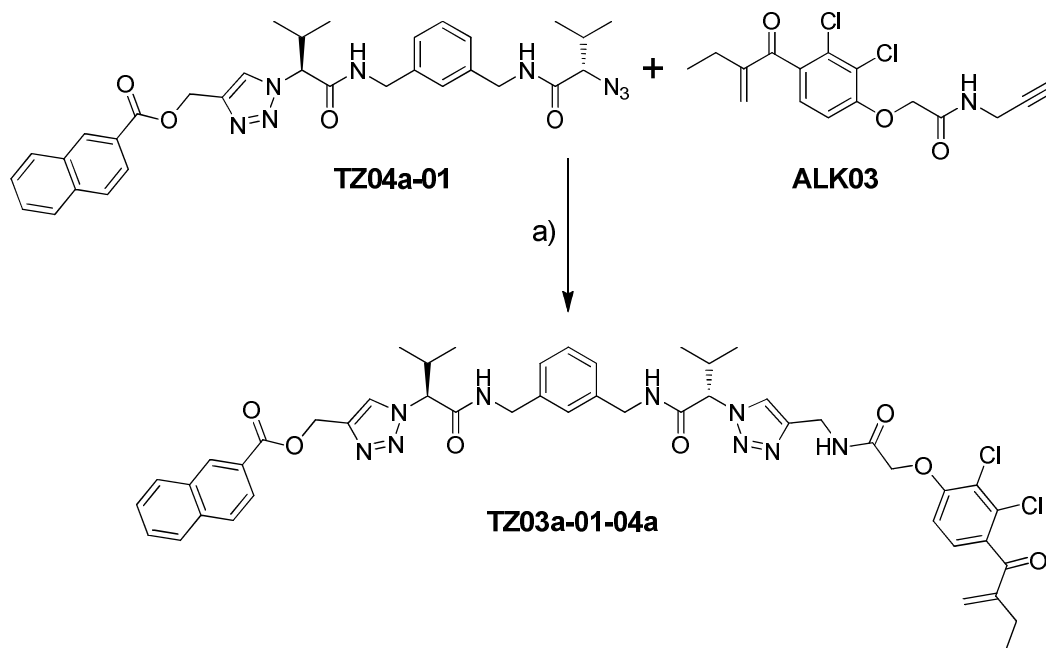
Figure 19: ISCC reaction between **TZ03a-01** and **ALK03** and between **TZ03a-01** and **ALK04** in the presence of various isoforms of GST. Graph shows the ratio of triazole observed in the presence of GST over two controls; the background reaction (block colour) and in the presence of BSA (diagonal stripe). Error bars represent standard deviation. For mGSTM1-1 and mGSTA4-4 only one peak observed. ISCC reaction conditions: Azide (50 μM final conc.), Alkyne (250 μM final conc.), DMSO (15 % by vol.), GST/BSA (15 μM final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μL of reaction quenched in 400 μL methanol, filtered and 10 μL injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for monotriazole and bistriazole products.

Surprisingly, rather than seeing an increase in triazole production in the presence of protein or even a 1:1 ratio indicating that it was only the background reaction being observed, the amount of triazole produced was reduced when compared with the background reaction, Figure 19. The greatest reduction was in the presence of mGSTM1-1 and hGSTP1-1.

The amount of triazole produced was also compared with the reaction in the presence of control protein, BSA. Again, the amount of triazole produced in the

presence of mGSTM1-1 was greatly reduced in comparison with the reaction in the presence of BSA, Figure 19. The other three GST isoforms had a close to 1:1 ratio with BSA.

The bistriazole product **TZ03a-01-03a** had been synthesized previously, Scheme 4. The bistriazole product **TZ03a-01-04a** was synthesized from the copper (I) iodide catalysed reaction between **TZ04a-01** and **ALK03**, Scheme 10.



Scheme 10: Synthesis of **TZ03a-01-04a**. a) 0.2 eq copper (I) iodide, acetonitrile, rt, 18 h.

The inhibitory activity of both 1,4 triazole products, along with the starting azide and alkynes were measured using the CDNB assay to try to elucidate the results observed in the ISCC reaction.

Table 4: A table showing the IC₅₀ values for starting fragments and triazole products observed in the ISCC screen with **TZ03a-01** measured using the CDNB assay. Values are calculated from three separate assays. Raw data described in experimental chapter. (-) = assay not performed.

Compound	hGSTP1-1 IC ₅₀ (μM)	mGSTM1-1 IC ₅₀ (μM)	SjGST IC ₅₀ (μM)	mGSTA1-1 IC ₅₀ (μM)
TZ03a-01	29.64	0.138	7.97	≥ 100
ALK03	27.18	1.101	10.49	≥ 100
TZ03a-01-03a	7.56	0.0038	0.014	≥ 100
ALK04	≥100	≥ 100	-	-
TZ03a-01-04a	24.45	0.055	13.08	≥ 100

The results from the CDNB binding assay identified a very potent inhibitor for both mGSTM1-1 and SjGST in the bivalent triazole **TZ03a-01-03a**, the triazole based on known inhibitor ethacrynic acid (IC₅₀ = 4 nM (mGSTM1-1) and 14 nM (SjGST)). This result, along with that for SjGST, highlights the power of the bivalent approach. In the example of mGSTM1-1, two monovalent inhibitors with IC₅₀ values of 0.1 and 1.0 μM have combined to make a bivalent inhibitor with low nanomolar activity. Even in the case of hGSTP1-1, an improvement in binding affinity was observed with the bivalent triazole.

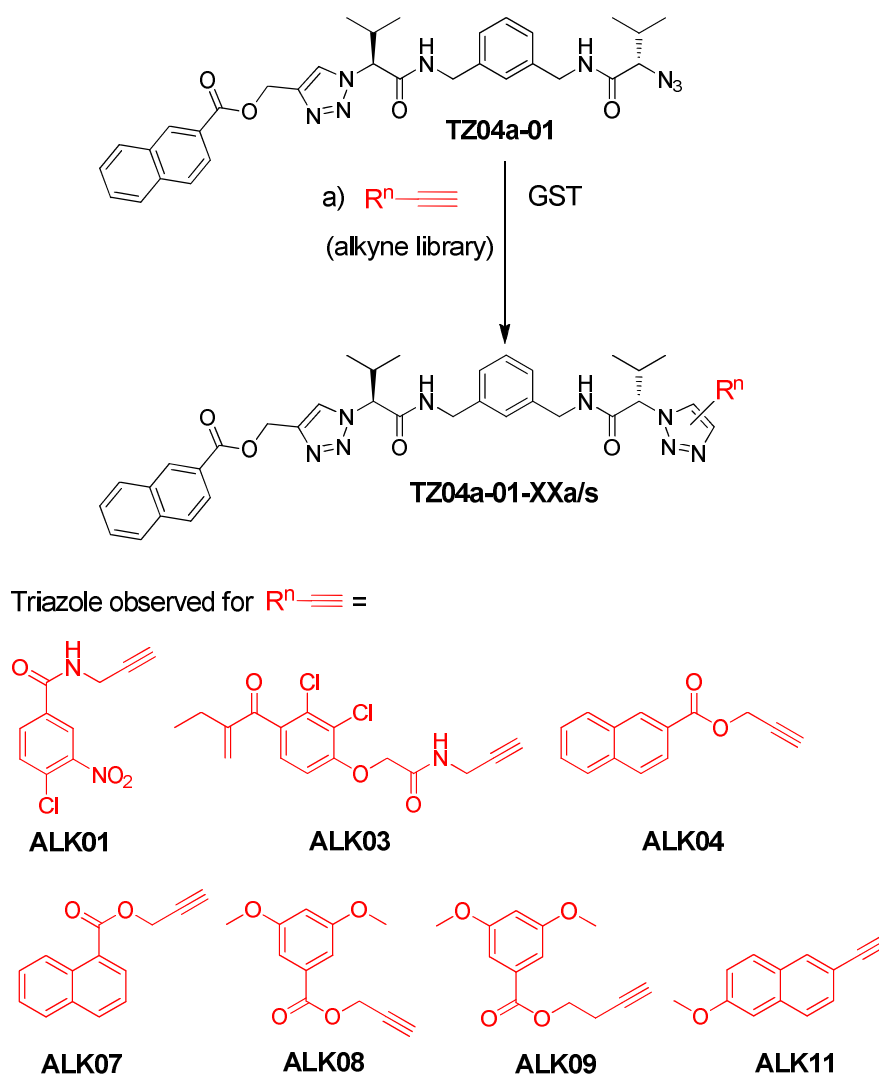
These results suggest that the ISCC reaction could work for **TZ03a-01** with **ALK03**. However, the results from the ISCC reaction were clearly negative with triazole formation actually being reduced relative to the control. This decrease might have been occurring because the azide and alkyne fragments were binding to the GST, reducing the concentration of those fragments participating in the background reaction, but being unable themselves to undergo triazole formation. A noted disadvantage of the ISCC approach is that false negatives can and do occur.^[27] This happens when the fragments bind in the enzyme in such a way that the reactive alkyne and azide groups are held at too great a distance apart and so the reaction cannot proceed, despite the fact that the final triazole product also binds well. This

could explain why the ISCC reaction was not working in this case. Another possible explanation for not seeing any amplification of the tight binding bistriazole product *in situ* was because the template binding pocket was symmetrical but the fragments were asymmetrical. One of the fragments may have been acting as a self inhibitor by binding in both halves of the GST binding pocket, so that azide plus azide or alkyne plus alkyne was present in the binding pocket and therefore no reaction could happen.

While **TZ03a-01-03a** was the most potent ligand, **TZ03a-01-04a** gave the highest GST isoform specificity, with an IC_{50} value 200 times greater for mGSTM1-1 than for any other GST isoform that was tested. The addition of the naphthyl alkyne **ALK04** had little effect on the binding affinity of the monotriazole **TZ03a-01** towards hGSTP1-1 or SjGST. This would suggest that these examples should not work in the ISCC reaction.

3.9 ISCC screen of TZ04a-01 with alkyne library using mGSTM1-1

The ISCC reaction using the monoazide **TZ03a-01** was clearly not working to produce products templated by GST, so a second monoazide **TZ04a-01** was selected to try in the ISCC reaction. **TZ04a-01** had much weaker binding affinity towards GST compared with **TZ03a-01**, which would reduce the possibility of the monoazide acting as a self-inhibitor. **TZ04a-01** was screened with the 20-member alkyne library in the ISCC reaction, under the same conditions as those described previously except for a reduction in the concentration of the azide to 50 μ M due to solubility problems. The screen was performed with mGSTM1-1 as the monoazide displayed a slightly more favourable activity for this isoform and because this isoform had produced the best results in the DCC approach. The resultant LC/MS-SIM traces were compared to the background reaction in the absence of protein, see appendix for traces.



Scheme 11: ISCC screen using **TZ04a-01** and alkyne library. Alkynes resulting in the triazole product being observed in the reaction are highlighted in red. a) **TZ04a-01** (50 μ M), ALK (250 μ M), hGSTP1-1 (15 μ M), 15 % DMSO, 10 mM ammonium acetate buffer, pH 7.0, 7 d, 25 $^{\circ}$ C.

From the 20-member library of alkynes, seven reactions resulted in peaks corresponding to expected triazole products. These products were confirmed by references from both the copper (I)-catalysed or Ru-catalysed reactions and *via* the thermal reaction. Four of the alkynes were the same as those forming triazole with the bisazide **AZ01** (section 3.5.1). These were **ALK03**, **ALK04**, **ALK07** and **ALK08**. The further three alkynes were **ALK01**, **ALK09** and **ALK11**. **ALK01** was based on known substrate CDNB. **ALK09** was the same as one of the original hits, **ALK08**, except for an extra carbon in the alkyne linker. **ALK11** was another fragment containing a naphthyl group. In all of the examples, except for those

containing **ALK08** and **ALK09**, the two triazole regioisomers, the *anti* (1,4) and the *syn* (1,5) triazole, could be separated on the LC/MS. The enzyme-templated reaction produced both triazole regioisomers in every example with the 1,4-triazole being the major isomer. This was unexpected as the confined spatial arrangement of the active site would be expected to strongly favour one isomer over the other.

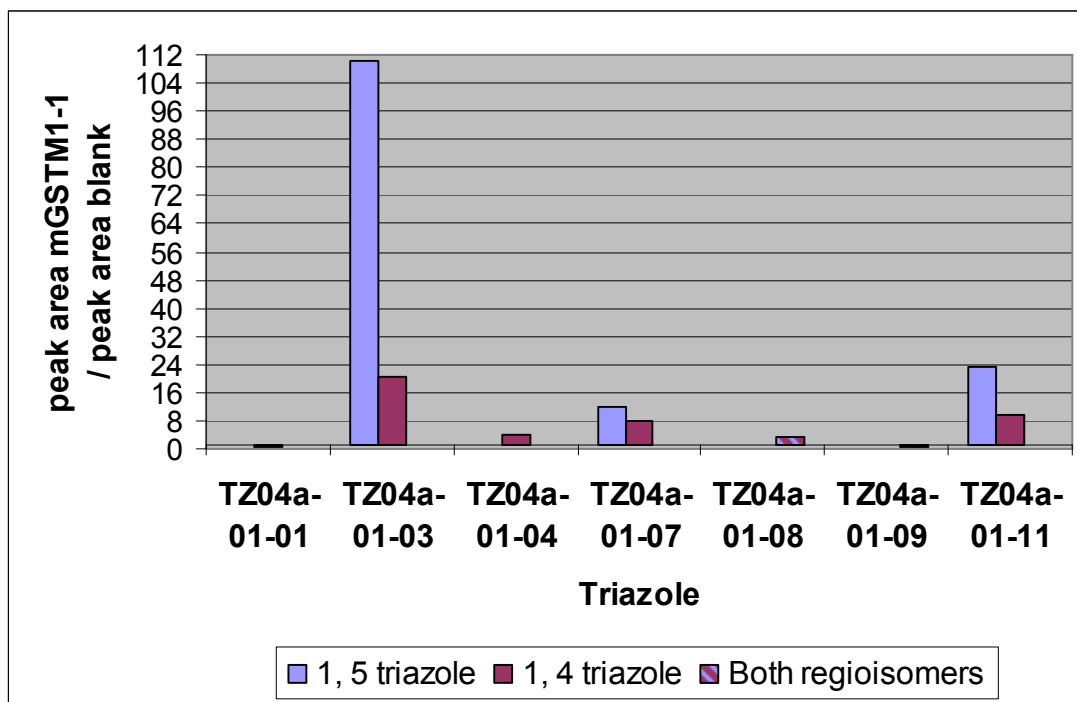


Figure 20: ISCC reaction between **TZ04a-01** and alkyne 'hits' in the presence of mGSTM1-1. Graph shows the ratio of triazole observed in the presence of GST over the background reaction. Data calculated from the ratio between the average value of the peak area for product formed in the presence of mGSTM1-1 and the average value for the peak area for product formed in the background reaction. ISCC reaction conditions: Azide (50 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

The comparison between the enzyme-containing reaction and the background was only undertaken between averages of the peak areas observed for the triazole products and not from an average of the ratios of reactions monitored on the same day. As described in section 3.8, comparison between reactions monitored at

different times resulted in less reliable results, however the important points to take from this experiment, Figure 20, are that the seven triazoles were observed. The alkyne based on the known inhibitor ethacrynic acid, **ALK03**, resulted in a significant increase in triazole formation in the presence of mGSTM1-1. There was a much greater increase in the formation of the 1,4 regioisomer over the 1,5 triazole product. Five of these triazoles would qualify as hits using the guideline 4:1 ratio. **TZ04a-01-01a/s** and **TZ04a-01-09a/s** would not qualify as being hits for mGSTM1-1.

3.9.1 ISCC screen of AZ04a-01 with alkyne 'hits' using mGSTM1-1 – further controls

The seven combinations which produced triazoles were subjected to the same controls described previously in this chapter.

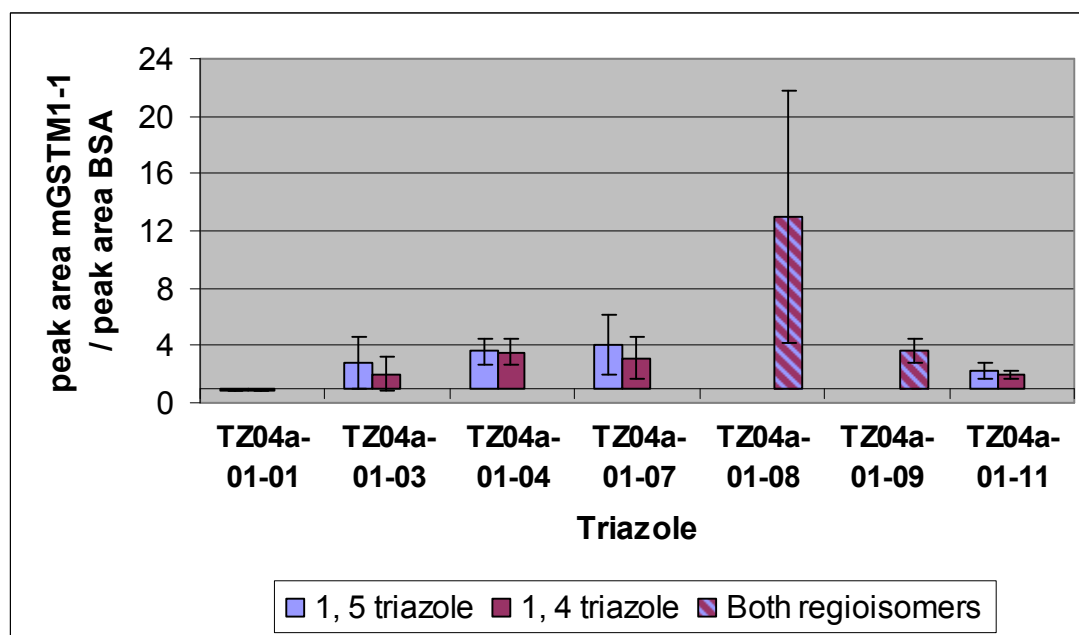


Figure 21: ISCC reaction between **TZ04a-01** and alkyne 'hits' in the presence of mGSTM1-1. Graph shows the ratio of triazole observed in the presence of GST over the reaction in the presence of BSA. Error bars represent standard deviation. ISCC reaction conditions: Azide (50 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST/BSA (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1).

Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

Comparison of the reactions in the presence of mGSTM1-1 with the reactions in the presence of BSA resulted in one reaction forming 12 times the amount of triazole, **TZ04a-01-08a/s**, than in the presence of BSA, Figure 21. Three more triazoles, **TZ04a-01-04a/s**, **TZ04a-01-07s** and **TZ04a-01-09a/s**, were observed at the proposed 4:1 ratio in favour of the mGSTM1-1 reaction. Only one reaction, for the production of **TZ04a-01-01a/s**, resulted in no increase at all in triazole formation between the mGSTM1-1 reaction and the BSA control. This matched the results seen with the background, so this reaction was not studied any further.

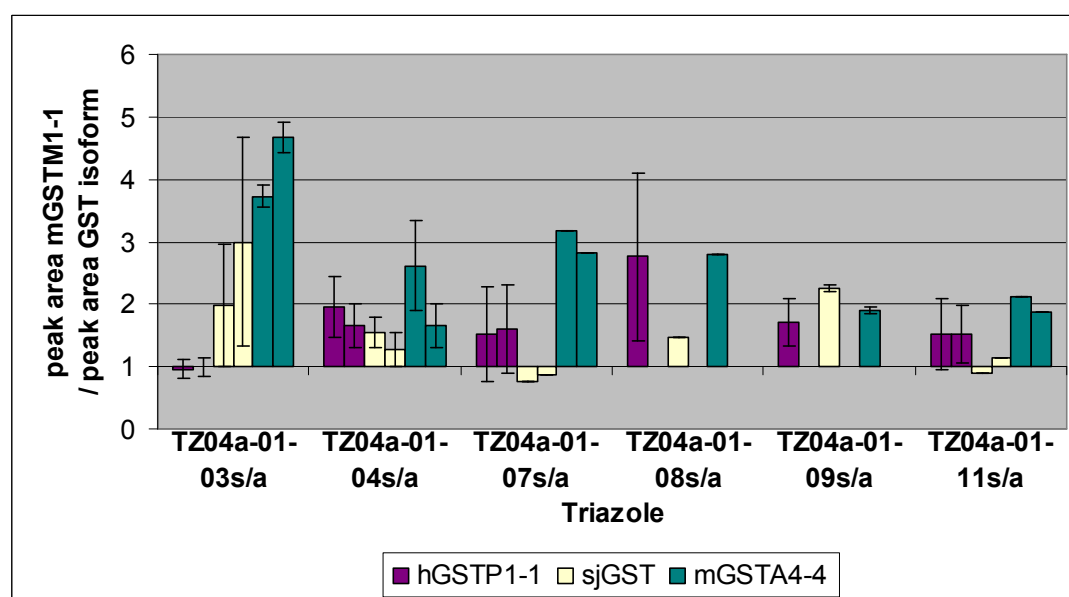


Figure 22: ISCC reaction between **TZ04a-01** and alkyne 'hits' in the presence of mGSTM1-1. Graph shows the ratio of triazole observed in the presence of mGSTM1-1 over the reaction in the presence of other GST isoforms. Error bars represent standard deviation. ISCC reaction conditions: Azide (50 μ M final conc.), Alkyne (250 μ M final conc.), DMSO (15 % by vol.), GST (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

When the reaction in the presence of mGSTM1-1 was compared with other GST isoforms the general trend was that mGSTM1-1 resulted in the greatest amount of triazole formation, Figure 22. The greatest difference between the amounts of triazole observed was between mGSTM1-1 and mGSTA4-4, with typically three times as much triazole forming in the presence of mGSTM1-1. On average twice as much triazole was observed in the reaction with mGSTM1-1 compared with the reaction with hGSTP1-1. The reaction in the presence of SjGST produced triazole in the most similar amounts to mGSTM1-1. These observations with the different GST isoforms follow the same pattern as in the DCLs described in the previous chapter. SjGST is most similar to mGSTM1-1 in its three dimensional structure and this could explain why the amount of triazole observed was very similar for mGSTM1-1 and SjGST.^[28-30]

The final control was to use a known inhibitor of mGSTM1-1 to block the binding pocket. If it was the binding pocket that was templating the click reaction then blocking it should prevent the click reaction. The same approach was used in the previous chapter to prove the amplification in the DCL was being templated by the GST binding pocket. Problems arose with this approach as the hydrophobic binding pocket of GST is so large and not very well defined.^[20, 31-34] As the expected binding orientation of the spacers was not known, a range of GST inhibitors were used as controls.

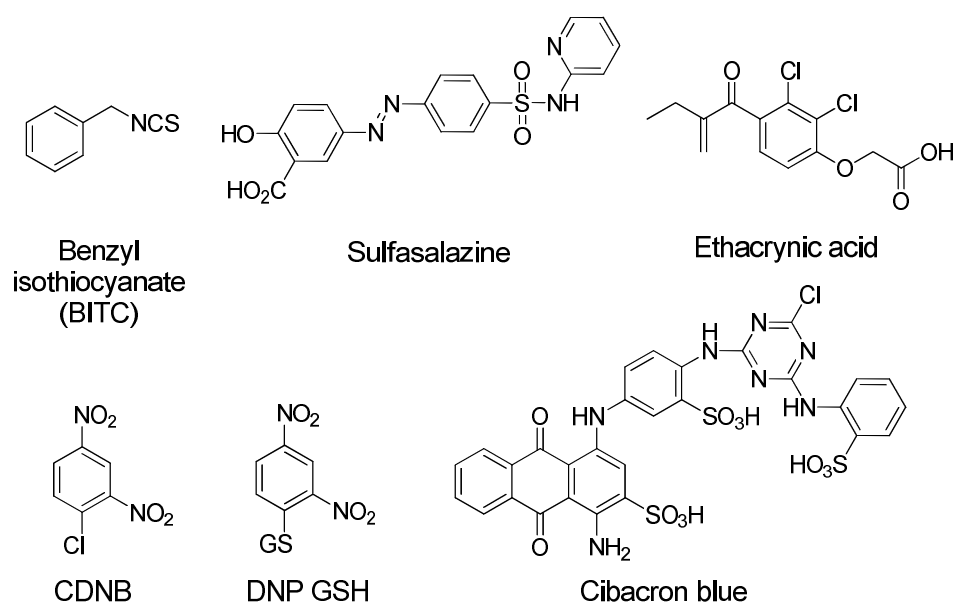


Figure 23: Known inhibitors of GST.

As described in detail in chapter 2.6, BITC, ethacrynic acid and DNP GSH all bind in different orientations within the hydrophobic binding pocket.^[35-38] Sulfasalazine binds in the hydrophobic binding pocket, but it also extends out into a further shallow pocket on the protein surface.^[38] Cibacron blue binds in the ligandin binding site.^[38] In the example of the DCL the only compound that blocked the amplification was DNP GSH. If the triazole compounds are binding in the same orientation as the acylhydrazone compounds described in chapter 2, it would be expected that DNP GSH would also block triazole formation.

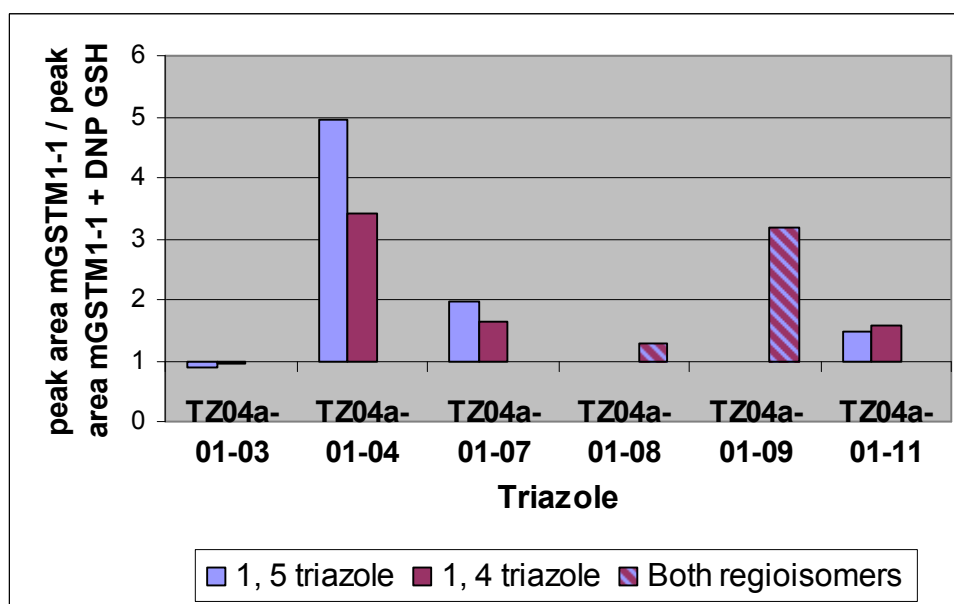


Figure 24: ISCC reaction between **TZ04a-01** and alkyne 'hits' in the presence of mGSTM1-1. Graph shows the ratio of triazole observed in the presence of mGSTM1-1 over the reaction in the presence of mGSTM1-1 plus GST inhibitor, DNP GSH. ISCC reaction conditions: Azide (50 μ M final conc.), Alkyne (250 μ M final conc.), GSH DNP (500 μ M final conc.) DMSO (15 % by vol.), GST (15 μ M final conc.) and ammonium acetate buffer (to total reaction volume of 0.3 mL, 10 mM, pH 7.0) mixed at room temperature. After 7 days, 200 μ L of reaction quenched in 400 μ L methanol, filtered and 10 μ L injected onto LC/MS using gradient B or C depending on the products being measured (see Chapter 5, Section 5.2.2, Table 1). Products detected in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

The full range of GST inhibitors was screened to block both the hydrophobic and ligandin binding pockets within mGSTM1-1. The results showed no effect on the triazole production with the exception of DNP GSH. These results are displayed in

Figure 24. Even with this inhibitor only two reactions, **TZ04a-01-04a/s** and **TZ04a-01-09a/s**, showed an increase in product formation in the presence of the free protein over the protein which was blocked with DNP GSH. Only one of these, **TZ04a-01-04a/s** gave an increase greater than 4:1. This data suggested, perhaps with the exception of **TZ04a-01-04a/s**, that it was not the hydrophobic binding pocket that was the site for the triazole synthesis.

3.10 Summary of the ISCC reaction between TZ04a-01 and alkyne ‘hits’ with mGSTM1-1 and comparison with the binding data for the triazole products

After trying several different controls no clear pattern emerged. The results from all of the control experiments have been summarized in Table 5. The ratio of the mGSTM1-1 value over the control has been listed and those that gave a value of four or above have been marked with a tick to provisionally indicate a ‘hit’ in the reaction.

Table 5: A summary of the results from the ISCC reaction between **TZ04a-01** and the alkyne 'hits', with mGSTM1-1. Data is tabulated by control reactions, with values quoted being the ratio of triazole formed in the presence of mGSTM1-1 over the control. A tick indicates a ratio above four. NT = no triazole observed. - = no reaction performed. The final column shows the IC₅₀ values measured using the CDNB assay. Values are calculated from three separate assays. Raw data described in experimental chapter.

Triazole	Background	BSA	Isoform	Inhibitor DNP GSH	IC ₅₀ μM
AZ04a-01	-	-	-	-	129.3
TZ04a-01-01s	0.8 X	1 X	-	-	-
TZ04a-01-01a	0.8 X	1 X	-	-	7.8
TZ04a-01-03s	110 ✓	3 X	4 (A4) ✓	1 X	-
TZ04a-01-03a	20 ✓	2 X	5 (A4) ✓	1 X	0.055
TZ04a-01-04s	4 ✓	4 ✓	3 (A4) X	5 ✓	16.9
TZ04a-01-04a	4 ✓	4 ✓	X	3 X	13.2
TZ04a-01-07s	12 ✓	4 ✓	3 (A4) X	2 X	3.8
TZ04a-01-07a	8 ✓	3 X	3 (A4) X	2 X	25.5
TZ04a-01-08s	4 ✓	13 ✓	3 (A4) X	1 X	62.1
TZ04a-01-08a	4 ✓	13 ✓	3 (A4) X	1 X	33.9
TZ04a-01-09s	0.4 X	4 ✓	X	3 X	64.9
TZ04a-01-09a	0.4 X	4 ✓	X	3 X	29.6
TZ04a-01-11s	23 ✓	2 X	X	1 X	35.1
TZ04a-01-11a	10 ✓	2 X	X	2 X	9.7
TZ04a-01-10a	NT	NT	NT	NT	38.8
TZ04a-01-18a	NT	NT	NT	NT	30.5

The results from the first two control reactions comparing the GST-templated reaction with the background reaction and the reaction in the presence of BSA, Table 5 column 2 and 3, show whether or not there was a specific feature of the GST enzyme that was promoting the cycloaddition reaction. The reaction between monoazide **Tz04a-01** and alkynes **ALK04**, **ALK07** and **ALK08** gave significant amplifications over both of these control reactions clearly indicating that a distinct region of GST was responsible for accelerating the click reaction. The

GST-templated reaction between monoazide **TZ04a-01** and alkynes **ALK03** and **ALK11** resulted in extremely large amplifications (greater than 10:1) in triazole formation over the background reaction, however a much less substantial increase, lower than the suggested 4:1 ratio, was observed over the BSA reaction.

A degree of improvement was observed in the formation of triazole in the presence of mGSTM1-1 over the other GST isoforms for these reactions, although never above a ratio of 4:1. The comparison between GST isoforms could only serve to indicate if the template-effect was as a result of a specific binding pocket on one isoform over the others. Smaller differences in triazole formation suggest that the binding site acting as the template is present in all of the GST isoforms and the small amplifications are as a result of subtle changes within this binding site between isoforms. The results of using the various GST inhibitors to block the hydrophobic binding pocket and the ligandin site had demonstrated that it was unlikely that it was these sites that were responsible for templating the click reaction.

To gain a better insight into the binding of the triazoles with mGSTM1-1, their inhibitory activity was measured in the CDNB assay. The IC_{50} values have been listed in Table 5. All of the triazoles showed an improvement in the inhibitory activity over the starting azide **AZ04a-01**. The products from the reactions that appeared to be amplified by the presence of GST (**TZ04a-01-03a/s**, **TZ04a-01-04a/s**, **TZ04a-01-07a/s**, **TZ04a-01-08a/s** and **TZ04a-01-11a/s**) displayed activities from 0.055 μ M to 62.1 μ M, a 1000-fold range. This is a large span of results, which appears to contradict the observations of the ISCC reactions. Two control triazoles not observed in the ISCC reaction, **TZ04a-01-10a** and **TZ04a-01-18a**, were also run in the CDNB assay and gave IC_{50} values of 39 and 31 μ M respectively, values that fall within the range observed for the ISCC hits. One explanation is that these two products happen to be false negatives in the ISCC reaction. Alternatively, and perhaps more likely, a three to four-fold improvement in activity for the ditriazole compound compared with the monotriazole precursor signifies a model that is not templated by the enzyme and those products that did show an increase in triazole formation are proof that the ratio used to determine a hit in ISCC needs to be re-evaluated.

While the inhibitory activity of a compound is not a direct measure of its binding affinity for the enzyme, it seems unlikely in this case that some of the compounds bind with good inhibitory activity and other compounds based on the same spacer design bind elsewhere on the enzyme, to exhibit good binding affinity but display little effect on the catalytic turnover. Perhaps the only explanation could be that the spacers bind into the solvent-accessible cleft, as designed, but the terminal groups bind at different angles, so that some bind into an area that can affect the enzymatic activity (possibly the hydrophobic pocket), while others bind in a different area, perhaps similar to the shallow secondary pocket where sulfasalazine is known to bind.^[38] The only way to prove these theories is to firstly determine the binding affinity each ditriazole product exhibits towards GST and compare these values with the data collected from the ISCC reactions. This can be achieved by using a technique such as isothermal titration calorimetry (ITC) or surface plasmon resonance (SPR). Secondly, getting crystal structures of the triazole products bound to the GST would explain how and where on the enzyme the template reaction was occurring. Alternatively, molecular modelling could be used to determine how these ditriazoles are likely to be binding to the GST dimers.

As previously mentioned in chapter 2.1, GSTM1-1 has been shown to bind to apoptosis signal-regulating kinase 1 (ASK1).^[39-41] This mitogen-activated protein kinase kinase kinase plays a crucial role in cytokine- and stress-induced apoptosis in the cell. It is the C-terminal area of GSTM1-1 that binds to ASK1. It could be this site that was responsible for templating the click reaction observed in the ISCC reactions. It has also been shown that the interaction of GSTM1-1 with ASK1 is independent of the glutathione-conjugating activity of GSTM1-1.^[39] This could explain why the bistriazole products that appeared to be amplified in the presence of mGSTM1-1 did not have good inhibition of the glutathione-conjugating activity measured in the CDNB assay.

3.11 Summary of the ISCC reaction between TZ04a-01 and alkyne 'hits' with hGSTP1-1 and comparison with the binding data for the product triazoles

As the original library was designed from bivalent inhibitors of hGSTP1-1, this GST isoform was also used in the ISCC reaction between **TZ04a-01** and the 20-member alkyne. The product formation in the presence of hGSTP1-1 was compared with the control reactions, data in Appendix, Section 6.3. A summary of these reactions has been collated in Table 6. The ratio of the hGSTP1-1 value over the control has been listed and those that gave a value of four or above have been marked with a tick to provisionally indicate a hit in the reaction.

The results obtained correlated well with those results described for the reactions using the mGSTM1-1 isoform. The same five reactions, between **TZ04a-01** and alkynes **ALK03**, **ALK04**, **ALK07**, **ALK08** and **ALK11**, resulted in large amplifications (greater than 4:1 increase) in triazole formation observed over the background and BSA control reactions, again indicating that a specific site on hGSTP1-1 was responsible for accelerating the rate of triazole synthesis. Some amplification was observed over other GST isoforms, typically SjGST and mGSTA4-4, although this was not true in every example. In this case, no decrease in triazole formation was observed with the use of any of the library of known GST inhibitors previously described. This provides evidence that it was not the hydrophobic binding pocket or the ligandin site of hGSTP1-1 that was templating the cycloaddition reaction.

Table 6: A summary of the results from the ISCC reaction between **TZ04a-01** and the alkyne 'hits', with hGSTP1-1. Data is tabulated by control reactions, with values quoted being the ratio of triazole formed in the presence of hGSTP1-1 over the control. A tick indicates a ratio above four. NT = no triazole observed. - = no reaction performed. The final column shows the IC₅₀ values measured using the CDNB assay. Values are calculated from three separate assays. Raw data described in experimental chapter.

Triazole	Background	BSA	Isoforms	Inhibitors	IC ₅₀ μM
AZ04a-01	-	-	-	-	≥ 100
TZ04a-01-01s	7 ✓	1 X	-	X	-
TZ04a-01-01a	7 ✓	2 X	-	X	-
TZ04a-01-03s	62 ✓	2 X	4 (A4) ✓	X	-
TZ04a-01-03a	53 ✓	2 X	5 (A4) ✓	X	24.5
TZ04a-01-04s	4 ✓	3 X	X	X	140.6
TZ04a-01-04a	4 ✓	4 ✓	X	X	166.6
TZ04a-01-07s	7 ✓	9 ✓	7 (A4) ✓	X	56.0
TZ04a-01-07a	6 ✓	5 ✓	5 (A4) ✓	X	55.2
TZ04a-01-08s	7 ✓	8 ✓	5 (Sj) ✓	X	-
TZ04a-01-08a	7 ✓	8 ✓	5 (Sj) ✓	X	-
TZ04a-01-09s	2 X	5 ✓	X	3 (CB)	-
TZ04a-01-09a	2 X	5 ✓	X	3 (CB)	-
TZ04a-01-11s	10 ✓	3 X	X	X	-
TZ04a-01-11a	3 X	2 X	X	X	-
TZ04a-01-10a	NT	NT	NT	NT	36.7
TZ04a-01-18a	NT	NT	NT	NT	2.2

The inhibitory activity of a selection of ligands was again determined and has been detailed in Table 6, column 6. The data showed that overall the ditriazole spacers were weaker inhibitors of hGSTP1-1 compared to mGSTM1-1. The best inhibitors of hGSTP1-1 were actually the two triazole compounds that were not observed in the ISCC reactions. Unless these are examples of false negative results, the triazoles with amplified formation in the presence of hGSTP1-1 must not be templated in a site that affects the enzyme activity, a theory supported by the lack of

an effect on triazole formation by the addition of known GST inhibitors to the GST-templated ISCC reaction. As well as its role in detoxification in the cell, hGSTP1-1 has also been shown to have antiapoptotic activity *via* a protein-protein interaction with c-*Jun* NH₂-terminal kinase (JNK), which is a key regulator of apoptosis within the cell.^[42-45] GSTP1-1 is known to bind to JNK in the C-terminal region of the protein. It is the same region that contributes to the shaping of the hydrophobic binding pocket. It could be that it was the site of the protein-protein interaction that was templating the click reaction in the ISCC reactions.

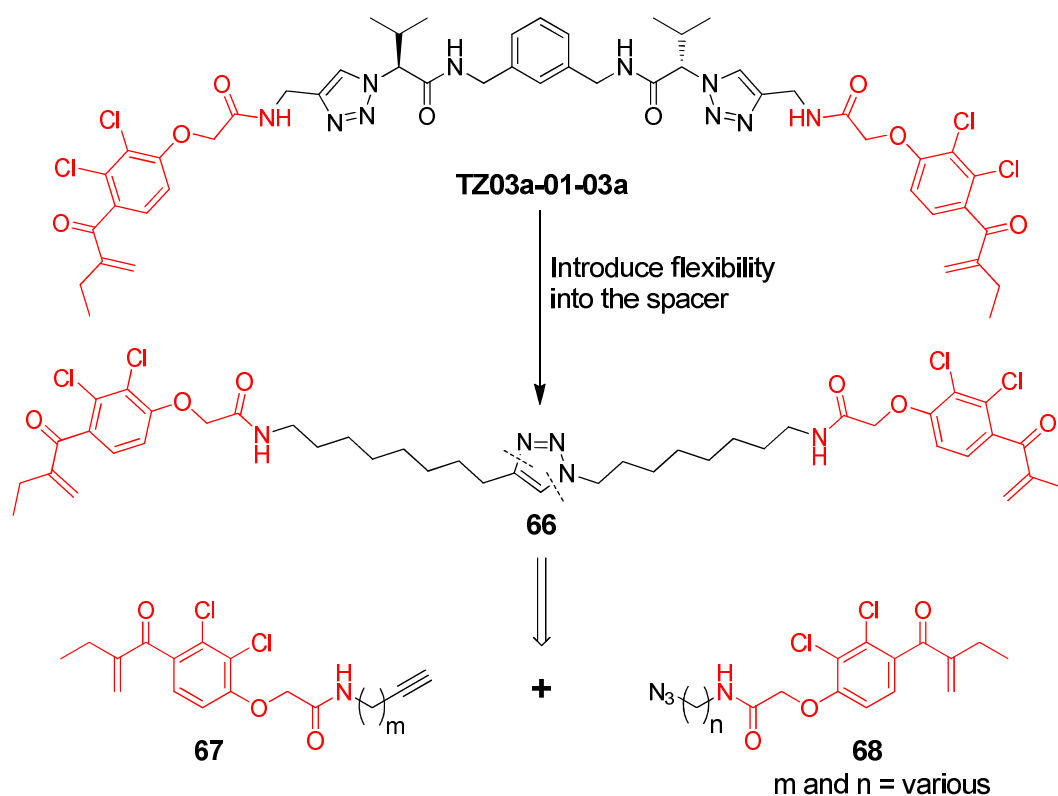
3.12 Implications for ISCC as an approach for TGS

ISCC screens based on three different azide scaffolds have been studied in this chapter and a number of weaknesses have been exposed. The use of a diazide spacer resulted in the observation of monotriazole products in four examples. The average amplification of triazole formation in the presence of GST over the controls was a 4:1 ratio or less. The inhibitory activity of these monotriazoles suggested that the small amplifications observed in the ISCC reaction were not as a result of a specific template effect of the target protein, demonstrating that a minimum of a 4:1 increase in triazole is desired to define an *in situ* hit.

The optimization reactions highlighted a number of inconsistencies with the observed amount of triazole formed in a reaction. The presence of a background reaction made this problematic in comparing the relative amounts of triazole forming in each reaction. As discussed, more recent examples of ISCC in the literature have reported the observation of the background click reaction.^[22, 24] As with this work, these were examples in which relatively weak-binding fragments were utilised in the ISCC library and so higher concentrations were required in the ISCC reaction, which seems to be the cause of the observed background reaction. If ISCC is going to be used as a valid and consistent tool in drug discovery, then it is reasonable to expect that weakly-binding fragments will be employed regularly, so suitable guidelines to define a 'hit' need to be determined. Clearly if the quantification of an *in situ* hit is going to be based on an amplification of product formation over a background reaction then the calculation of the amount of product observed needs to be both accurate and reliable. The results observed in this work suggested the ionisation of

compounds in the LC/MS was unreliable over long time periods. While reactions run on the same day could be reasonably compared, reactions run weeks apart could not be. It would be recommended to introduce an internal standard to the sample prior to monitoring on the LC/MS, which would then allow the concentration of triazole to be calculated from the peak areas observed.

The second azide scaffold to be studied, the monoazide based on the original diazide spacer with ethacrynic acid attached at one end, highlighted another drawback of ISCC, the problem of false negatives. For the reaction between the monoazide and the alkyne-derivatised ethacrynic acid, a 100-fold increase in the IC₅₀ value of the product over the monoazide was observed, yet the triazole formation observed in the GST-templated reaction was reduced compared with the background reaction. These observations could imply that a degree of flexibility is essential between the fragments and the reactive functional groups. The central spacer unit had a peptide backbone which could reduce the flexibility of movement of the reactive azide and alkyne functional groups. Future work could involve redesigning the fragment libraries to introduce more flexibility, e.g. by removing the peptidic spacer and/or changing from two triazole connections to one **66**, as shown in Scheme 12. The azide and alkyne functionality could then be linked to the ligand, which is designed to dock in the hydrophobic binding pocket, by a more flexible alkyl or polyethylene glycol (PEG) chain **67** and **68**. This approach, however, is vulnerable to problems with self-inhibition due to the symmetrical binding pockets of the GST as the template, but may give some indication of the amount of flexibility required by fragments used in ISCC.



Scheme 12: A scheme to outline the potential for adding more flexibility into the spacer design to aid the ISCC approach.

The final azide scaffold analysed, the monoazide with a terminal naphthyl group, identified several products that were showing large amplifications in triazole formation as a result of a specific template effect from the GST proteins. The use of known GST inhibitors and the lack of correlation between the inhibitory activity of those triazoles amplified by the GST suggested that it was not the hydrophobic binding pocket that was acting as the template, despite the original library being designed to span the two H-sites located in the GST dimer. As discussed previously, it is now important to get a direct measure of the binding affinity of these triazoles with the GST proteins. Even without the binding affinity the results of the inhibitory activity reveal another potential limitation to ISCC as a drug discovery technique. This ISCC library was designed, based on limited structural knowledge, to identify novel GST inhibitors. The results of the final ISCC screen indicated several hits and while these may still demonstrate good binding affinity for GST, some of them did not display good inhibition of GST. One of the advantages of ISCC over traditional drug discovery approaches is that it should not require much, if any structural

information about the biological target. The results from this ISCC screen clearly demonstrate the importance of the control reaction in which the target binding site is blocked with a known inhibitor, to reveal the exact location within the biological target that is templating the click reaction. The large, flexible hydrophobic binding site of GST has made this control reaction very difficult in this work. The ability to perform this control reaction is an important consideration in the future selection of biological targets to investigate with an ISCC approach.

3.13 Inconsistencies in the ISCC reaction

An unexpected observation in the ISCC reaction was the formation of both triazole regioisomers. The restricted three dimensional structure of the binding pocket of the biological template should direct the synthesis of only one regioisomer. Indeed in all previous literature reports of ISCC, a single triazole regioisomer is observed. In an article by Omura and co-workers, using ISCC to identify novel chitinase inhibitors, the second regioisomer was observed in the background reaction; however this product showed no amplification in the presence of the biological target.^[24] In the ISCC reactions described in this thesis, both regioisomers are observed in the background and are shown to be amplified in the presence of GST. While this is a novel observation in ISCC, it is important to note that inhibitory activity of both regioisomers was similar, with typically only a 2-fold difference in inhibition, suggesting the GST protein is tolerant of both triazoles. Again, obtaining an image of how the ligands are binding to GST, either by molecular modelling or X-ray crystallography, will explain this observation.

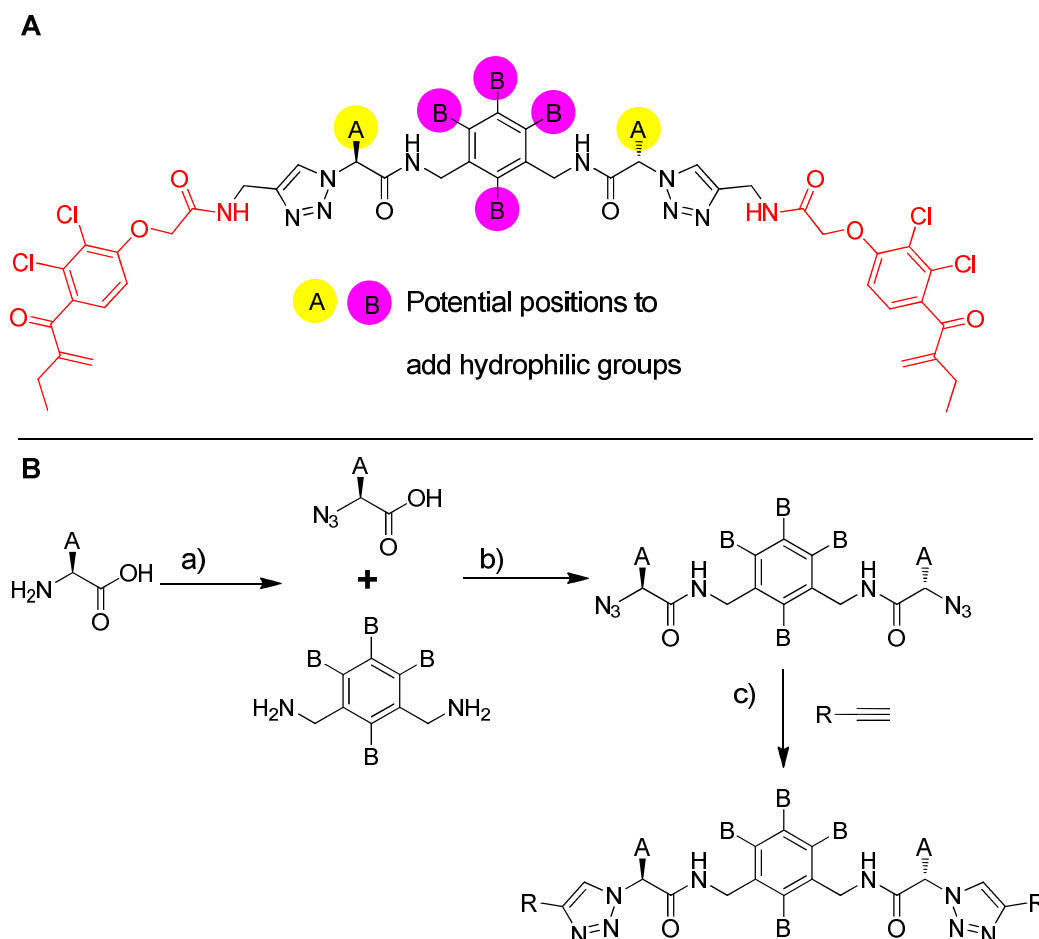
A second point to consider from the ISCC screens, is that the same four alkynes (**ALK03**, **ALK04**, **ALK07** and **ALK08**) are always observed to form triazole, with a further three related alkynes (**ALK01**, **ALK09** and **ALK11**) producing triazole in the final ISCC screen with the naphthyl monoazide **TZ04a-01**. The remaining 13 alkynes do not result in triazole formation even in the background reaction. This might suggest a source of metal catalyst is present in these alkyne samples which is responsible for the click reaction. The metal trace can only be present in the alkyne samples or triazole would be observed for every reaction. The observation of both triazole regioisomers means that a trace amount of both copper

and ruthenium must be present in these samples if they are the cause of the click reaction. While this seems unlikely, the alkyne samples could be analysed for trace copper and ruthenium.

3.14 Implications for multivalent drug design

The ditriazole products were poorly water soluble, which made them unsuitable for assays to measure their binding affinity for GST. The inhibitory activity of the triazoles could only be determined by using 15 % DMSO to solubilise the ligands in the CDNB activity assay. The hydrophobic nature of these ligands highlights a drawback in the multivalent approach to drug discovery, which becomes limited by the size and hydrophobicity of the molecules that are designed. The library fragments also had limited solubility in the reaction buffer, which resulted in the reactions being performed at almost saturated conditions to try to obtain the maximum amount of triazole being produced in the protein-templated reaction. It would perhaps help the ISCC reaction if the water solubility of the reacting fragments was increased.

The best area for the introduction of some hydrophilic groups would appear to be the peptidic spacer as this is binding in the solvent-accessible cleft, and might be expected to be more tolerant to the introduction of added functionality. The two areas that could be modified are the side chain of the peptide backbone, marked A in Scheme 13. Alternatively hydrophilic groups could be introduced onto the central phenyl ring of the spacer, marked B in Scheme 13. These modifications could be investigated using the ISCC reaction, to see if the increased solubility results in clearer results from the reaction. The identification of the novel bivalent inhibitors **TZ03a-01-03a** and **TZ03a-01-04a**, combined with the lack of success with the ISCC reaction for the GST target so far, may warrant a more traditional approach by synthesis of a combinatorial library in the lab, followed by biological testing. The synthesis of analogues of the symmetrical inhibitor **TZ03a-01-03a** could be achieved in three simple steps using reliable chemistry already demonstrated in this chapter.



Scheme 13: A scheme to outline the potential for adding more water soluble groups to bivalent product **TZ03a-01-03a**. **A)** Sites A and B highlight the positions on **TZ03a-01-03a** that could be modified to add hydrophilic groups, **B)** A scheme to outline a simple three step procedure to synthesize analogues of **TZ03a-01-03a**. a) diazotransfer step, b) amide coupling step, c) Cu (I) catalysed click reaction step.

3.15 Conclusions and future work

Two novel bivalent inhibitors of GST with low nanomolar inhibitory activity have been identified in this work.

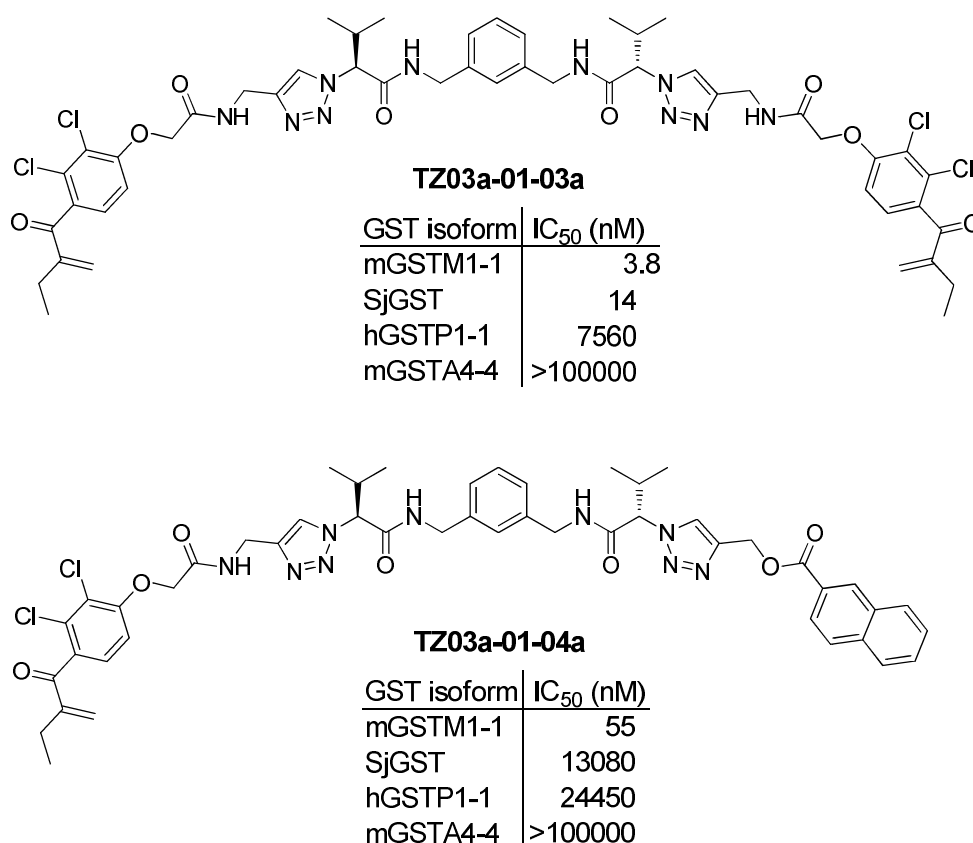


Figure 25: Two novel bivalent GST inhibitors **TZ03a-01-03a** and **TZ03a-01-04a** and their IC₅₀ binding data for four GST isoforms determined using the CDNB binding assay.

TZ03a-01-03a is a bivalent compound linking together two ethacrynic acid moieties with a triazolo-peptide spacer. The compound had excellent binding affinity for both mGSTM1-1 and SjGST, with IC₅₀ values of 4 nM and 14 nM respectively. This is the most potent GSTM1-1 inhibitor reported to date (previous best is cibacron blue, IC₅₀ = 700 nM for mGSTM1-1).^[20] **TZ03a-01-04a** linked one ethacrynic acid group with a naphthyl group utilizing the same triazolo-peptide spacer. This compound was highly isoform selective for mGSTM1-1 over three other GST isoforms tested, hGSTP1-1, mGSTA4-4 and SjGST, with an IC₅₀ value 200 times greater for mGSTM1-1 than for any other isoform. This is unprecedented selectivity for GSTM1-1, with the previous best selectivity exhibited being a 3-fold increase in potency for a glutathione analogue over GSTP1-1 and GSTA1-1.^[46] Whilst these ditriazole compounds were discovered as part of an ISCC approach, the selection / non-selection of triazoles in the *in situ* reaction did not correlate well to inhibitory activity.

As has already been discussed, the immediate work that is required is the measurement of the binding affinity of the observed ditriazoles for GST, either by SPR or ITC. Unfortunately the poor water solubility of these ligands makes this very difficult to achieve and so the hydrophilicity of these triazoles needs to be improved. Obtaining a crystal structure of a ditriazole product bound to GST or alternatively the use of molecular modelling is also important to understand where the spacers are binding on the protein. The reactions in which triazoles are observed need to be analysed to ensure there is no trace of copper (I) or ruthenium present, which might promote triazole synthesis.

In summary, a new ISCC screen has been developed to investigate novel bivalent inhibitors of GST. The results from the ISCC reactions have highlighted some potential problems with the ISCC approach to target-guided synthesis, including difficulties when the background reaction is observed, the possibility of false negatives and the need for flexibility of the reactive functional groups in the fragment library design. The results from the ISCC reaction suggested that several examples were being specifically amplified in the presence of one or more isoforms of GST over the background and the control protein reactions. The evidence indicated that the hydrophobic pocket of GST, which was the expected site of binding, was not the template for the increase in triazole formation. Work now needs to be done to identify the binding site on GST for the bistriazole compounds. In the course of the project two of the bistriazole compounds synthesized, **TZ03a-01-03a** and **TZ03a-01-04a**, were found to be potent inhibitors of mGSTM1-1 and in the latter case, **TZ03a-01-04a**, highly isoform specific.

3.16 References

- [1] A. V. Eliseev, *Drug Discovery Today* **2004**, *9*, 348.
- [2] G. E. Boldt, T. J. Dickerson, K. D. Janda, *Drug Discovery Today* **2006**, *11*, 143.
- [3] X. Hu, R. Manetsch, *Chem. Soc. Rev.* **2010**, *39*, 1316.
- [4] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 2004.
- [5] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057.

- [6] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2596.
- [7] H. C. Kolb, K. B. Sharpless, *Drug Discovery Today* **2003**, *8*, 1128.
- [8] J. E. Moses, A. D. Moorhouse, *Chem. Soc. Rev.* **2007**, *36*, 1249.
- [9] G. C. Tron, T. Pirali, R. A. Billington, P. L. Canonico, G. Sorba, A. A. Genazzani, *Med. Res. Rev.* **2008**, *28*, 278.
- [10] Guest Editors, M. G. Finn, V. Fokin, *Chem. Soc. Rev.* **2010**, *39*, 1221.
- [11] W. G. Lewis, L. G. Green, F. Grynszpan, Z. Radic, P. R. Carlier, P. Taylor, M. G. Finn, K. B. Sharpless, *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 1053.
- [12] R. P. Lyon, J. J. Hill, W. M. Atkins, *Biochemistry* **2003**, *42*, 10418.
- [13] D. Y. Maeda, S. S. Mahajan, W. M. Atkins, J. A. Zebala, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3780.
- [14] S. S. Mahajan, L. Hou, C. Doneanu, R. Paranj, D. Maeda, J. Zebala, W. M. Atkins, *J. Am. Chem. Soc.* **2006**, *128*, 8615.
- [15] A. Brik, J. Alexandratos, Y.-C. Lin, J. H. Elder, A. J. Olson, A. Wlodawer, D. S. Goodsell, C.-H. Wong, *ChemBioChem* **2005**, *6*, 1167.
- [16] V. D. Bock, R. Perciaccante, T. P. Jansen, H. Hiemstra, J. H. van Maarseveen, *Org. Lett.* **2006**, *8*, 919.
- [17] V. D. Bock, D. Speijer, H. Hiemstra, J. H. van Maarseveen, *Org. Biomol. Chem.* **2007**, *5*, 971.
- [18] E. D. Goddard-Borger, R. V. Stick, *Org. Lett.* **2007**, *9*, 3797.
- [19] W. H. Habig, M. J. Pabst, W. B. Jakoby, *J. Biol. Chem.* **1974**, *249*, 7130.
- [20] B. Mannervik, U. Helena Danielson, *Crit. Rev. Biochem. Mol. Biol.* **1988**, *23*, 283.
- [21] T. Shimada, *Drug Metabolism and Pharmacokinetics* **2006**, *21*, 257.
- [22] M. Whiting, J. Muldoon, Y. C. Lin, S. M. Silverman, W. Lindstrom, A. J. Olson, H. C. Kolb, M. G. Finn, K. B. Sharpless, J. H. Elder, V. V. Fokin, *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 1435.
- [23] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat Chem* **2010**, *2*, 490.

- [24] T. Hirose, T. Sunazuka, A. Sugawara, A. Endo, K. Iguchi, T. Yamamoto, H. Ui, K. Shiomi, T. Watanabe, K. B. Sharpless, S. Omura, *J. Antibiot.* **2009**, *62*, 277.
- [25] X. Hu, J. Sun, H.-G. Wang, R. Manetsch, *J. Am. Chem. Soc.* **2008**, *130*, 13820.
- [26] L. Zhang, X. G. Chen, P. Xue, H. H. Y. Sun, I. D. Williams, K. B. Sharpless, V. V. Fokin, G. C. Jia, *J. Am. Chem. Soc.* **2005**, *127*, 15998.
- [27] V. P. Mocharla, B. Colasson, L. V. Lee, S. Roper, K. B. Sharpless, C. H. Wong, H. C. Kolb, *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 116.
- [28] K. Lim, J. X. Ho, K. Keeling, G. L. Gilliland, X. Ji, F. Rüker, D. C. Carter, *Protein Science* **1994**, *3*, 2233.
- [29] M. A. McTigue, D. R. Williams, J. A. Tainer, *J. Mol. Biol.* **1995**, *246*, 21.
- [30] R. M. F. Cardoso, D. S. Daniels, C. M. Bruns, J. A. Tainer, *Proteins: Structure, Function, and Bioinformatics* **2003**, *51*, 137.
- [31] T. H. Rushmore, C. B. Pickett, *J. Biol. Chem.* **1993**, *268*, 11475.
- [32] R. N. Armstrong, *Chem. Res. Toxicol.* **1997**, *10*, 2.
- [33] D. Sheehan, G. Meade, V. M. Foley, C. A. Dowd, *Biochem. J.* **2001**, *360*, 1.
- [34] J. D. Hayes, J. U. Flanagan, I. R. Jowsey, *Annual Review of Pharmacology & Toxicology* **2005**, *45*, 51.
- [35] L. A. Ralat, R. F. Colman, *J. Biol. Chem.* **2004**, *279*, 50204.
- [36] A. J. Oakley, M. Lo Bello, A. P. Mazzetti, G. Federici, M. W. Parker, *FEBS Lett.* **1997**, *419*, 32.
- [37] A. J. Oakley, J. Rossjohn, M. Lo Bello, A. M. Caccuri, G. Federici, M. W. Parker, *Biochemistry* **1997**, *36*, 576.
- [38] A. J. Oakley, M. Lo Bello, M. Nuccetelli, A. P. Mazzetti, M. W. Parker, *J. Mol. Biol.* **1999**, *291*, 913.
- [39] S.-G. Cho, Y. H. Lee, H.-S. Park, K. Ryoo, K. W. Kang, J. Park, S.-J. Eom, M. J. Kim, T.-S. Chang, S.-Y. Choi, J. Shim, Y. Kim, M.-S. Dong, M.-J. Lee, S. G. Kim, H. Ichijo, E.-J. Choi, *J. Biol. Chem.* **2001**, *276*, 12749.
- [40] S. Dorion, H. Lambert, J. Landry, *J. Biol. Chem.* **2002**, *277*, 30792.
- [41] K. Ryoo, S.-H. Huh, Y. H. Lee, K. W. Yoon, S.-G. Cho, E.-J. Choi, *J. Biol. Chem.* **2004**, *279*, 43589.

- [42] D. M. Townsend, K. D. Tew, *Oncogene* **2003**, *22*, 7369.
- [43] C. C. McIlwain, D. M. Townsend, K. D. Tew, *Oncogene* **2006**, *25*, 1639.
- [44] P. Ruzza, A. Rosato, C. R. Rossi, M. Floreani, L. Quintieri, *Anti-Cancer Agents Med. Chem.* **2009**, *9*, 763.
- [45] E. Laborde, *Cell Death Differ* **2010**, *17*, 1373.
- [46] T. Kunze, S. Heps, *Biochem. Pharmacol.* **2000**, *59*, 973.

Chapter 4 Conclusions

Three novel, potent, bivalent GST inhibitors have been reported in this thesis. The linkers were based on a design first described by Atkins and co-workers, which connected two substrates that demonstrated binding affinity for the hydrophobic binding pocket of GST with a central linker unit, as outlined in Figure 1.

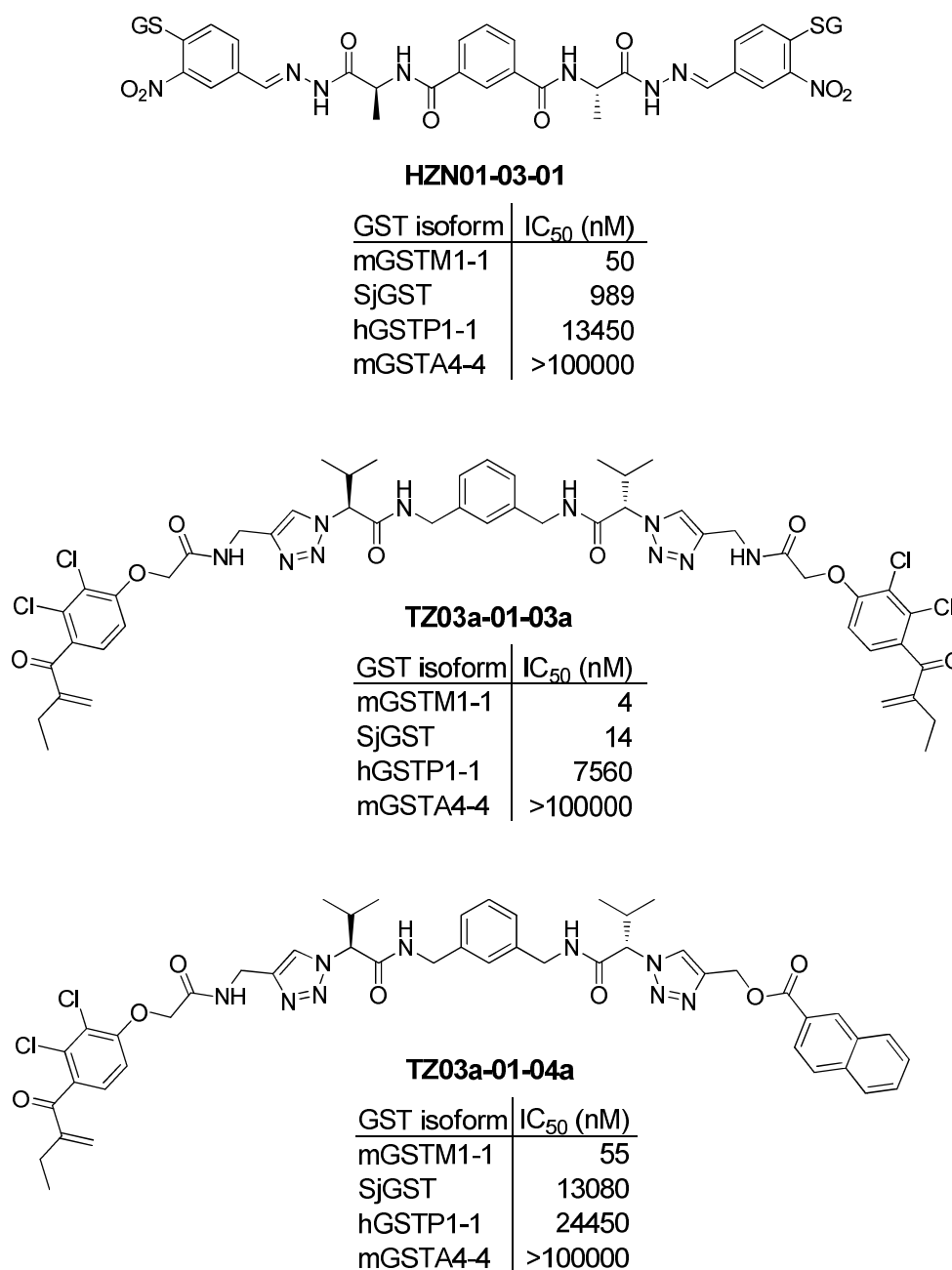


Figure 1: Three novel bivalent GST inhibitors **HZN01-03-01**, **TZ03a-01-03a** and **TZ03a-01-04a** and their IC₅₀ binding data for four GST isoforms.

The first potent inhibitor described in this thesis linked two molecules of DNP GSH using an acylhydrazone linkage. The best inhibitor **HZN01-03-01** exhibited an IC_{50} value of 50 nM towards mGSTM1-1, which was a 20-fold improvement in inhibition over any other GST isoform tested. The other two inhibitors were designed based on a triazole connection; a symmetrical ligand **TZ03a-01-03a** connecting two molecules of ethacrynic acid and an asymmetrical ligand **TZ03a-01-04a** connecting a molecule of ethacrynic acid with a naphthalene species. The symmetrical ligand **TZ03a-01-03a** is the most potent inhibitor of GSTM1-1 reported to date with an IC_{50} of 5 nM. The asymmetrical ligand **TZ03a-01-04a** ($IC_{50} = 55$ nM for GSTM1-1) offers excellent isoform specific with a 200-fold improvement in inhibitory activity for GSTM1-1 over SjGST, hGSTP1-1 and mGSTA4-4.

The acylhydrazone ligand **HZN01-03-01** was discovered using a DCC approach, employing aniline-catalysed acylhydrazone formation as the reversible reaction. Acylhydrazone linker **HZN01-03-01** was selectively amplified in the presence of a GST from a library composing of four bis hydrazide linkers and three aldehyde fragments. The amplifications observed in the DCL correlated well with the binding data for the product compounds.

A second TGS approach has also been established using *in situ* click chemistry (ISCC) to try to identify novel GST inhibitors. This approach proved to be much more complicated than using DCC. In a direct comparison with the DCC approach, attempts to synthesise the symmetrical ditriazole ligand **TZ03a-01-03a** from a bis azide linker **AZ01** and two alkyne fragments in the presence of the GST template proved to be unsuccessful. There was evidence of the formation of the monotriazole product **TZ03a-01** but the observation of the background reaction in the absence of GST made the interpretation of this data difficult. Some amplification in monotriazole product was observed in the presence of GST for several alkynes, however these amplifications did not correlate with the inhibitory activity of the monotriazole products. Attempts to synthesise **TZ03a-01-03a** from the monoazide **TZ03a-01** in the presence of GST were also unsuccessful, with no increase in product formation observed in the presence of GST over the background. The inhibitory activity of the ditriazole product showed a significant improvement over

the starting azide and alkyne fragments (35-fold improvement over azide) suggesting that an *in situ* hit should be observed. The failure to template the click reaction in this example may be because a symmetrical template was being used with asymmetric reactive fragments. Replacing the ethacrynic acid-tagged azide with a less potent azide **TZ04a-01** resulted in the formation of several triazole products being amplified in the presence of GST, but again the inhibitory activity of these compounds did not correspond with the observed amplifications. This result suggests that a different binding site on the GST could be responsible for templating the click reaction and work is now being done to identify where these ligands could be binding to GST. The primary goal of identifying novel inhibitors through the use of ISCC has failed. The binding affinity of these ligands for GST needs to be calculated to find out if, as the ISCC data suggests, the ligands are binding to a different site, not associated with the enzyme's activity.

This work has identified several drawbacks to the use of ISCC as a tool in drug design, which need to be addressed. The major problem that was encountered during the ISCC screens was the observation of the background reaction. The ISCC approach was originally hailed a successful reaction for kinetic TGS as no background reaction was observed, however as ISCC has been applied to more ambitious targets, so more weakly binding fragments have been employed. Because of the weak binding affinity of the azide and alkyne fragments, much greater concentrations are required in the ISCC reaction and this seems to be the cause for observing the background reaction. Clearly for ISCC to be a useful technique the use of weakly binding fragments will be a necessity and therefore a clear set of rules defining how much amplification in triazole formation results in a hit need to be established. To avoid the problems associated with defining a hit, it might be advisable to pursue the use of ISCC with a solid phase approach; that is to immobilise the protein template on a solid support which would allow any triazole products formed in the template to be easily separated from the reaction mixture and identified, thus negating the problems observed with comparisons to the background reaction.

One of the aims of this thesis was to compare the thermodynamic TGS approach with the kinetic TGS approach using a difficult protein target whilst also

employing a multicomponent reaction. With respect to this DCC has clearly proved to be a much more suitable approach, identifying novel inhibitors from a three component reaction. The confusing results observed in the ISCC reactions indicate that there are limitations to the biological targets that can be used with ISCC.

The advantages of a multivalent approach to drug discovery have clearly been demonstrated within this thesis. Linking two molecules of the substrate product DNP GSH to synthesise **HZN01-03-01** resulted in a 6800-fold increase in inhibition of mGSTM1-1 (IC₅₀ of DNP GSH-aldehyde **ALD01** = 340 μM, IC₅₀ of **HZN01-03-01** = 50 nM). The potency of ethacrynic acid for mGSTM1-1 was improved 250-fold by linking two molecules together to yield **TZ03a-01-03a** (IC₅₀ of alkyne-functionalised ethacrynic acid **ALK03** = 1.1 μM, IC₅₀ of **TZ03a-01-03a** = 4 nM).

Atkins utilised the idea of multivalency to overcome the challenge of designing isoform specific inhibitors of GST and demonstrated this in the design of specific inhibitors of GSTA1-1 and GSTP1-1. This thesis has now reported specific inhibitors of GSTM1-1, a target in anti-cancer therapy, and demonstrated the ease with which a DCC approach can be used to identify these inhibitors based on a similar motif to the Atkins inhibitors (ligand-linker-ligand). This design can easily be exploited by varying both the ligands used and the length of the linkers to identify more new isoform specific GST inhibitors.

Chapter 5 Experimental Data

5.1 Product synthesis and characterisation

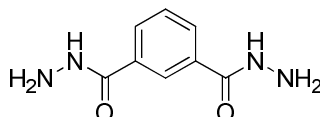
5.1.1 General methods

^1H and ^{13}C NMR spectra were recorded on a Bruker AVA500 (500 MHz) instrument. All NMR spectra were referenced internally to residual protio solvent (^1H) or solvent (^{13}C) resonances and are reported relative to tetramethylsilane ($\delta = 0$ ppm). Chemical shifts are quoted in δ (ppm) and coupling constants in Hertz. Electrospray high resolution mass spectrometry was performed by the EPSRC National Mass Spectrometry Service Centre, Swansea, using a Thermofisher LTQ Orbitrap XL mass spectrometer. The data is recorded as the ionisation method followed by the calculated and measured masses. Liquid chromatography low resolution mass spectrometry was performed on an Agilent 1200 series LC-MS and recorded as the calculated and measured masses followed by the LC purity measured by the UV spectrum (at 254 nm) and finally the retention time (with gradient described in section 5.2.2). Infrared spectra were obtained using a Shimadzu FT-IR Affinity-1 spectrophotometer or a JASCO FT/IR-460 using sodium chloride disks. $[\alpha]_{\text{D}}$ values were calculated using an AA-55 Series Automatic Polarimeter. Melting points are uncorrected. TLC was performed on Merck 60F254 silica plates and visualised by UV light and/or anisaldehyde or potassium permanganate stains. All chemicals were purchased from a chemical supplier and used as received.

5.1.2 Synthesis of compounds described in Chapter 2

Hydrazides

HZD02 Isophthalic acid dihydrazide

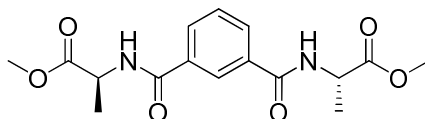


Dimethylisophthalate (1.0 g, 5.2 mmol) was dissolved in dry methanol (25 mL). Hydrazine hydrate (2.3 mL, 30.9 mmol) was added and the reaction was stirred at room temperature overnight. The resulting precipitate was collected by filtration and the white solid was washed with methanol (3 x 25 mL). The final product was obtained as a white solid in 98 % yield. Mass = 1.0 g.

^1H NMR (500 MHz, d_6 -DMSO) δ 9.82 (s, 2H, NH), 8.26 (t, $J = 1.6$ Hz, 1H, ArH), 7.91 (dd, $J = 7.7, 1.8$ Hz, 2H, ArH), 7.52 (t, $J = 7.7$ Hz, 1H, ArH), 4.53 (s, 4H, NH₂); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 165.5 (C=O), 133.6 (ArC-C), 129.3 (ArC-H), 128.5 (ArC-H), 126.1 (ArC-H); IR (film) 3287, 1657, 1620, 1587, 1522, 1479, 1321, 1308, 1109 cm^{-1} .

The spectroscopic data is in agreement with that previously reported.^[1]

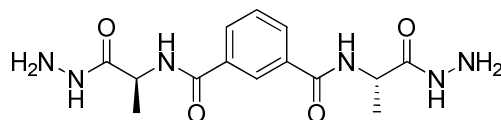
Isophthalic acid alanine methyl ester



To a solution of isophthalic acid (897 mg, 5.4 mmol) in chloroform (50 mL), L-alanine methyl ester (1.5 g, 10.8 mmol) and DMAP (269 mg, 2.2 mmol) were added at 0 °C. After 10 minutes of stirring EDC.HCl (2.53 g, 13.2 mmol) was added. The reaction was stirred overnight at room temperature. The reaction mixture was quenched with water (100 mL), extracted with chloroform (50 mL) and washed with water (3 x 100mL). The organic layer was dried (MgSO_4), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with ethyl acetate (100 %) to yield a white solid, yield = 53 %. Mass = 963 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.16 (t, $J = 1.5$ Hz, 1H, ArH), 7.88 (dd, $J = 7.7, 1.7$ Hz, 2H, ArH), 7.42 (t, $J = 7.6$ Hz, 1H, ArH), 7.07 (d, $J = 7.1$ Hz, 2H, NH), 4.87 – 4.71 (m, 2H, CHCH₃), 3.77 (s, 6H, OCH₃), 1.51 (d, $J = 7.3$ Hz, 6H, CH₃CH); ^{13}C NMR (126 MHz, CDCl_3) δ 173.9 (C=O), 166.3 (C=O), 134.3 (ArC-C), 130.6 (ArC-H), 129.2 (ArC-H), 125.5 (ArC-H), 52.8 (OCH₃), 48.8 (CHCH₃), 18.5 (CH₃-CH); HRMS (ESI+) calcd for $\text{C}_{16}\text{H}_{21}\text{O}_6\text{N}_2$ $[\text{M} + \text{H}]^+$ 337.1394, found 337.1397; IR (film) 3321, 1742, 1636, 1522, 1454, 1207, 1175, 1156 cm^{-1} ; MP 125 – 127 °C; $[\alpha]_{\text{D}} = +60.0$ ($c = 0.5$ in CHCl_3).

HZD03 Isophthalic acid alanine dihydrazide

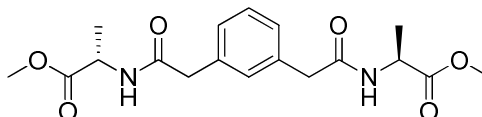


Isophthalic acid alanine methyl ester (500 mg, 1.5 mmol) was dissolved in

dry methanol (20 mL). Hydrazine hydrate (0.68 mL, 8.9 mmol) was added and the reaction was stirred at room temperature overnight. The resulting precipitate was collected by filtration and the white solid was washed with methanol (3 x 20 mL). The final product was obtained as a white solid in 74 % yield. Mass = 370 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 9.19 (s, 2H, NH), 8.57 (d, $J = 7.6$ Hz, 2H, NH), 8.35 (t, $J = 1.7$ Hz, 1H, ArH), 8.01 (dd, $J = 7.7, 1.7$ Hz, 2H, ArH), 7.54 (t, $J = 7.7$ Hz, 1H, ArH), 4.54 – 4.43 (m, 2H, CHCH₃), 4.22 (s, 4H, NH₂), 1.33 (d, $J = 7.2$ Hz, 6H, CH₃CH); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 171.6 (C=O), 165.6 (C=O), 134.1 (ArC-C), 130.3 (ArC-H), 128.1 (ArC-H), 126.8 (ArC-H), 47.8 (CHCH₃), 18.2 (CH₃-CH); HRMS (ESI+) calcd for C₁₄H₂₁O₄N₆ [M + H]⁺ 337.1619, found 337.1616; IR (film) 3291, 1645, 1634, 1607, 1559, 1526, 1271, 1258 cm⁻¹; MP 228 – 232 °C; $[\alpha]_{\text{D}} = +33.3$ (c = 0.57 in H₂O).

1,3-phenylenediacetic acid alanine methyl ester

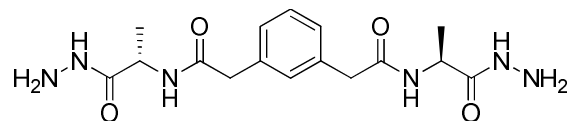


To a solution of 1,3-diphenylenediacetic acid (1.0 g, 5.4 mmol) in chloroform (50 mL), alanine methyl ester (1.5 g, 10.8 mmol) and DMAP (269 mg, 2.2 mmol) were added at 0 °C. After 10 minutes of stirring EDC.HCl (2.5 g, 13.2 mmol) was added. The reaction was stirred overnight at room temperature. The reaction mixture was quenched with water (100 mL), extracted with chloroform (50 mL) and washed with water (3 x 100mL). The organic layer was dried (MgSO₄), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with ethyl acetate (100 %) to yield a white solid, yield = 48 %. Mass = 953 mg.

^1H NMR (500 MHz, CDCl₃) δ 7.35 – 7.30 (m, 1H, ArH), 7.22 – 7.16 (m, 3H, ArH), 6.05 (d, $J = 6.8$ Hz, 2H, NH), 4.60 – 4.52 (m, 2H, CHCH₃), 3.68 (s, 6H, OCH₃), 3.56 (s, 4H, CH₂C(=O)), 1.33 (d, $J = 7.3$ Hz, 6H, CH₃CH); ^{13}C NMR (126 MHz, CDCl₃) δ 173.6 (C=O), 170.5 (C=O), 135.5 (ArC-C), 130.6 (ArC-H), 129.8 (ArC-H), 128.6 (ArC-H), 52.7 (OCH₃), 48.3 (CHCH₃), 43.6 (CH₂), 18.4 (CH₃CH); HRMS (ESI+) calcd for C₁₈H₂₅O₆N₂ [M + H]⁺ 365.1707, found 365.1709; IR (film) 3304, 1749,

1738, 1640, 1530, 1437, 1358, 1211, 1165, 1148 cm^{-1} ; MP 141 – 144 $^{\circ}\text{C}$;
 $[\alpha]_{\text{D}} = +30.0$ ($c = 0.5$ in CHCl_3).

HZD04 1,3-phenylenediacetic acid alanine hydrazide

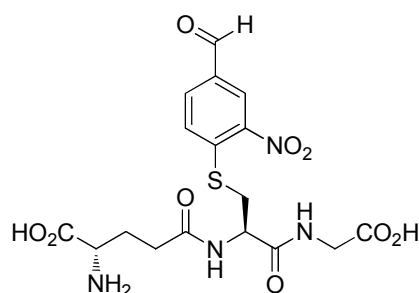


1,3-phenylenediacetic acid alanine methyl ester (500 mg, 1.4 mmol) was dissolved in dry methanol (20 mL). Hydrazine hydrate (0.6 mL, 8.2 mmol) was added and the reaction was stirred at room temperature overnight. The resulting precipitate was collected by filtration and the white solid was washed with methanol (3 x 20 mL). The final product was obtained as a white solid in 68 % yield. Mass = 339 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 9.10 (s, 2H, NH), 8.23 (d, $J = 7.8$ Hz, 2H, NH), 7.22 – 7.15 (m, 1H, ArH), 7.14 – 7.07 (m, 3H, ArH), 4.31 – 4.12 (m, 6H, CHCH₃ and NH₂), 3.41 (d, $J = 2.0$ Hz, 4H, CH₂C(=O)), 1.17 (d, $J = 7.1$ Hz, 6H, CH₃CH); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 171.5 (C=O), 169.6 (C=O), 136.2 (ArC-C), 129.9 (ArC-H), 127.9 (ArC-H), 127.0 (ArC-H), 46.8 (CHCH₃), 41.9 (CH₂), 18.7 (CH₃CH); HRMS (ESI+) calcd for C₁₆H₂₅O₄N₆ [M + H]⁺ 365.1932, found 365.1941; IR (film) 3291, 1661, 1620, 1539, 1441, 1418, 1354, 1240 cm^{-1} ; MP 214 – 218 $^{\circ}\text{C}$;
 $[\alpha]_{\text{D}} = -40.9$ ($c = 0.44$ in DMSO).

Aldehyde

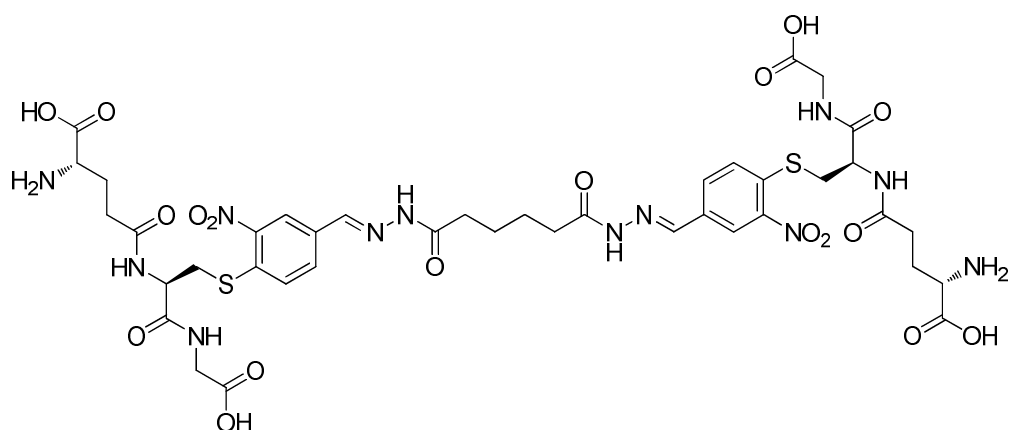
ALD01 Para Glutathione conjugated 3-nitrobenzaldehyde



Synthesised by Venu Bhat.^[2]

Acyldiazide products

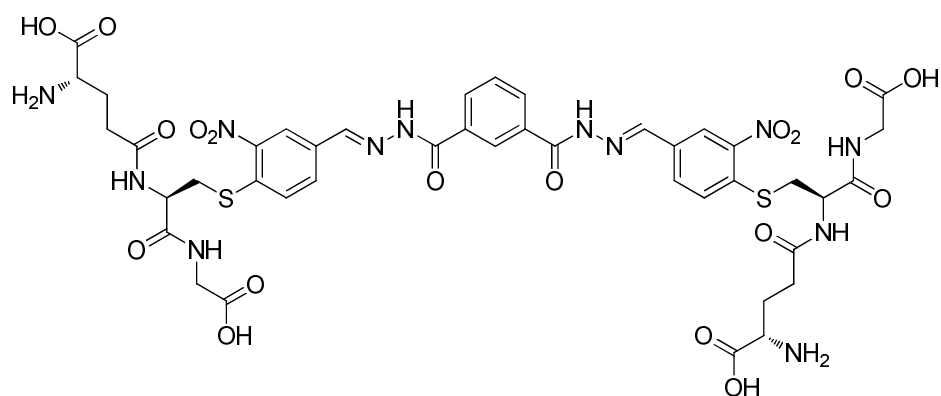
HZN01-01-01



Commercially available bishydrazide **HZD01** (50 mg, 0.3 mmol) and aldehyde **ALD01** (262 mg, 0.6 mmol) were dissolved in water (5 mL). A few drops of glacial acetic acid were added to the stirred mixture until a yellow precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a yellow solid in 58 % yield. Mass = 175 mg.

^1H NMR could not be determined due to poor solubility of sample. HRMS (ESI $^-$) calcd for $\text{C}_{40}\text{H}_{48}\text{O}_{18}\text{N}_{12}\text{S}_2$ $[\text{M} - 2\text{H}]^{2-}$ 524.1331, found 524.1346; HPLC (UV λ = 254 nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 16.6 min; IR (film) 1643, 1607, 1514, 1400, 1337, 1221, 1107 cm^{-1} ; MP decomposition observed > 200 $^\circ\text{C}$.

HZN01-02-01

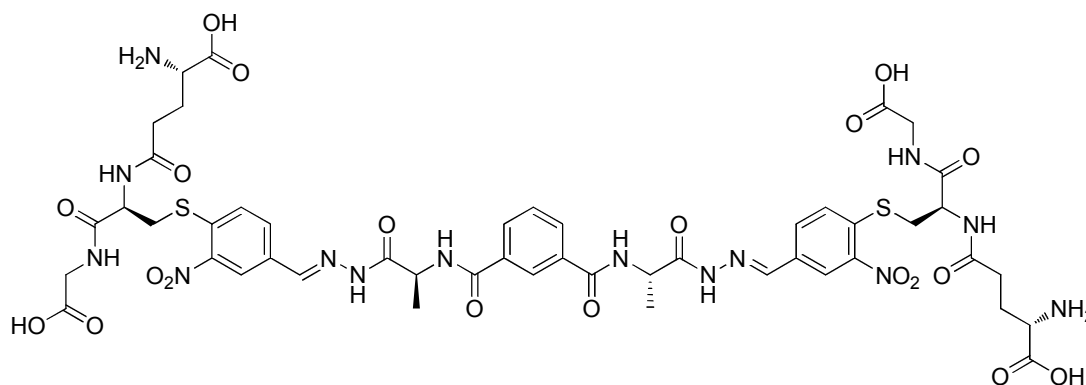


Bishydrazide **HZD02** (50 mg, 0.3 mmol) and aldehyde **ALD01** (235 mg, 0.5 mmol) were dissolved in water (5 mL). A few drops of glacial acetic acid were

added to the stirred mixture until a yellow precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a yellow solid in 71 % yield. Mass = 195 mg.

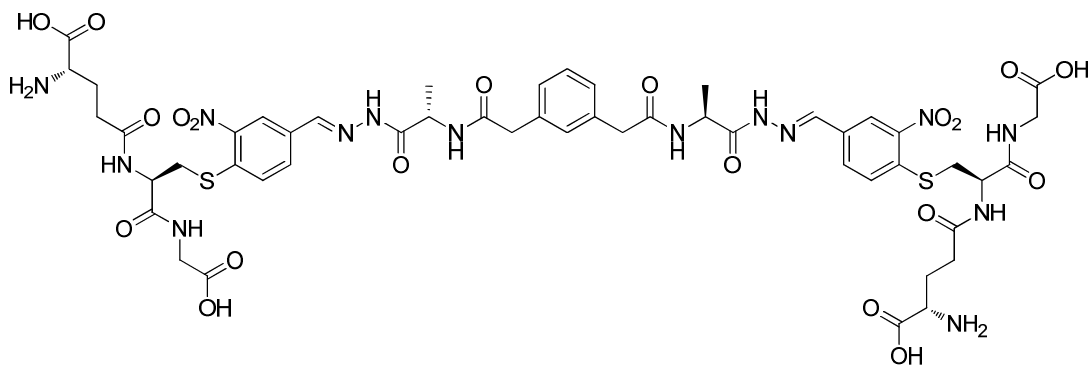
^1H NMR could not be determined due to poor solubility of sample. LC/MS mass calcd for $\text{C}_{42}\text{H}_{47}\text{O}_{18}\text{N}_{12}\text{S}_2$ $[\text{M} + \text{H}]^+$ 1071.2, found 1071.0; HPLC (UV $\lambda = 254$ nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 19.4 min; IR (film) 1643, 1607, 1514, 1398, 1337, 1227 cm^{-1} ; MP decomposition observed > 200 $^\circ\text{C}$.

HZN01-03-01



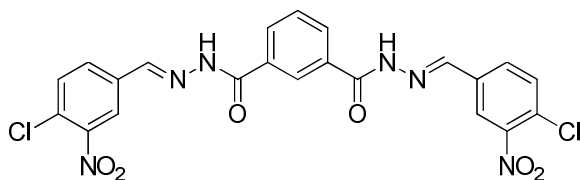
Bishydrazide **HZD03** (100 mg, 0.3 mmol) and aldehyde **ALD01** (271 mg, 0.6 mmol) were dissolved in water (10 mL). A few drops of glacial acetic acid were added to the stirred mixture until a yellow precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a yellow solid in 24 % yield. Mass = 87 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 12.00 – 11.81 (m, 1H), 11.71 – 11.52 (m, 1H), 9.02 – 8.86 (m, 1H), 8.86 – 8.61 (m, 4H), 8.51 – 8.37 (m, 3H), 8.36 – 8.30 (m, 1H), 8.18 – 7.94 (m, 5H), 7.80 (m, 3H), 7.66 – 7.50 (m, 1H), 5.45 – 5.26 (m, 1H), 4.64 – 4.45 (m, 3H), 3.84 – 2.86 (m, 13H), 2.40 – 2.20 (m, 3H), 2.04 – 1.71 (m, 4H), 1.44 (t, $J = 7.8$ Hz, 4H), 1.23 (s, 3H). HRMS (ESI+) calcd for $\text{C}_{48}\text{H}_{57}\text{O}_{20}\text{N}_{14}\text{S}_2$ $[\text{M} + \text{H}]^+$ 1213.3309, found 1213.3286; HPLC (UV $\lambda = 254$ nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 18.5 min; IR (film) 1643, 1609, 1514, 1398, 1337, 1229, 1105 cm^{-1} ; MP decomposition observed > 200 $^\circ\text{C}$.

HZN01-04-01

Bishydrazide **HZD04** (50 mg, 0.1 mmol) and aldehyde **ALD01** (125 mg, 0.3 mmol) were dissolved in water (15 mL). A few drops of glacial acetic acid were added to the stirred mixture until a yellow precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a yellow solid in 44 % yield. Mass = 75 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 11.81 (d, $J = 45.4$ Hz, 1H), 11.61 (s, 1H), 8.81 (s, 2H), 8.65 (s, 2H), 8.60 – 8.52 (m, 1H), 8.44 (s, 1H), 8.41 – 8.32 (m, 2H), 8.27 (d, $J = 11.2$ Hz, 1H), 8.04 (d, $J = 8.7$ Hz, 1H), 7.96 (d, $J = 7.5$ Hz, 2H), 7.83 – 7.68 (m, 2H), 7.24 – 7.05 (m, 4H), 5.19 – 5.07 (m, 1H), 4.61 – 4.51 (m, 2H), 4.39 – 4.28 (m, 1H), 3.79 – 3.66 (m, 5H), 3.58 – 3.30 (m, 9H), 3.30 – 3.14 (m, 4H), 2.32 (s, 4H), 2.02 – 1.78 (m, 4H), 1.36 (m, 6H). HRMS (ESI+) calcd for $\text{C}_{50}\text{H}_{61}\text{O}_{20}\text{N}_{14}\text{S}_2$ $[\text{M} + \text{H}]^+$ 1241.3622, found 1241.3596; HPLC (UV $\lambda = 254$ nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 17.5 min; IR (film) 1641, 1609, 1516, 1398, 1337, 1227, 1107 cm^{-1} ; MP decomposition observed > 200 $^\circ\text{C}$.

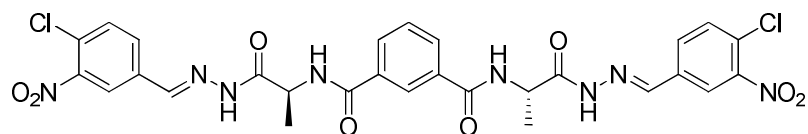
HZN02-02-02

Bishydrazide **HZD02** (50 mg, 0.3 mmol) and aldehyde **ALD02** (96 mg, 0.5 mmol) were dissolved in an ethanol/water mix (10 mL, 1:1). A few drops of glacial acetic acid were added to the stirred mixture until a precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The

solid was resuspended in water and lyophilised. The final product was obtained as a white solid in 79 % yield. Mass = 107 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 12.35 (s, 2H, NH), 8.54 (s, 2H, CH=N), 8.49 (s, 1H, ArH), 8.42 (s, 2H, ArH), 8.16 (d, $J = 7.5$ Hz, 2H, ArH), 8.09 (d, $J = 8.3$ Hz, 2H, ArH), 7.89 (d, $J = 8.4$ Hz, 2H, ArH), 7.74 (t, $J = 7.6$ Hz, 1H, ArH); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 191.3, 163.3, 148.4, 145.2, 136.2, 135.4, 134.0, 133.9, 133.3, 132.7, 132.1, 131.6, 129.4, 127.7, 126.7, 126.4, 124.1; LC/MS mass calcd for $\text{C}_{22}\text{H}_{15}\text{O}_6\text{N}_6\text{Cl}_2$ $[\text{M} + \text{H}]^+$ 529.1, found 529.0; HPLC (UV $\lambda = 254$ nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 34.2 min; IR (film) 3190, 3028, 1659, 1605, 1557, 1535, 1343, 1304, 1277, 1256, 1225, 1171, 1125 cm^{-1} .

HZN02-03-02

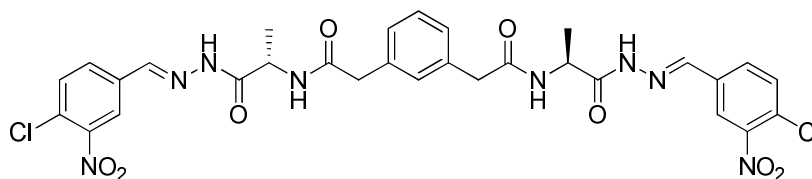


Bishydrazide **HZD03** (50 mg, 0.15 mmol) and aldehyde **ALD02** (55 mg, 0.3 mmol) were dissolved in an ethanol/water mix (10 mL, 1:1). A few drops of glacial acetic acid were added to the stirred mixture until a precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a white solid in 64 % yield. Mass = 64 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 11.88 (s, 1H, NH), 11.67 (s, 1H, NH), 8.87 – 8.80 (m, 1H, CH=N), 8.80 (dd, $J = 6.9, 5.0$ Hz, 1H, CH=N), 8.44 – 8.34 (m, 3H, ArH), 8.33 (s, 1H, NH), 8.10 – 7.97 (m, 5H, 4 x ArH and 1 x NH), 7.84 (dd, $J = 8.4, 3.5$ Hz, 2H, ArH), 7.62 – 7.53 (m, 1H, ArH), 5.41 – 5.30 (m, 1H, CHCH₃), 4.61 – 4.51 (m, 1H, CHCH₃), 1.43 (dd, $J = 7.0, 5.3$ Hz, 6H, CH₃CH); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 174.0, 169.5, 166.0, 165.7, 148.1, 147.9, 143.5, 140.0, 135.0, 134.9, 134.2, 134.0, 132.2, 132.1, 131.4, 131.1, 130.3, 128.2, 126.9, 125.7, 125.3, 123.6, 123.3, 48.6, 46.3, 17.5, 16.7; HRMS (ESI+) calcd for $\text{C}_{28}\text{H}_{28}\text{O}_8\text{N}_9\text{Cl}_2$ $[\text{M} + \text{NH}_4]^+$ 688.1432, found 688.1424; HPLC (UV $\lambda = 254$ nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 33.5 min; IR (film) 3300, 3071, 1682, 1645,

1532, 1478, 1348, 1258, 1233, 1130 cm^{-1} ; MP 198 – 204 $^{\circ}\text{C}$; $[\alpha]_{\text{D}} = +216.6$ ($c = 0.3$ in 2:1 $\text{CHCl}_3 / \text{DMSO}$).

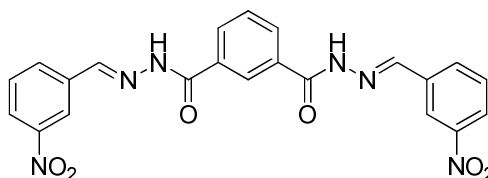
HZN02-04-02



Bishydrazide **HZD04** (50 mg, 0.1 mmol) and aldehyde **ALD02** (51 mg, 0.3 mmol) were dissolved in an ethanol/water mix (10 mL, 1:1). A few drops of glacial acetic acid were added to the stirred mixture until a precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a white solid in 76 % yield. Mass = 73 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 11.81 (s, 1H, NH), 11.63 (s, 1H, NH), 8.54 – 8.46 (m, 1H, CH=N), 8.44 – 8.39 (m, 1H, CH=N), 8.37 – 8.30 (m, 2H, ArH), 8.26 (s, 1H, NH), 8.07 – 7.92 (m, 3H, 2 x ArH and 1 x NH), 7.84 (dd, $J = 8.4, 3.9$ Hz, 1H, ArH), 7.79 (d, $J = 8.4$ Hz, 1H, ArH), 7.25 – 7.05 (m, 4H, ArH), 5.13 (dq, $J = 7.8, 7.1$ Hz, 1H, CHCH₃), 4.34 (dq, $J = 7.8, 7.1$ Hz, 1H, CHCH₃), 3.51 – 3.39 (m, 4H, CH₂), 1.28 (d, $J = 7.2$ Hz, 6H, CH₃CH); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 174.6, 170.5, 170.1, 169.9, 148.5, 148.3, 144.0, 140.4, 136.5, 135.4, 135.3, 132.6, 132.5, 132.0, 131.5, 130.3, 128.4, 127.4, 126.2, 125.7, 124.0, 123.7, 48.1, 45.72, 42.3, 42.2, 18.4, 17.6; LC/MS mass calcd for $\text{C}_{30}\text{H}_{29}\text{O}_8\text{N}_8\text{Cl}_2$ $[\text{M} + \text{H}]^+$ 699.1, found 699.2; HPLC (UV $\lambda = 254$ nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 33.0 min; IR (film) 3269, 3063, 2976, 1678, 1645, 1609, 1532, 1449, 1397, 1354, 1263, 1132 cm^{-1} .

HZN03-02-03

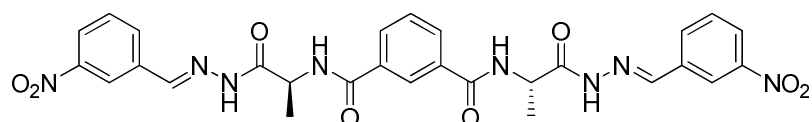


Bishydrazide **HZD02** (50 mg, 0.3 mmol) and aldehyde **ALD03** (78 mg, 0.6 mmol) were dissolved in an ethanol/water mix (10 mL, 1:1). A few drops of

glacial acetic acid were added to the stirred mixture until a precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a white solid in 46 % yield. Mass = 55 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 12.33 (s, 2H, NH), 8.62 (s, 2H, CH=N or ArH), 8.60 (s, 2H, CH=N or ArH), 8.51 (s, 1H, ArH), 8.30 (d, J = 7.6 Hz, 2H, ArH), 8.23 – 8.14 (m, 4H, ArH), 7.83 – 7.72 (m, 3H, ArH); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 163.3, 148.7, 146.3, 136.6, 134.0, 131.6, 131.0, 129.4, 127.6, 124.9, 121.5; LC/MS mass calcd for $\text{C}_{22}\text{H}_{17}\text{O}_6\text{N}_6$ $[\text{M} + \text{H}]^+$ 461.1, found 461.1; HPLC (UV λ = 254 nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 29.6 min; IR (film) 3181, 3017, 1665, 1649, 1564, 1524, 1350, 1306, 1273, 1169 cm^{-1} .

HZN03-03-03

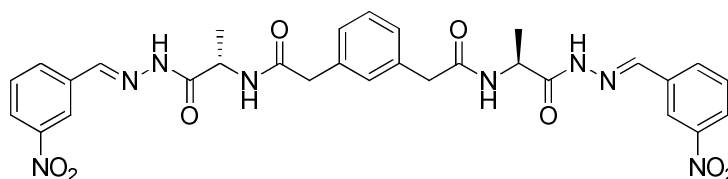


Bishydrazide **HZD03** (50 mg, 0.15 mmol) and aldehyde **ALD03** (45 mg, 0.30 mmol) were dissolved in an ethanol/water mix (10 mL, 1:1). A few drops of glacial acetic acid were added to the stirred mixture until a precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a white solid in 70 % yield. Mass = 62 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 11.88 (s, 1H, NH), 11.67 (s, 1H, NH), 8.88 (t, J = 6.0 Hz, 1H, CH=N), 8.78 (dd, J = 7.2, 3.3 Hz, 1H, CH=N), 8.53 (s, 1H, ArH), 8.50 (s, 1H, ArH), 8.43 (t, J = 6.8 Hz, 1H, ArH), 8.40 (s, 1H, NH), 8.26 (d, J = 8.1 Hz, 2H, ArH), 8.20 – 8.12 (m, 3H, ArH and NH), 8.10 – 8.04 (m, 2H, ArH), 7.75 (td, J = 8.0, 1.8 Hz, 2H, ArH), 7.64 – 7.55 (m, 1H, ArH), 5.39 (dq, J = 8.4, 7.3 Hz, 1H, CHCH₃), 4.59 (dq, J = 8.4, 7.3 Hz, 1H, CHCH₃), 1.46 (dd, J = 13.1, 7.2 Hz, 6H, CH₃CH); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 174.5, 169.9, 166.5, 166.2, 148.8, 148.7, 145.0, 141.6, 136.7, 136.5, 134.7, 134.5, 133.7, 133.2, 130.9, 130.8, 128.7, 127.4, 124.7, 124.6, 121.5, 49.1, 46.7, 18.0, 17.2; LC/MS mass calcd for $\text{C}_{28}\text{H}_{27}\text{O}_8\text{N}_8$ $[\text{M} + \text{H}]^+$ 603.2, found 603.2; HPLC (UV λ = 254 nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 29.0 min; IR (film) 3225, 3075, 1682,

1632, 1570, 1530, 1350, 1242 cm^{-1} .

HZN03-04-03



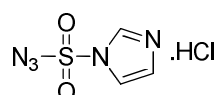
Bishydrazide **HZD04** (50 mg, 0.15 mmol) and aldehyde **ALD03** (42 mg, 0.3 mmol) were dissolved in an ethanol/water mix (10 mL, 1:1). A few drops of glacial acetic acid were added to the stirred mixture until a precipitate was observed. The solid was collected by filtration and washed with ice cold ether (2 x 10 mL). The solid was resuspended in water and lyophilised. The final product was obtained as a white solid in 45 % yield. Mass = 39 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 11.77 (s, 1H, NH), 11.59 (s, 1H, NH), 8.55 – 8.47 (m, 2H, CH=N and ArH), 8.45 – 8.38 (m, 2H, CH=N and ArH), 8.32 (s, 1H, NH), 8.28 – 8.20 (m, 2H, ArH), 8.15 – 8.07 (m, 3H, ArH and NH), 7.79 – 7.66 (m, 2H, ArH), 7.26 – 7.06 (m, 4H, ArH), 5.16 (dq, $J = 7.6, 7.2$ Hz, 1H, CHCH₃), 4.36 (dq, $J = 7.6, 7.2$ Hz, 1H, CHCH₃), 3.53 – 3.38 (m, 4H, CH₂), 1.31 (dd, $J = 11.0, 7.1$ Hz, 6H, CH₃CH); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 174.5, 170.5, 170.1, 169.8, 148.7, 148.7, 145.0, 141.5, 136.6, 136.6, 136.4, 133.7, 133.1, 130.8, 130.3, 128.4, 127.5, 127.4, 124.7, 124.5, 121.5, 48.1, 45.7, 42.3, 42.2, 18.5, 17.6; LC/MS mass calcd for C₃₀H₃₁O₈N₈ [M + H]⁺ 631.2, found 631.2; HPLC (UV $\lambda = 254$ nm, Gradient A (section 5.2.1)) purity > 95 %, retention time 28.0 min; IR (film) 3262, 3208, 3065, 1686, 1668, 1647, 1528, 1348, 1234 cm^{-1} .

5.1.3 Synthesis of compounds described in Chapter 3

Azide linker library

Imidazole sulfonylazide



Sulfonyl chloride (16 mL, 200 mol) was added drop-wise to an ice-cooled suspension of sodium azide (13.0 g, 200 mmol) in acetonitrile (200 mL). The reaction mixture was stirred at room temperature overnight. Imidazole (25.9 g,

380 mmol) was added portion-wise to the ice-cooled reaction mixture. The resulting slurry was stirred at room temperature for 3 hours. The reaction mixture was then diluted with ethyl acetate (400 mL) and washed with water (2 x 400 mL) followed by saturated aqueous sodium bicarbonate solution (2 x 400 mL). The organic layer was dried (MgSO_4) and filtered. A solution of HCl in ethanol was made by the drop-wise addition of acetyl chloride (21.3 mL, 300 mmol) to ice-cooled dry ethanol. This solution of HCl in ethanol was added to the filtrate with stirring. The mixture was chilled in an ice-bath to obtain colourless needles that were collected by filtration and washed with ethyl acetate (3 x 100 mL). Yield = 65 %. Mass = 27.2 g.

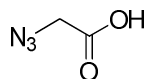
^1H NMR (500 MHz, D_2O) δ 9.45 (dd, $J = 1.8, 1.4$ Hz, 1H, ArH), 8.08 (dd, $J = 2.2, 1.9$ Hz, 1H, ArH), 7.67 (dd, $J = 2.2, 1.3$ Hz, 1H, ArH); ^{13}C NMR (126 MHz, D_2O) δ 137.7 (ArCH), 123.3 (ArCH), 120.1 (ArCH); IR (thin film, NaCl plate): 2920, 2853, 2171, 1456, 1428, 1231, 1192, 1162, 1139 cm^{-1}

The spectroscopic data is in agreement with that previously reported.^[3]

General procedure A- diazotransfer step

Amino acid (250 mg, 2.8 mmol), potassium carbonate (853 mg, 6.2 mmol) and copper (II) sulfate pentahydrate (7 mg, 28 μmol) were dissolved in methanol (15 mL) at room temperature. Imidazole-1-sulfonyl azide.HCl (704 mg, 3.4 mmol) was added to this solution, which was stirred at room temperature overnight. The mixture was concentrated under reduced pressure and then diluted with water (50 mL), which was then acidified with HCl (conc.). The product was extracted in to ethyl acetate (3 x 30 mL). The organic layer was dried (MgSO_4), filtered and concentrated under reduced pressure.

Azidoglycine

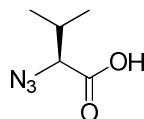


Azidoglycine was synthesised by general procedure A, starting from glycine (1.0 g, 13.3 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane/acetic acid (50 / 48 / 2) to yield a colourless oil. Yield = 82 %. Mass = 1.1 g.

^1H NMR (500 MHz, CDCl_3) δ 11.20 (s, 1H, COOH), 3.94 (s, 2H, CH_2); ^{13}C NMR (126 MHz, CDCl_3) δ 175.1 (COOH), 50.5 (CH_2); IR (film) 3104, 2108, 1722, 1418, 1281, 1223 cm^{-1} .

The spectroscopic data is in agreement with that previously reported.^[4]

Azidovaline

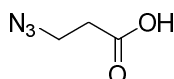


Azidovaline was synthesised by general procedure A, starting from L-valine (5.0 g, 43 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane/acetic acid (20 / 79 / 1) to yield a colourless oil. Yield = 88 %. Mass = 4.9 g.

^1H NMR (500 MHz, CDCl_3) δ 11.27 (s, 1H, COOH), 3.75 (d, $J = 5.7$ Hz, 1H, N_3CH), 2.29 – 2.17 (m, 1H, $\text{CH}(\text{CH}_3)_2$), 1.04 (d, $J = 6.8$ Hz, 3H, CH_3), 0.99 (d, $J = 6.8$ Hz, 3H, CH_3); ^{13}C NMR (126 MHz, CDCl_3) δ 176.8 (COOH), 68.1 (CH), 31.1 (CH), 19.6 (CH_3), 17.9 (CH_3); IR (film) 2970, 2106, 1713, 1267, 1227 cm^{-1} ; $[\alpha]_{\text{D}} = -53.0$ ($c = 1.0$ in CHCl_3).

The spectroscopic data is in agreement with that previously reported.^[4]

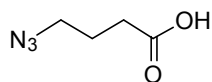
Azido- β -alanine



Azido- β -alanine was synthesised by general procedure A, starting from β -alanine (250 mg, 2.8 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane/acetic acid (20 / 79 / 1) to yield a colourless oil. Yield = 86 %. Mass = 277 mg.

^1H NMR (500 MHz, CDCl_3) δ 10.19 (s, 1H, COOH), 3.50 (t, $J = 6.4$ Hz, 2H, CH_2), 2.62 (t, $J = 6.4$ Hz, 2H, CH_2); ^{13}C NMR (126 MHz, CDCl_3) δ 177.2 (COOH), 46.6 (CH_2), 34.0 (CH_2); IR (film) 2943, 2099, 1713, 1402, 1244, 1179 cm^{-1} .

The spectroscopic data is in agreement with that previously reported.^[5]

4-azidobutyric acid

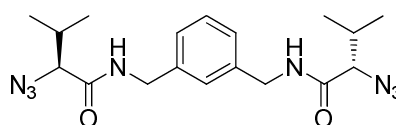
4-azidobutyric acid was synthesised by general procedure A, starting from 4-aminobutyric acid (1.0 g, 9.7 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane/acetic acid (50 / 49 / 1) to yield a colourless oil. Yield = 87 %. Mass = 1.1 g.

^1H NMR (500 MHz, CDCl_3) δ 11.43 (s, 1H, COOH), 3.39 (t, J = 6.7 Hz, 2H, CH_2), 2.49 (t, J = 7.2 Hz, 2H, CH_2), 1.93 (m, 2H, CH_2); ^{13}C NMR (126 MHz, CDCl_3) δ 179.3 (COOH), 50.4 (CH_2), 30.9 (CH_2), 23.9 (CH_2); IR (film) 2940, 2097, 1707, 1414, 1281, 1242 cm^{-1} .

The spectroscopic data is in agreement with that previously reported.^[6]

General procedure B – EDC-mediated amide coupling step

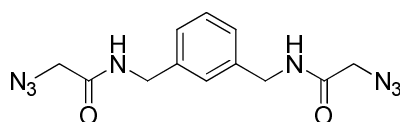
Azido acid (100 mg, 1.0 mmol) and diamine (67 mg, 0.5 mmol) were dissolved in chloroform (10 mL) at 0 °C and DMAP (24 mg, 0.2 mmol) was added. The solution was stirred for 10 minutes and then EDC.HCl (230 mg, 1.2 mmol) was added. The solution was allowed to warm to room temperature and stirred overnight. The reaction mixture was quenched with water (25 mL), extracted with dichloromethane (25 mL) and washed with water (3 x 50 mL). The organic layer was dried (MgSO_4), filtered and concentrated under reduced pressure.

AZ01

AZ01 was synthesised by general procedure B, starting from azidovaline (2.0 g, 14 mmol) and *m*-xylylenediamine (0.9 g, 7 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (7:3) to yield a white solid. Yield = 66 %. Mass = 1.8 g.

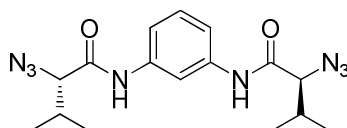
^1H NMR (500 MHz, CDCl_3) δ 7.39 – 7.33 (m, 1H, ArH), 7.26 – 7.20 (m, 3H, ArH), 6.69 (s, 2H, NH), 4.53 – 4.42 (m, 4H, CH_2), 3.92 (d, J = 3.9 Hz, 2H, N_3CH), 2.50 – 2.39 (m, 2H, $\text{CH}(\text{CH}_3)_2$), 1.12 (d, J = 6.9 Hz, 6H, CH_3), 0.94 (d, J = 6.8 Hz,

6H, CH₃); ¹³C NMR (126 MHz, CDCl₃) δ 168.9 (C=O), 138.4 (ArC), 129.3 (ArCH), 127.3 (ArCH), 127.1 (ArCH), 70.8 (CH), 43.3 (CH₂), 32.1 (CH), 19.7 (CH₃), 16.7 (CH₃); HRMS (ESI+) calcd for C₁₈H₃₀O₂N₉ [M + NH₄]⁺ 404.2517, found 404.2520; IR (film) 3277, 2093, 1641, 1547, 1225 cm⁻¹; MP 148 – 151 °C; [α]_D = - 49.0 (c = 1.0 in CHCl₃)

AZ02

AZ02 was synthesised by general procedure B, starting from azidoglycine (100 mg, 1.0 mmol) and m-xylylenediamine (67 mg, 0.5 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 32 %. Mass = 48 mg.

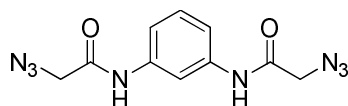
¹H NMR (500 MHz, CDCl₃) δ 7.33 – 7.29 (m, 1H, ArH), 7.21 – 7.15 (m, 3H, ArH), 6.72 (s, 2H, NH), 4.46 (d, *J* = 5.9 Hz, 4H, CH₂), 4.02 (s, 4H, CH₂); ¹³C NMR (126 MHz, CDCl₃) δ 166.6 (C=O), 138.2 (ArC), 129.3 (ArCH), 127.3 (ArCH), 127.2 (ArCH), 52.7 (CH₂), 43.2 (CH₂); LC/MS mass calcd for C₁₂H₁₅O₂N₈ [M + H]⁺ 303.1, found 303.3; HPLC (UV λ = 254 nm, Gradient B (section 5.2.2)) purity > 90 %, retention time 2.92 min; IR (film) 3275, 2924, 2097, 1649, 1557, 1452, 1271, 1207 cm⁻¹.

AZ03

AZ03 was synthesised by general procedure B, starting from azidovaline (500 mg, 3.5 mmol) and 1,3-phenylenediamine (189 mg, 1.8 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 57 %. Mass = 358 mg.

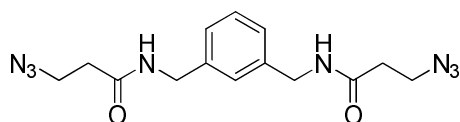
¹H NMR (500 MHz, CDCl₃) δ 8.19 (s, 2H, NH), 7.92 (t, *J* = 2.0 Hz, 1H, ArH), 7.43 – 7.38 (m, 2H, ArH), 7.36 – 7.30 (m, 1H, ArH), 4.00 (d, *J* = 4.0 Hz, 2H, N₃CH), 2.60 – 2.40 (m, 2H, CH(CH₃)₂), 1.16 (d, *J* = 6.9 Hz, 6H, CH₃), 0.99 (d, *J* = 6.8 Hz, 6H, CH₃); ¹³C NMR (126 MHz, CDCl₃) δ 167.4 (C=O), 137.6 (ArC), 129.8 (ArCH),

116.2 (ArCH), 111.3 (ArCH), 70.9 (CH), 32.4 (CH), 19.7 (CH₃), 16.8 (CH₃); LC/MS mass calcd for C₁₆H₂₃O₂N₈ [M + H]⁺ 359.2, found 359.2; HPLC (UV λ = 254 nm, Gradient B (section 5.2.2)) purity > 95 %, retention time 6.30 min; IR (film) 3283, 2967, 2097, 1661, 1607, 1533, 1489, 1424, 1219 cm⁻¹.

AZ04

AZ04 was synthesised by general procedure B, starting from azidoglycine (100 mg, 1 mmol) and 1,3-phenylenediamine (54 mg, 0.5 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (7:3) to yield a white solid. Yield = 39 %. Mass = 53 mg.

¹H NMR (500 MHz, CDCl₃) δ 8.06 (s, 2H, NH), 7.89 (t, *J* = 1.9 Hz, 1H, ArH), 7.37 – 7.26 (m, 3H, ArH), 4.12 (s, 4H, CH₂); ¹³C NMR (126 MHz, CDCl₃) δ 164.9 (C=O), 137.7 (ArC), 130.0 (ArCH), 116.6 (ArCH), 111.7 (ArCH), 53.2 (CH₂); LC/MS mass calcd for C₁₀H₁₁O₂N₈ [M + H]⁺ 275.1, found 275.1; HPLC (UV λ = 254 nm, Gradient B (section 5.2.2)) purity > 90 %, retention time 3.17 min; IR (film) 3266, 2099, 1665, 1605, 1541, 1447, 1267, 1240 cm⁻¹.

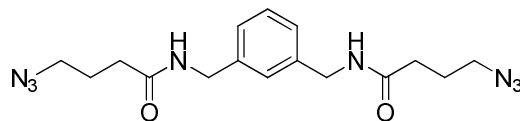
AZ05

AZ05 was synthesised by general procedure B, starting from azido-β-alanine (150 mg, 1.3 mmol) and *m*-xylylenediamine (89 mg, 0.65 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (7:3) to yield a white solid. Yield = 47 %. Mass = 101 mg.

¹H NMR (500 MHz, *d*₆-DMSO) δ 7.64 (t, *J* = 5.6 Hz, 2H, NH), 6.45 – 6.39 (m, 1H, ArH), 6.31 – 6.26 (m, 3H, ArH), 2.70 (t, *J* = 6.4 Hz, 4H, CH₂), 1.67 – 1.64 (m, 4H, CH₂), 1.63 (t, *J* = 6.4 Hz, 4H, CH₂); ¹³C NMR (126 MHz, *d*₆-DMSO) δ 169.5 (C=O), 139.4 (ArC), 128.3 (ArCH), 126.1 (ArCH), 125.7 (ArCH), 47.0 (CH₂), 42.1 (CH₂), 34.5 (CH₂); LC/MS mass calcd for C₁₄H₁₉O₂N₈ [M + H]⁺ 331.2, found 331.4;

HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 90 %, retention time 3.40 min; IR (film) 3289, 2928, 2089, 1634, 1553, 1439, 1281, 1260 cm^{-1} .

AZ06

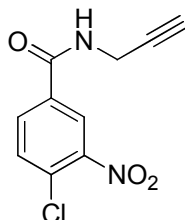


AZ06 was synthesised by general procedure B, starting from 4-azidobutyric acid (300 mg, 2.3 mmol) and m-xylylenediamine (158 mg, 1.2 mmol). The product was purified by flash column chromatography, eluting with ethyl acetate (100 %) to yield a white solid. Yield = 44 %. Mass = 183 mg.

^1H NMR (500 MHz, CDCl_3) δ 7.30 – 7.26 (m, 1H, ArH), 7.19 – 7.09 (m, 3H, ArH), 6.00 (s, 2H, NH), 4.40 (d, $J = 5.8$ Hz, 4H, CH_2), 3.36 (t, $J = 6.5$ Hz, 4H, CH_2), 2.31 (t, $J = 7.2$ Hz, 4H, CH_2), 1.96 – 1.89 (m, 4H, CH_2); ^{13}C NMR (126 MHz, CDCl_3) δ 171.9 (C=O), 139.0 (ArC), 129.4 (ArCH), 127.3 (ArCH), 127.2 (ArCH), 51.0 (CH_2), 43.7 (CH_2), 33.3 (CH_2), 25.0 (CH_2); LC/MS mass calcd for $\text{C}_{16}\text{H}_{23}\text{O}_2\text{N}_8$ $[\text{M} + \text{H}]^+$ 359.2, found 359.4; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 90 %, retention time 4.21 min; IR (film) 3264, 2095, 1634, 1541, 1352, 1248 cm^{-1} .

Alkyne library

ALK01 4-Chloro-3-nitro-*N*-prop-2-ynyl-Benzamide

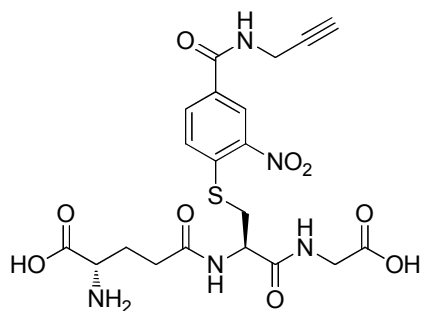


To a solution of 4-chloro-3-nitrobenzoic acid (500 mg, 2.5 mmol) and propargyl amine (0.2 mL, 2.5 mmol) in DCM (20 mL) at 0 °C, DMAP (61 mg, 0.5 mmol) was added. The solution was stirred for 10 minutes and then EDC.HCl (571 mg, 3.0 mmol) was added. The solution was allowed to warm to room temperature and stirred overnight. The reaction mixture was quenched with water (50 mL), extracted with dichloromethane (20 mL) and washed with water (3 x 50mL). The organic layer was dried (MgSO_4), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting

with hexane/ethyl acetate (3:2) to yield a white solid. Yield = 89 %. Mass = 531 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.27 (d, $J = 2.1$ Hz, 1H, ArH), 7.94 (dd, $J = 8.4$, 2.1 Hz, 1H, ArH), 7.66 (d, $J = 8.4$ Hz, 1H, ArH), 6.46 (s, 1H, NH), 4.25 (dd, $J = 5.2$, 2.6 Hz, 2H, $\text{CH}_2\text{C}\equiv\text{CH}$), 2.30 (t, $J = 2.6$ Hz, 1H, $\text{C}\equiv\text{CH}$); ^{13}C NMR (126 MHz, CDCl_3) δ 164.0 (C=O), 148.0 (ArC), 133.7 (ArC), 132.7 (ArCH), 131.8 (ArCH), 130.9 (ArC), 124.4 (ArCH), 78.8 ($-\text{C}\equiv$), 72.8 ($\equiv\text{CH}$), 30.4 (CH_2); IR (film) 3306, 3254, 1632, 1605, 1528, 1470, 1443, 1333, 1290, 1244 cm^{-1} .

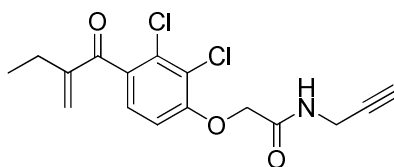
ALK02 2-Amino-4-[1-(carboxymethyl-carbamoyl)-2-(2-nitro-4-prop-2-ynylcarbamoyl-phenylsulfanyl)-ethylcarbamoyl]-butyric acid



Reduced glutathione (58 mg, 0.2 mmol) and sodium carbonate (20 mg, 0.2 mmol) were dissolved in water (2.0 mL) and added to a solution of 4-chloro-3-nitro-*N*-prop-2-ynyl-benzamide (50 mg, 0.2 mmol) in acetonitrile (2.0 mL). The reaction mixture was stirred at room temperature for 48 hours. The solvent was concentrated under reduced pressure. The product was purified by flash column chromatography using reverse phase silica and eluting with 95% water / 5% methanol to yield a yellow solid after lyophilisation. Yield = 11 %. Mass = 11 mg.

LC/MS mass calcd for $\text{C}_{20}\text{H}_{24}\text{O}_9\text{N}_5\text{S}$ $[\text{M} + \text{H}]^+$ 510.1, found 510.1; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 95 %; IR (film) 3269, 1641, 1609, 1543, 1516, 1395, 1333, 1317 cm^{-1} .

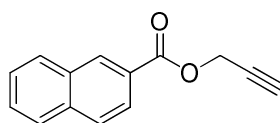
ALK03 **2-[2,3-Dichloro-4-(2-methylene-butyl)-phenoxy]-N-prop-2-ynyl-acetamide**



To a solution of ethacrynic acid (100 mg, 0.3 mmol) and propargyl amine (23 μ L, 0.3 mmol) in DCM (5.0 mL) at 0 °C, DMAP (8 mg, 0.07 mmol) was added. The solution was stirred for 10 minutes and then EDC.HCl (76 mg, 0.4 mmol) was added. The solution was allowed to warm to room temperature and stirred overnight. The reaction mixture was quenched with water (20 mL), extracted with dichloromethane (10 mL) and washed with water (3 x 20mL). The organic layer was dried (MgSO_4), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with hexane/ethyl acetate (1:1) to yield a white solid. Yield = 85 %. Mass = 87 mg.

^1H NMR (500 MHz, CDCl_3) δ 7.22 (d, $J = 8.5$ Hz, 1H, ArH), 6.97 (s, 1H, NH), 6.89 (d, $J = 8.5$ Hz, 1H, ArH), 5.99 (s, 1H, C=CH), 5.62 (s, 1H, C=CH), 4.63 (s, 2H, OCH_2), 4.22 (dd, $J = 5.5, 2.5$ Hz, 2H, $\text{NHCH}_2\text{C}=\text{CH}$), 2.50 (q, $J = 7.4$ Hz, 2H, CH_2CH_3), 2.32 (t, $J = 2.5$ Hz, 1H, C=CH), 1.18 (t, $J = 7.5$ Hz, 3H, CH_3CH_2); ^{13}C NMR (126 MHz, CDCl_3) δ 195.6, 166.5, 154.4, 150.2, 134.4, 131.6, 128.8, 127.2, 123.1, 111.0, 78.7, 72.1, 68.2, 29.0, 23.4, 12.4; HRMS (ESI+) calcd for $\text{C}_{16}\text{H}_{19}\text{O}_3\text{N}_2\text{Cl}_2$ [$\text{M} + \text{NH}_4$] $^+$ 357.0767, found 357.0768; IR (film) 3285, 1651, 1586, 1554.6, 1468, 1447, 1379, 1294, 1246 cm^{-1} ; MP 118 – 120 °C.

ALK04

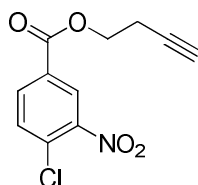


To a solution of 2-naphthoic acid (1.0 g, 5.8 mmol) and propargyl alcohol (338 μ L, 5.8 mmol) in chloroform (40 mL) at 0 °C, DMAP (142 mg, 1.2 mmol) was added. The solution was stirred for 10 minutes and then EDC.HCl (1.3 g, 7.0 mmol) was added. The solution was allowed to warm to room temperature and stirred overnight. The reaction mixture was quenched with water (100 mL),

extracted with dichloromethane (50 mL) and washed with water (3 x 100mL). The organic layer was dried (MgSO₄), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with hexane/ethyl acetate (3:2) to yield a white solid. Yield = 93 %. Mass = 1.1 g.

¹H NMR (500 MHz, CDCl₃) δ 8.63 (s, 1H, ArH), 8.06 (dd, *J* = 8.6, 1.4 Hz, 1H, ArH), 7.95 (d, *J* = 8.1 Hz, 1H, ArH), 7.90 – 7.83 (m, 2H, ArH), 7.59 (t, *J* = 7.2 Hz, 1H, ArH), 7.54 (t, *J* = 7.5 Hz, 1H, ArH), 4.98 (d, *J* = 2.3 Hz, 2H, CH₂C≡CH), 2.53 (t, *J* = 2.3 Hz, 1H, C≡CH); ¹³C NMR (126 MHz, CDCl₃) δ 166.2 (C=O), 135.9 (ArC), 132.7 (ArC), 131.7 (ArCH), 129.6 (ArCH), 128.7 (ArCH), 128.5 (ArCH), 128.0 (ArCH), 127.0 (ArCH), 126.8 (ArC), 125.4 (ArCH), 78.0 (-C≡), 75.3 (≡CH), 52.8 (CH₂); IR (film) 3273, 1707, 1435, 1370, 1287, 1229 cm⁻¹.

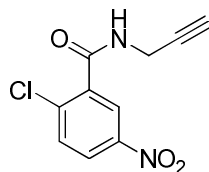
ALK05



To a solution of 4-chloro-3-nitrobenzoic acid (250 mg, 1.2 mmol) and propargyl amine (94 μL, 1.2 mmol) in DCM (10.0 mL) at 0 °C, DMAP (30 mg, 0.2 mmol) was added. The solution was stirred for 10 minutes and then EDC.HCl (285 mg, 1.5 mmol) was added. The solution was allowed to warm to room temperature and stirred overnight. The reaction mixture was quenched with water (50 mL), extracted with dichloromethane (20 mL) and washed with water (3 x 50mL). The organic layer was dried (MgSO₄), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with hexane/ethyl acetate (9:1) to yield a white solid. Yield = 83 %. Mass = 252 mg.

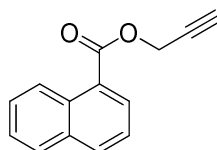
¹H NMR (500 MHz, CDCl₃) δ 8.50 (d, *J* = 2.0 Hz, 1H, ArH), 8.16 (dd, *J* = 8.4, 2.0 Hz, 1H, ArH), 7.66 (d, *J* = 8.3 Hz, 1H, ArH), 4.51 (t, *J* = 6.7 Hz, 2H, CH₂), 2.67 (td, *J* = 6.7, 2.7 Hz, 2H, CH₂C≡CH), 2.02 (t, *J* = 2.7 Hz, 1H, C≡CH); ¹³C NMR (126 MHz, CDCl₃) δ 163.7 (C=O), 148.2 (ArC), 133.9 (ArCH), 132.4 (ArCH), 132.1 (ArC), 130.1 (ArC), 126.9 (ArCH), 79.7 (-C≡), 70.6 (≡CH), 63.7 (CH₂), 19.3 (CH₂); LC/MS mass calcd for C₁₁H₉O₄NCl [M + H]⁺ 254.1, found 254.2; HPLC (UV

$\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 95 %; IR (film) 3275, 1719, 1603, 1537, 1354, 1296, 1240 cm^{-1} .

ALK06

To a solution of 2-chloro-5-nitrobenzoic acid (500 mg, 2.5 mmol) and propargyl amine (0.2 mL, 2.5 mmol) in chloroform (20.0 mL) at 0 °C, DMAP (61 mg, 0.5 mmol) was added. The solution was stirred for 10 minutes and then EDC.HCl (571 mg, 3.0 mmol) was added. The solution was allowed to warm to room temperature and stirred overnight. The reaction mixture was quenched with water (50 mL), extracted with dichloromethane (20 mL) and washed with water (3 x 50mL). The organic layer was dried (MgSO_4), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with hexane/ethyl acetate (1:1) to yield a white solid. Yield = 87 %. Mass = 519 mg.

^1H NMR (500 MHz, d_6 -DMSO) δ 9.18 (t, $J = 5.3$ Hz, 1H, NH), 8.30 (dd, $J = 8.8, 2.8$ Hz, 1H, ArH), 8.24 (d, $J = 2.7$ Hz, 1H, ArH), 7.84 (d, $J = 8.8$ Hz, 1H, ArH), 4.09 (dd, $J = 5.5, 2.5$ Hz, 2H, $\text{CH}_2\text{C}\equiv\text{CH}$), 3.22 (t, $J = 2.5$ Hz, 1H, $\text{C}\equiv\text{CH}$); ^{13}C NMR (126 MHz, d_6 -DMSO) δ 164.5 (C=O), 146.5 (ArC), 137.6 (ArC), 137.5 (ArC), 131.9 (ArCH), 126.1 (ArCH), 124.2 (ArCH), 80.8 ($-\text{C}\equiv$), 74.1 ($\equiv\text{CH}$), 29.0 (CH_2); IR (film) 3279, 1649, 1613, 1537, 1435, 1352, 1310, 1277, 1238 cm^{-1} .

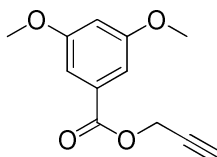
ALK07

To a solution of 1-naphthoic acid (250 mg, 1.5 mmol) and propargyl alcohol (84 μL , 1.5 mmol) in dichloromethane (10.0 mL) at 0 °C, DMAP (35 mg, 0.3 mmol) was added. The solution was stirred for 10 minutes and then EDC.HCl (334 mg, 1.7 mmol) was added. The solution was allowed to warm to room temperature and stirred overnight. The reaction mixture was quenched with water (50 mL), extracted with dichloromethane (20 mL) and washed with water (3 x

50mL). The organic layer was dried (MgSO₄), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with hexane/ethyl acetate (9:1) to yield a white solid. Yield = 91 %. Mass = 287 mg.

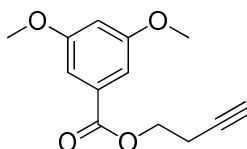
¹H NMR (500 MHz, CDCl₃) δ 8.93 (d, *J* = 8.7 Hz, 1H, ArH), 8.26 – 8.23 (m, 1H, ArH), 8.03 (d, *J* = 8.2 Hz, 1H, ArH), 7.87 (d, *J* = 8.2 Hz, 1H, ArH), 7.64 – 7.59 (m, 1H, ArH), 7.56 – 7.46 (m, 2H, ArH), 5.00 (d, *J* = 2.5 Hz, 2H, CH₂C≡CH), 2.53 (t, *J* = 2.2 Hz, 1H, C≡CH); ¹³C NMR (126 MHz, CDCl₃) δ 166.7 (C=O), 134.1 (ArCH), 134.0 (ArC), 131.6 (ArC), 131.0 (ArCH), 128.8 (ArCH), 128.2 (ArCH), 126.5 (ArCH), 126.2 (ArC), 125.9 (ArCH), 124.7 (ArCH), 78.1 (-C≡), 75.3 (≡CH), 52.7 (CH₂); IR (film) 3264, 1695, 1510, 1277, 1234 cm⁻¹.

ALK08 3,5-Dimethoxy-benzoic acid prop-2-ynyl ester



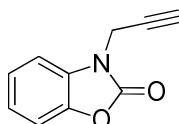
To a solution of 3,5-dimethoxybenzoic acid (300 mg, 1.7 mmol) in dry dichloromethane (10.0 mL), propargyl alcohol (77 mg, 1.4 mmol), DMAP (34 mg, 0.27 mmol) and *N,N'*-dicyclohexylcarbodiimide (424 mg, 2.1 mmol) was added at 0 °C. The reaction was stirred overnight at room temperature. The reaction mixture was filtered to remove the solid *N, N'*-dicyclohexylcarbodiimide urea waste and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with hexane/ethyl acetate (4:1) to yield a white solid. Yield = 76 %. Mass = 285 mg.

¹H NMR (500 MHz, CDCl₃) δ 7.19 (d, *J* = 2.4 Hz, 2H, ArH), 6.64 (t, *J* = 2.4 Hz, 1H, ArH), 4.89 (d, *J* = 2.5 Hz, 2H, CH₂C≡CH), 3.81 (s, 6H, OCH₃), 2.49 (t, *J* = 2.5 Hz, 1H, C≡CH); ¹³C NMR (126 MHz, CDCl₃) δ 165.8 (C=O), 160.9 (ArC), 131.4 (ArC), 107.6 (ArCH), 106.4 (ArCH), 77.9 (-C≡), 75.3 (≡CH), 55.8 (OCH₃), 52.8 (CH₂); IR (film) 2932, 2122, 1709, 1599, 1452, 1424, 1364, 1346, 1302, 1227, 1206 cm⁻¹.

ALK09 3,5-Dimethoxy-benzoic acid but-3-ynyl ester

To a solution of 3,5-dimethoxybenzoic acid (250 mg, 1.4 mmol) in dry dichloromethane (10.0 mL), butyne-3-ol (96.0 mg, 1.4 mmol), DMAP (34 mg, 0.3 mmol) and EDC.HCl (403 mg, 2.1 mmol) was added at 0 °C. The reaction was stirred overnight at room temperature. The reaction mixture was quenched with water (50 mL), extracted with dichloromethane (20 mL) and washed with water (3 x 50mL). The organic layer was dried (MgSO₄), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with hexane/ethyl acetate (9:1) to yield a white solid, yield = 74%. Mass = 243 mg.

¹H NMR (500 MHz, CDCl₃) δ 7.18 (d, *J* = 2.4 Hz, 2H, ArH), 6.64 (t, *J* = 2.4 Hz, 1H, ArH), 4.40 (t, *J* = 6.9 Hz, 2H, CH₂), 3.81 (s, 6H, OCH₃), 2.65 (td, *J* = 6.9, 2.7 Hz, 2H, CH₂C≡CH), 2.01 (t, *J* = 2.7 Hz, 1H, C≡CH); ¹³C NMR (126 MHz, CDCl₃) δ 166.3 (C=O), 160.9 (ArC), 132.0 (ArC), 107.5 (ArCH), 106.0 (ArCH), 80.2 (-C≡), 70.3 (≡CH), 62.9 (CH₂), 55.8 (OCH₃), 19.3 (CH₂); LC/MS mass calcd for C₁₃H₁₅O₄ [M + H]⁺ 235.1, found 235.2; HPLC (UV λ = 254 nm, Gradient B (section 5.2.2)) purity > 95 %; IR (thin film, NaCl plate) 3251, 2361, 1715, 1600, 1428, 1351, 1304, 1273, 1236, 1209, 1160 cm⁻¹.

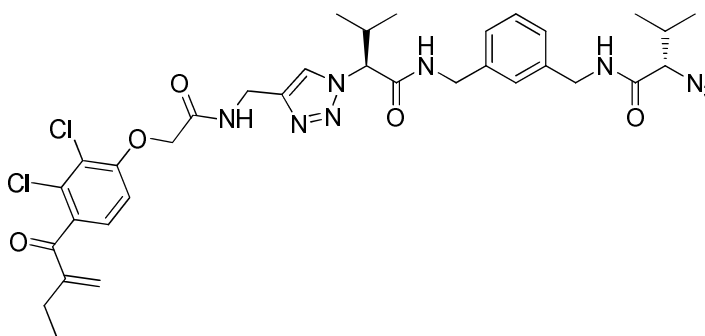
ALK10 3-Prop-2-ynyl-3H-benzoxazol-2-one

To a solution of 2-benzoxazolinone (250 mg, 1.9 mmol) in acetonitrile (10.0 mL), propargyl bromide 80 wt % in toluene (303 mg, 2.0 mmol) and Huenig's base (359 mg, 2.8 mmol) was added. The reaction mixture was stirred overnight at 60 °C. The reaction mixture was quenched with water (50 mL), extracted with dichloromethane (20 mL) and washed with water (3 x 50mL). The organic layer was dried (MgSO₄), filtered and concentrated under reduced pressure. The product was purified by flash column chromatography, eluting with hexane/ethyl acetate (1:1) to yield an orange solid, yield = 70 %. Mass = 230 mg.

^1H NMR (500 MHz, CDCl_3) δ 7.23 – 7.10 (m, 4H, ArH), 4.62 (d, $J = 2.6$ Hz, 2H, $\text{CH}_2\text{C}\equiv\text{CH}$), 2.35 (t, $J = 2.6$ Hz, 1H, $\text{C}\equiv\text{CH}$); ^{13}C NMR (126 MHz, CDCl_3) δ 153.9 (C=O), 142.9 (ArC), 130.2 (ArC), 124.2 (ArCH), 123.2 (ArCH), 110.4 (ArCH), 109.4 (ArCH), 75.7 ($-\text{C}\equiv$), 74.2 ($\equiv\text{CH}$), 32.0 (CH_2); IR (thin film, NaCl plate): 3269, 2359, 1767, 1614, 1488, 1395, 1274, 1241 cm^{-1} .

Triazole Library

TZ03a-01

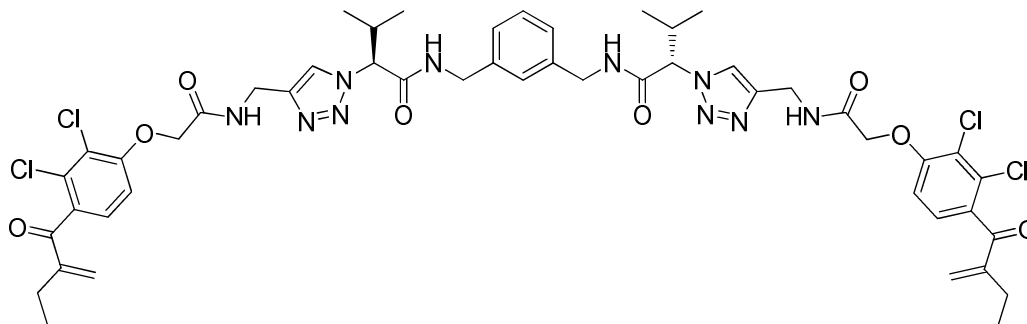


Azide **AZ01** (500 mg, 1.3 mmol) and alkyne **ALK03** (440 mg, 1.3 mmol) were dissolved in acetonitrile (30 mL). Copper (I) iodide (49 mg, 0.3 mmol) was added and the reaction was placed under a nitrogen atmosphere and stirred overnight at room temperature. On completion of the reaction, the solvent was removed under reduced pressure and the reaction mixture was absorbed onto silica prior to purification by flash column chromatography, eluting with ethyl acetate/hexane (7:3) to yield the monotriazole product **TZ03a-01** and the bistriazole product **TZ03a-01-03a** both of which were white solids. Yield of monotriazole **TZ03a-01** = 13 %. Mass = 123 mg.

^1H NMR (500 MHz, CDCl_3) δ 7.99 – 7.83 (m, 2H), 7.57 – 7.48 (m, 1H), 7.48 – 7.36 (m, 1H), 7.22 – 7.15 (m, 1H), 7.15 – 6.96 (m, 3H), 6.96 – 6.86 (m, 1H), 6.85 – 6.72 (m, 1H), 5.91 (s, 1H), 5.54 (s, 1H), 5.03 – 4.78 (m, 1H), 4.62 – 4.19 (m, 8H), 3.78 (d, $J = 4.4$ Hz, 1H), 2.57 – 2.37 (m, 3H), 2.35 – 2.23 (m, 1H), 1.10 (t, $J = 7.4$ Hz, 3H), 1.05 – 0.90 (m, 6H), 0.88 (d, $J = 6.7$ Hz, 3H), 0.75 (d, $J = 6.6$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 195.8, 169.4, 167.8, 167.2, 154.7, 150.3, 144.5, 138.6, 138.2, 134.3, 131.5, 129.3, 129.1, 127.4, 127.2, 127.0, 123.2, 122.3, 111.2, 71.0, 70.7, 68.4, 43.7, 43.3, 35.0, 32.4, 32.0, 23.6, 19.8, 19.6, 19.5, 18.9, 17.1, 12.5; HRMS (ESI+) calcd for $\text{C}_{34}\text{H}_{42}\text{O}_5\text{N}_9\text{Cl}_2$ $[\text{M} + \text{H}]^+$ 726.2680, found 726.2668; IR (film) 3289, 2967,

2938, 2102, 1659, 1586, 1537, 1468, 1439, 1385, 1371, 1292, 1242 cm^{-1} ;
MP 120 – 124 $^{\circ}\text{C}$.

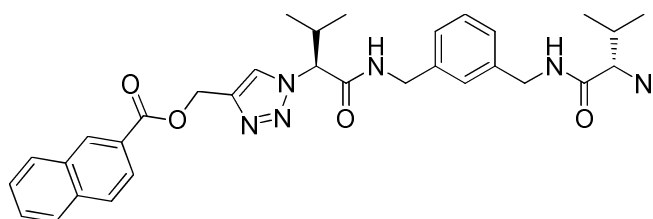
TZ03a-01-03a



Synthesis of **TZ03a-01-03a** described above in the synthesis of monotriazole **TZ03a-01**. Yield = 26 %. Mass = 361 mg.

^1H NMR (500 MHz, CDCl_3) δ 7.96 (s, 2H), 7.72 (s, 2H), 7.52 (s, 2H), 7.13 (d, $J = 8.4$ Hz, 2H), 7.08 – 7.01 (m, 1H), 6.95 (d, $J = 7.2$ Hz, 2H), 6.82 (d, $J = 8.4$ Hz, 2H), 6.69 (s, 1H), 5.91 (s, 2H), 5.55 (s, 2H), 5.00 (s, 2H), 4.53 (s, 8H), 4.38 – 4.11 (m, 4H), 2.58 – 2.46 (m, 2H), 2.46 – 2.35 (m, 4H), 1.10 (t, $J = 7.4$ Hz, 6H), 1.05 (d, $J = 6.2$ Hz, 6H), 0.74 (d, $J = 5.3$ Hz, 6H); ^{13}C NMR (126 MHz, CDCl_3) δ 195.9, 167.9, 167.4, 154.8, 150.3, 138.2, 134.3, 131.5, 129.1, 127.4, 127.1, 126.2, 123.2, 111.3, 70.9, 68.5, 43.5, 35.0, 32.1, 23.6, 19.5, 19.0, 12.6; HRMS (ESI+) calcd for $\text{C}_{50}\text{H}_{57}\text{O}_8\text{N}_{10}\text{Cl}_4$ $[\text{M} + \text{H}]^+$ 1065.3109, found 1065.3100; IR (film) 3306, 3084, 2969, 2936, 1663, 1586, 1533, 1470, 1437, 1386, 1341, 1290, 1256 cm^{-1} ; MP 78 – 82 $^{\circ}\text{C}$.

TZ04a-01

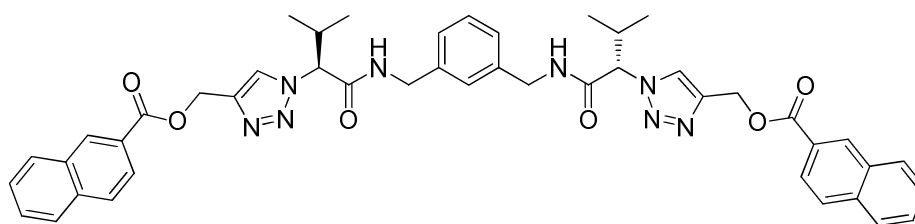


Azide **AZ01** (500 mg, 1.3 mmol) and alkyne **ALK04** (272 mg, 1.3 mmol) were dissolved in acetonitrile (75 mL). Copper (I) iodide (49 mg, 0.3 mmol) was added and the reaction was placed under a nitrogen atmosphere and stirred overnight at room temperature. On completion of the reaction, the solvent was removed under reduced pressure and the reaction mixture was absorbed onto silica prior to purification by flash column chromatography, eluting with ethyl acetate/hexane (3:2)

to yield the monotriazole product **TZ04a-01** and the bistriazole product **TZ04a-01-04a** both of which were white solids. Yield of monotriazole **TZ04a-01** = 28 %. Mass = 217 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.61 (s, 1H), 8.13 (s, 1H), 8.04 (dd, $J = 8.6, 1.7$ Hz, 1H), 7.94 (d, $J = 8.1$ Hz, 1H), 7.87 (dd, $J = 8.3, 3.6$ Hz, 2H), 7.70 (t, $J = 5.8$ Hz, 1H), 7.63 – 7.51 (m, 2H), 7.22 (t, $J = 7.8$ Hz, 1H), 7.16 – 7.09 (m, 3H), 6.92 (t, $J = 5.7$ Hz, 1H), 5.49 (q, $J = 12.8$ Hz, 2H), 5.00 (d, $J = 10.0$ Hz, 1H), 4.49 – 4.31 (m, 4H), 3.83 (d, $J = 4.4$ Hz, 1H), 2.64 – 2.53 (m, 1H), 2.42 – 2.31 (m, 1H), 1.07 (dd, $J = 6.7, 5.0$ Hz, 6H), 0.90 (d, $J = 6.7$ Hz, 3H), 0.79 (d, $J = 6.7$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 169.3, 167.8, 166.7, 143.3, 138.5, 138.2, 135.8, 132.6, 131.6, 129.5, 129.3, 128.59, 128.4, 127.9, 127.3, 127.1, 127.0, 126.9, 125.4, 124.1, 71.1, 70.7, 58.3, 43.7, 43.34, 32.5, 32.1, 19.8, 19.5, 19.0, 17.1; HRMS (ESI+) calcd for $\text{C}_{32}\text{H}_{37}\text{O}_4\text{N}_8$ $[\text{M} + \text{H}]^+$ 597.2932, found 597.2946; IR (film) 3337, 2965, 2099, 1707, 1674, 1522, 1466, 1395, 1287, 1229 cm^{-1} ; MP 141 – 145 $^\circ\text{C}$; $[\alpha]_{\text{D}} = +17.8$ ($c = 0.45$ in CHCl_3).

TZ04a-01-04a

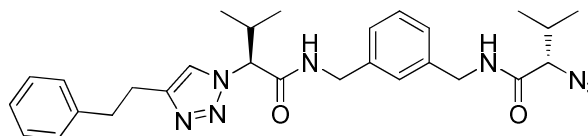


Synthesis of **TZ04a-01-04a** described above in the synthesis of monotriazole **TZ04a-01**. Yield = 28 %. Mass = 294 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.56 (s, 2H), 8.05 (s, 2H), 7.99 (dd, $J = 8.6, 1.7$ Hz, 2H), 7.92 (d, $J = 8.1$ Hz, 2H), 7.86 – 7.81 (m, 4H), 7.59 – 7.54 (m, 2H), 7.53 – 7.48 (m, 2H), 7.26 (t, $J = 6.0$ Hz, 2H), 7.10 – 7.05 (m, 1H), 6.97 (d, $J = 1.3$ Hz, 1H), 6.96 (s, 1H), 6.79 (s, 1H), 5.47 (q, $J = 12.8$ Hz, 4H), 4.96 (d, $J = 10.1$ Hz, 2H), 4.35 (dd, $J = 15.0, 6.2$ Hz, 2H), 4.22 (dd, $J = 15.0, 5.5$ Hz, 2H), 2.67 – 2.54 (m, 2H), 1.08 (d, $J = 6.7$ Hz, 6H), 0.81 (d, $J = 6.6$ Hz, 6H); ^{13}C NMR (126 MHz, CDCl_3) δ 167.8, 166.8, 143.4, 138.1, 135.8, 132.6, 131.6, 129.6, 129.2, 128.7, 128.4, 128.0, 127.1, 127.0, 127.0, 126.1, 125.4, 124.3, 71.5, 58.3, 43.6, 32.1, 19.6, 19.1; HRMS (ESI+) calcd for $\text{C}_{46}\text{H}_{47}\text{O}_6\text{N}_8$ $[\text{M} + \text{H}]^+$ 807.3613, found 807.3626; IR (film) 3395, 2969,

1721, 1705, 1688, 1655, 1528, 1458, 1354, 1279, 1233 cm^{-1} ; MP 190 – 193 $^{\circ}\text{C}$;
 $[\alpha]_{\text{D}} = +68.0$ ($c = 1.0$ in CHCl_3).

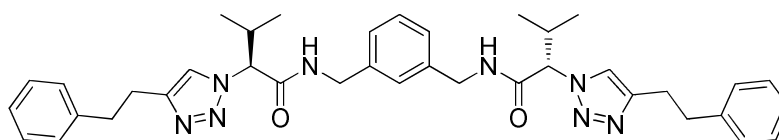
TZ18a-01



Azide **AZ01** (100 mg, 0.3 mmol) and alkyne **ALK18** (36.4 μL , 0.3 mmol) were dissolved in acetonitrile (10 mL). Copper (I) iodide (5 mg, 0.03 mmol) was added and the reaction was placed under a nitrogen atmosphere and stirred overnight at room temperature. On completion of the reaction, the solvent was removed under reduced pressure and the reaction mixture was absorbed onto silica prior to purification by flash column chromatography, eluting with ethyl acetate/hexane (3:2) to yield the monotriazole product **TZ18a-01** and the bistriazole product **TZ18a-01-18a** both of which were white solids. Yield of monotriazole **TZ18a-01** = 15 %. Mass = 23 mg.

^1H NMR (500 MHz, CDCl_3) δ 7.66 – 7.61 (m, 1H), 7.39 – 7.34 (m, 1H), 7.30 – 7.28 (m, 1H), 7.28 – 7.25 (m, 1H), 7.23 (s, 1H), 7.19 – 7.15 (m, 1H), 7.13 (d, $J = 6.9$ Hz, 2H), 7.11 – 7.05 (m, 2H), 6.81 (s, 1H), 6.68 (s, 1H), 4.54 (d, $J = 9.7$ Hz, 1H), 4.46 – 4.33 (m, 4H), 3.83 (d, $J = 4.1$ Hz, 1H), 3.08 – 2.93 (m, 4H), 2.60 – 2.48 (m, 1H), 2.44 – 2.33 (m, 1H), 1.04 (d, $J = 4.1$ Hz, 3H), 1.00 (d, $J = 6.7$ Hz, 3H), 0.88 (d, $J = 6.7$ Hz, 3H), 0.74 (d, $J = 6.6$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 169.2, 167.9, 141.2, 138.6, 138.2, 129.5, 128.7, 128.6, 128.1, 127.4, 127.3, 127.1, 126.4, 121.9, 71.6, 70.9, 43.8, 43.5, 35.7, 32.3, 32.3, 27.8, 19.9, 19.6, 18.9, 17.0; LC/MS mass calcd for $\text{C}_{28}\text{H}_{37}\text{O}_2\text{N}_8$ $[\text{M} + \text{H}]^+$ 517.3, found 517.3; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 90 %, retention time 6.43 min; IR (film) 3302, 2961, 2928, 2101, 1659, 1547, 1454, 1371, 1225 cm^{-1} .

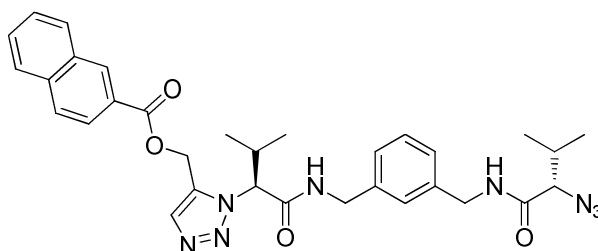
TZ18a-01-18a



Synthesis of **TZ18a-01-18a** described above in the synthesis of monotriazole **TZ18a-01**. Yield = 42 %. Mass = 81 mg.

^1H NMR (500 MHz, CDCl_3) δ 7.56 (d, $J = 19.6$ Hz, 4H), 7.33 – 7.26 (m, 4H), 7.26 – 7.20 (m, 3H), 7.17 (d, $J = 7.2$ Hz, 4H), 7.10 (d, $J = 7.5$ Hz, 2H), 6.79 (s, 1H), 4.99 (d, $J = 8.9$ Hz, 2H), 4.39 – 4.21 (m, 4H), 3.11 – 2.87 (m, 8H), 2.65 – 2.50 (m, 2H), 1.11 (d, $J = 6.6$ Hz, 6H), 0.81 (d, $J = 6.6$ Hz, 6H); ^{13}C NMR (126 MHz, CDCl_3) δ 168.2, 141.2, 138.1, 129.3, 128.7, 128.6, 127.3, 126.8, 126.3, 71.0, 43.6, 35.6, 32.1, 27.8, 19.6, 19.0; LC/MS mass calcd for $\text{C}_{38}\text{H}_{47}\text{O}_2\text{N}_8$ $[\text{M} + \text{H}]^+$ 647.4, found 647.4; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 90 %, retention time 6.88 min.

TZ04s-01

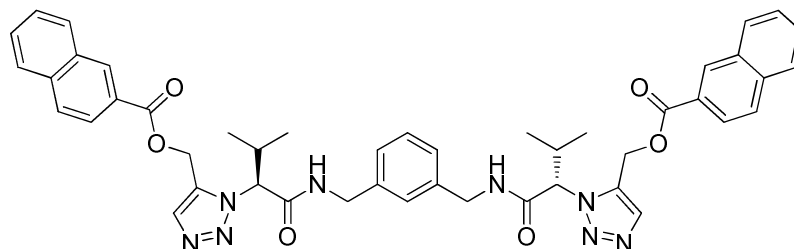


Alkyne **ALK04** (82 mg, 0.4 mmol) was dissolved in dioxane (2.5 mL) followed by azide **AZ01** (150 mg, 0.4 mmol) and this mixture was added to a solution of $\text{Cp}^*\text{RuCl}(\text{PPh}_3)_2$ (6 mg, 8.0 μmol) in dioxane (2.5 mL). The reaction was placed under a nitrogen atmosphere and heated at 60 $^\circ\text{C}$ overnight. On completion of the reaction the reaction mixture was absorbed onto silica for purification by flash column chromatography, eluting with ethyl acetate/hexane (3:2) to yield the monotriazole product **TZ04s-01** and the bistriazole product **TZ04s-01-04s** both of which were white solids. Yield of monotriazole **TZ04s-01** = 10 %. Mass = 24 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.55 (s, 1H), 7.95 (dd, $J = 8.6, 1.6$ Hz, 1H), 7.90 (d, $J = 8.0$ Hz, 1H), 7.87 (s, 1H), 7.86 – 7.82 (m, 2H), 7.61 – 7.55 (m, 1H), 7.55 – 7.49 (m, 1H), 7.27 – 7.21 (m, 1H), 7.18 – 7.12 (m, 1H), 7.08 (d, $J = 7.7$ Hz, 1H), 6.99 (d, $J = 7.6$ Hz, 1H), 6.95 (s, 1H), 6.64 (s, 1H), 5.58 (d, $J = 13.7$ Hz, 1H), 5.42 (d, $J = 13.7$ Hz, 1H), 4.85 (d, $J = 10.1$ Hz, 1H), 4.49 – 4.24 (m, 4H), 3.84 (d, $J = 4.1$ Hz, 1H), 2.90 – 2.78 (m, 1H), 2.43 – 2.29 (m, 1H), 1.14 (d, $J = 6.7$ Hz, 3H), 1.05 (d, $J = 6.9$ Hz, 3H), 0.88 (d, $J = 6.7$ Hz, 3H), 0.83 (d, $J = 6.6$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 169.1, 167.4, 166.1, 138.5, 138.3, 136.0, 135.3, 133.8, 132.6, 131.9, 129.7, 129.4, 128.9, 128.6, 128.0, 127.2, 127.1, 127.0, 126.8, 126.1, 125.2,

70.8, 70.6, 54.0, 43.5, 43.4, 32.7, 32.2, 19.9, 19.6, 19.5, 17.0; LC/MS mass calcd for $C_{32}H_{37}O_4N_8$ $[M + H]^+$ 597.3, found 597.3; HPLC (UV $\lambda = 254$ nm, Gradient C (section 5.2.2)) purity > 90 %, retention time 10.12 min; IR (film) 3281, 2967, 2099, 1721, 1647, 1543, 1466, 1451, 1371, 1354, 1279, 1225 cm^{-1} .

TZ04s-01-04s



Synthesis of **TZ04s-01-04s** described above in the synthesis of monotriazole **TZ04s-01**. Yield = 17 %. Mass = 55 mg.

1H NMR (500 MHz, $CDCl_3$) δ 8.54 (s, 2H), 7.96 – 7.80 (m, 10H), 7.61 – 7.47 (m, 4H), 7.18 (t, $J = 5.6$ Hz, 2H), 7.07 – 7.02 (m, 1H), 6.95 – 6.89 (m, 2H), 6.80 (s, 1H), 5.57 (d, $J = 13.7$ Hz, 2H), 5.45 (d, $J = 13.6$ Hz, 2H), 4.85 (d, $J = 10.2$ Hz, 2H), 4.36 (dd, $J = 15.0, 6.1$ Hz, 2H), 4.24 (dd, $J = 15.0, 5.7$ Hz, 2H), 2.91 – 2.75 (m, 2H), 1.25 (s, 6H), 1.09 (d, $J = 6.7$ Hz, 3H), 0.83 (d, $J = 6.6$ Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.4, 166.1, 138.3, 136.0, 135.3, 133.8, 132.6, 131.9, 129.7, 129.3, 128.9, 128.6, 128.0, 127.1, 126.7, 126.5, 126.2, 125.2, 70.7, 54.1, 43.5, 32.6, 19.6, 19.5.

General procedure C - for copper (I) catalysed 1,4 triazole synthesis

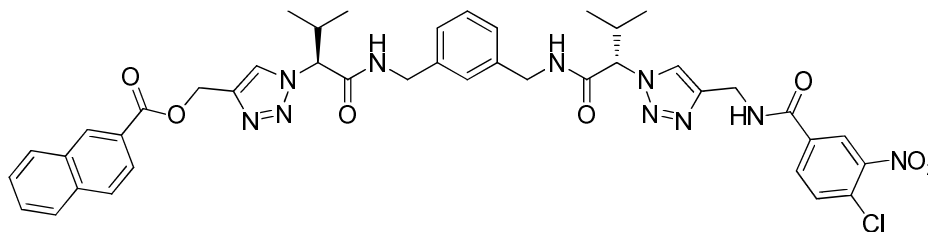
Azide **TZ04a-01** (50 mg, 83.8 μ mol) and alkyne (1 eq, 83.8 μ mol) were dissolved in acetonitrile (3 mL). Copper (I) iodide (3 mg, 16.8 μ mol) was added and the reaction was placed under a nitrogen atmosphere and stirred overnight at room temperature. On completion of the reaction, the solvent was removed under reduced pressure and the reaction mixture was absorbed onto silica prior to purification by flash column chromatography.

General procedure D - for ruthenium catalysed 1,5 triazole synthesis

Alkyne (1 eq, 83.8 μ mol) was dissolved in dioxane (0.5 mL) followed by azide **TZ04a-01** (50 mg, 83.8 μ mol) and this mixture was added to a solution of $Cp^*RuCl(PPh_3)_2$ (7 mg, 8.4 μ mol) in dioxane (2 mL). The reaction was placed under a nitrogen atmosphere and heated at 60 $^{\circ}C$ overnight. On completion of the reaction

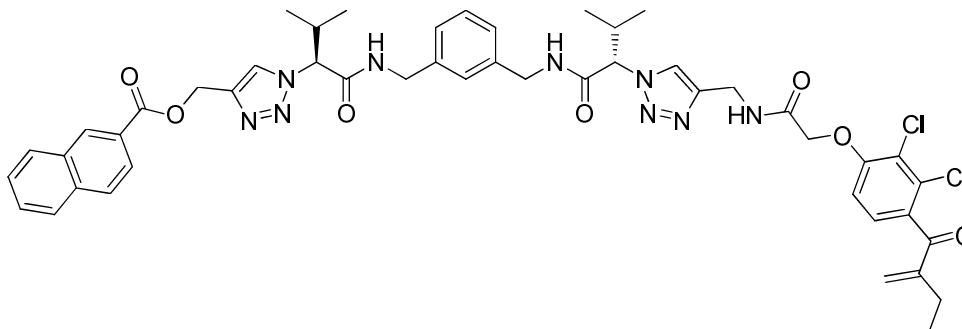
the reaction mixture was absorbed onto silica for purification by flash column chromatography.

TZ04a-01-01a



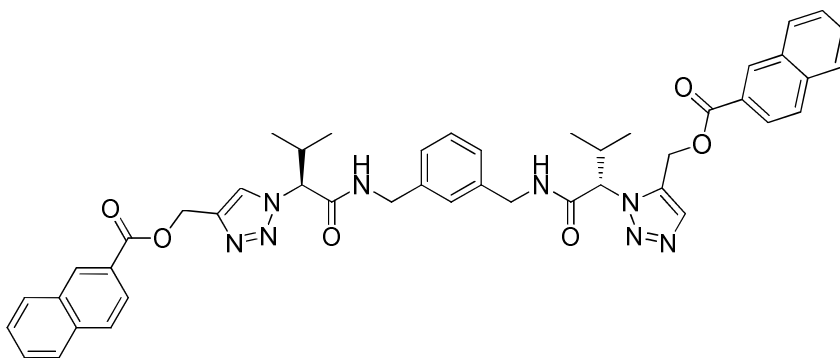
TZ04a-01-01a was synthesized following general procedure C, starting from azide **TZ04a-01** and alkyne **ALK01**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 39 %. Mass = 27 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.52 (s, 1H), 8.45 (s, 1H), 8.36 (d, $J = 1.8$ Hz, 1H), 8.23 (s, 1H), 8.05 (s, 1H), 7.96 (ddd, $J = 15.1, 8.5, 1.6$ Hz, 2H), 7.87 (d, $J = 7.9$ Hz, 2H), 7.82 (t, $J = 7.4$ Hz, 2H), 7.78 (s, 1H), 7.60 – 7.44 (m, 3H), 7.01 – 6.90 (m, 2H), 6.83 (d, $J = 7.4$ Hz, 1H), 6.70 (s, 1H), 5.53 – 5.37 (m, 2H), 5.14 (d, $J = 10.1$ Hz, 1H), 4.96 (d, $J = 10.1$ Hz, 1H), 4.68 – 4.49 (m, 2H), 4.33 – 4.04 (m, 4H), 2.63 – 2.42 (m, 2H), 1.02 (dd, $J = 9.2, 6.8$ Hz, 6H), 0.73 (dd, $J = 24.1, 6.6$ Hz, 6H); ^{13}C NMR (126 MHz, CDCl_3) δ 168.0, 168.0, 167.0, 164.5, 147.8, 144.7, 143.4, 138.1, 138.0, 135.8, 133.8, 132.6, 132.2, 132.1, 131.6, 130.1, 129.6, 129.0, 128.7, 128.5, 128.0, 127.1, 127.0, 126.9, 126.8, 125.9, 125.2, 124.8, 124.5, 123.4, 71.1, 70.9, 58.3, 43.4, 35.7, 32.2, 32.0, 19.5, 19.4, 19.1, 19.0; LC/MS mass calcd for $\text{C}_{42}\text{H}_{44}\text{O}_7\text{N}_{10}\text{Cl}$ $[\text{M} + \text{H}]^+$ 835.3, found 835.2; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 95 %, retention time 7.02 min.

TZ04a-01-03a/TZ03a-01-04a

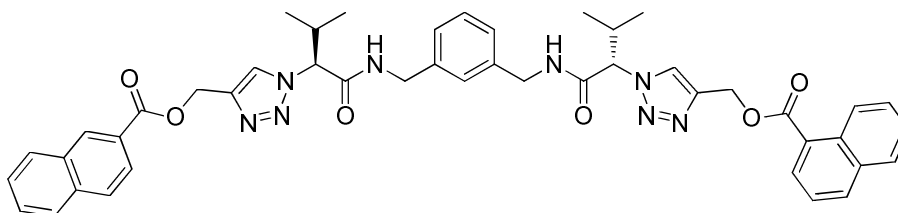
TZ04a-01-03a was synthesized following general procedure C, starting from azide **TZ04a-01** and alkyne **ALK03**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 80 %. Mass = 63 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.55 (s, 1H), 8.12 (s, 1H), 7.97 (dd, $J = 8.6, 1.6$ Hz, 1H), 7.94 (s, 1H), 7.88 (d, $J = 8.1$ Hz, 1H), 7.82 (t, $J = 7.6$ Hz, 2H), 7.73 (s, 1H), 7.62 – 7.53 (m, 2H), 7.50 (t, $J = 7.0$ Hz, 1H), 7.46 (s, 1H), 7.13 (d, $J = 8.5$ Hz, 1H), 7.06 – 7.00 (m, 1H), 6.96 (d, $J = 7.5$ Hz, 1H), 6.90 (d, $J = 7.4$ Hz, 1H), 6.81 (d, $J = 8.5$ Hz, 1H), 6.75 (s, 1H), 5.90 (s, 1H), 5.54 (s, 1H), 5.44 (q, $J = 12.8$ Hz, 2H), 5.11 – 4.94 (m, 2H), 4.62 – 4.43 (m, 4H), 4.34 – 4.10 (m, 4H), 2.64 – 2.47 (m, 2H), 2.43 (q, $J = 7.3$ Hz, 2H), 1.11 (t, $J = 7.4$ Hz, 3H), 1.08 – 0.98 (m, 6H), 0.77 (dd, $J = 14.4, 6.6$ Hz, 6H); ^{13}C NMR (126 MHz, CDCl_3) δ 195.8, 167.9, 167.9, 167.4, 166.8, 154.7, 150.3, 138.2, 138.1, 135.8, 134.3, 132.6, 131.6, 129.6, 129.2, 129.1, 128.6, 128.4, 128.0, 127.4, 127.1, 127.0, 126.9, 126.9, 126.4, 125.4, 124.0, 123.2, 111.3, 71.0, 68.4, 58.3, 43.5, 43.5, 35.0, 32.1, 32.1, 23.6, 19.5, 19.1, 19.0, 12.6; HRMS (ESI+) calcd for $\text{C}_{48}\text{H}_{52}\text{O}_7\text{N}_9\text{Cl}_2$ $[\text{M} + \text{H}]^+$ 936.3361, found 936.3375; IR (film) 3310, 2969, 2940, 1663, 1586, 1532, 1468, 1437, 1387, 1354, 1283, 1225 cm^{-1} ; MP 96 – 99 °C.

TZ04a-01-04s

TZ04a-01-04s was synthesized following general procedure D, starting from azide **TZ04a-01** and alkyne **ALK04**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 12 %. Mass = 8 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.59 (s, 1H), 8.56 (s, 1H), 8.07 – 8.01 (m, 2H), 7.96 – 7.88 (m, 4H), 7.87 – 7.81 (m, 4H), 7.62 – 7.54 (m, 2H), 7.54 – 7.49 (m, 2H), 7.17 (s, 1H), 7.10 – 7.02 (m, 2H), 7.01 – 6.96 (m, 1H), 6.94 (d, $J = 7.5$ Hz, 1H), 6.84 (s, 1H), 5.63 – 5.36 (m, 4H), 4.84 (dd, $J = 10.1, 3.8$ Hz, 2H), 4.43 – 4.18 (m, 4H), 2.92 – 2.79 (m, 1H), 2.65 – 2.50 (m, 1H), 1.19 (d, $J = 6.7$ Hz, 3H), 1.06 (d, $J = 6.7$ Hz, 3H), 0.86 (d, $J = 6.5$ Hz, 3H), 0.77 (d, $J = 6.7$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 167.7, 167.3, 166.7, 166.3, 138.4, 138.0, 136.0, 135.8, 132.6, 132.0, 131.6, 129.7, 129.6, 129.3, 129.0, 128.7, 128.6, 128.4, 128.0, 127.2, 127.1, 126.9, 126.4, 126.1, 125.5, 125.2, 124.1, 71.3, 70.7, 58.4, 54.1, 43.7, 43.4, 32.5, 32.4, 19.6, 19.6, 19.5, 19.0; HRMS (ESI+) calcd for $\text{C}_{46}\text{H}_{47}\text{O}_6\text{N}_8$ $[\text{M} + \text{H}]^+$ 807.3613, found 807.3612; IR (film) 3316, 3069, 2928, 1715, 1672, 1651, 1526, 1466, 1447, 1354, 1279, 1227, 1194, 1128, 1088 cm^{-1} ; MP 94 – 98 $^\circ\text{C}$.

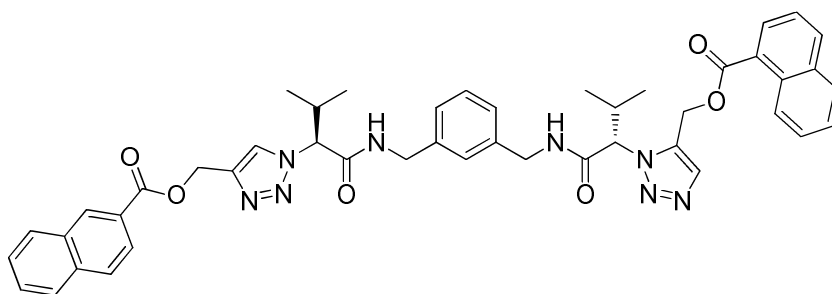
TZ04a-01-07a

TZ04a-01-07a was synthesized following general procedure C, starting from azide **TZ04a-01** and alkyne **ALK07**. The product was purified by flash column

chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 50 %. Mass = 34 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.84 (d, $J = 8.7$ Hz, 1H), 8.55 (s, 1H), 8.15 – 8.07 (m, 3H), 8.00 – 7.94 (m, 2H), 7.88 (d, $J = 7.8$ Hz, 1H), 7.82 (dd, $J = 11.6, 6.2$ Hz, 3H), 7.59 – 7.46 (m, 6H), 7.44 – 7.37 (m, 1H), 7.06 – 6.99 (m, 1H), 6.96 – 6.88 (m, 2H), 6.79 (s, 1H), 5.52 – 5.36 (m, 4H), 5.05 (dd, $J = 10.1, 2.6$ Hz, 2H), 4.33 – 4.11 (m, 4H), 2.65 – 2.49 (m, 2H), 1.07 (dd, $J = 6.6, 2.6$ Hz, 6H), 0.78 (d, $J = 6.6$ Hz, 6H); ^{13}C NMR (126 MHz, CDCl_3) δ 167.9, 167.4, 166.8, 143.4, 138.1, 135.8, 134.0, 133.9, 132.6, 131.6, 131.5, 130.8, 129.6, 129.2, 128.8, 128.6, 128.4, 128.1, 127.9, 127.1, 127.1, 127.0, 126.9, 126.5, 126.5, 126.4, 125.8, 125.4, 124.7, 124.1, 124.1, 71.1, 71.1, 58.3, 43.6, 43.5, 32.1, 19.5, 19.0, 19.0, 14.4; LC/MS mass calcd for $\text{C}_{46}\text{H}_{47}\text{O}_6\text{N}_8$ $[\text{M} + \text{H}]^+$ 807.4, found 807.3; HPLC (UV $\lambda = 254$ nm, Gradient C (section 5.2.2)) purity > 95 %, retention time 11.50 min.

TZ04a-01-07s

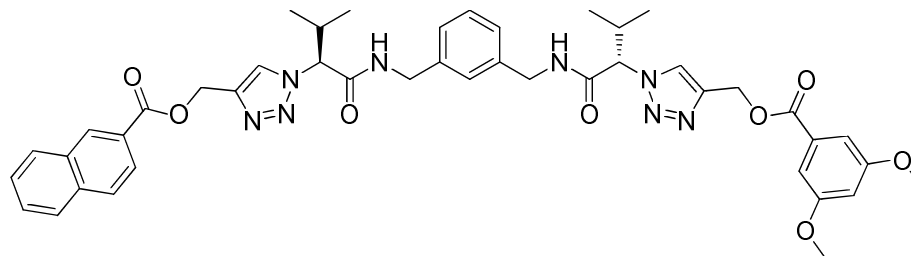


TZ04a-01-07s was synthesized following general procedure D, starting from azide **TZ04a-01** and alkyne **ALK07**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 9 %. Mass = 5 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.85 (d, $J = 8.4$ Hz, 1H), 8.60 (s, 1H), 8.13 (d, $J = 7.3$ Hz, 1H), 8.06 – 7.98 (m, 3H), 7.97 – 7.89 (m, 2H), 7.85 (t, $J = 7.5$ Hz, 3H), 7.62 – 7.48 (m, 4H), 7.46 – 7.40 (m, 1H), 7.15 (s, 1H), 7.09 (t, $J = 7.6$ Hz, 1H), 7.04 – 6.93 (m, 3H), 6.80 (s, 1H), 5.60 – 5.44 (m, 4H), 4.85 (t, $J = 10.1$ Hz, 2H), 4.45 – 4.19 (m, 4H), 2.92 – 2.78 (m, 1H), 2.68 – 2.47 (m, 1H), 1.16 (d, $J = 6.7$ Hz, 3H), 1.03 (d, $J = 6.7$ Hz, 3H), 0.84 (d, $J = 6.6$ Hz, 3H), 0.77 (d, $J = 6.7$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 167.7, 167.3, 166.7, 166.7, 143.4, 138.4, 138.0, 135.9, 135.4, 134.6, 134.0, 133.9, 132.7, 131.6, 131.6, 131.2, 129.6, 129.3, 129.0,

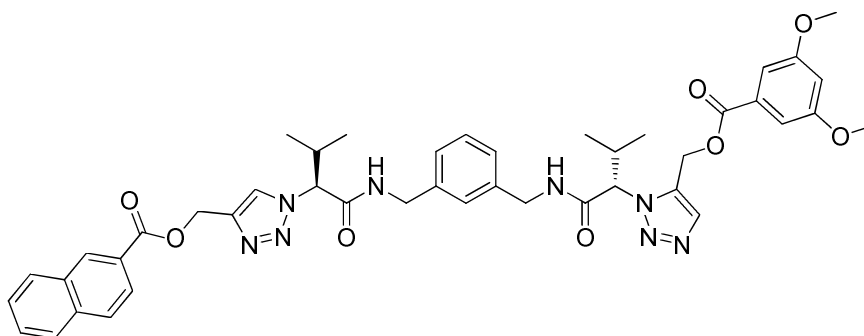
128.6, 128.5, 128.4, 128.0, 127.2, 127.1, 127.0, 126.9, 126.7, 126.4, 125.6, 125.5, 125.3, 124.7, 124.1, 71.3, 70.5, 58.4, 54.2, 43.7, 43.4, 32.4, 32.3, 19.6, 19.6, 19.5, 19.0; LC/MS mass calcd for $C_{46}H_{47}O_6N_8$ $[M + H]^+$ 807.4, found 807.3; HPLC (UV λ = 254 nm, Gradient C (section 5.2.2)) purity > 95 %, retention time 11.42 min.

TZ04a-01-08a



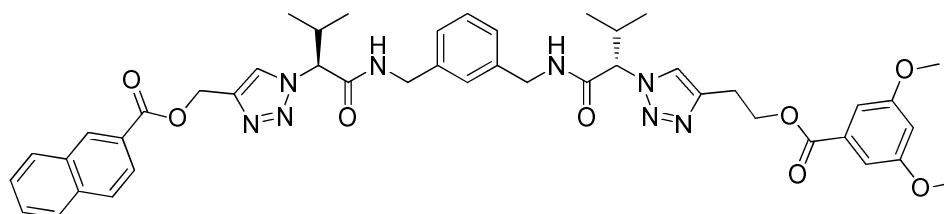
TZ04a-01-08a was synthesized following general procedure C, starting from azide **TZ04a-01** and alkyne **ALK08**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 57 %. Mass = 39 mg.

1H NMR (500 MHz, $CDCl_3$) δ 8.54 (s, 1H), 8.12 (s, 1H), 8.06 (s, 1H), 7.97 (dd, J = 8.6, 1.5 Hz, 1H), 7.88 (d, J = 8.1 Hz, 1H), 7.85 (t, J = 7.3 Hz, 2H), 7.64 – 7.46 (m, 4H), 7.11 (d, J = 2.3 Hz, 2H), 7.06 (t, J = 7.6 Hz, 1H), 6.97 (t, J = 7.5 Hz, 2H), 6.82 (s, 1H), 6.62 (t, J = 2.3 Hz, 1H), 5.47 – 5.30 (m, 4H), 5.09 (t, J = 10.9 Hz, 2H), 4.34 – 4.16 (m, 4H), 3.79 (s, 6H), 2.64 – 2.46 (m, 2H), 1.08 – 1.00 (m, 6H), 0.80 – 0.73 (m, 6H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.9, 167.8, 166.8, 166.4, 160.8, 138.1, 135.8, 132.6, 131.6, 129.6, 129.2, 128.6, 128.4, 127.9, 127.2, 126.9, 126.9, 126.6, 125.4, 124.2, 107.5, 106.1, 71.1, 58.3, 55.8, 53.6, 43.6, 32.2, 19.5, 19.0; LC/MS mass calcd for $C_{44}H_{49}O_8N_8$ $[M + H]^+$ 817.4, found 817.3; HPLC (UV λ = 254 nm, Gradient B (section 5.2.2)) purity > 95 %, retention time 7.26 min.

TZ04a-01-08s

TZ04a-01-08s was synthesized following general procedure D, starting from azide **TZ04a-01** and alkyne **ALK08**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 15 %. Mass = 10 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.59 (s, 1H), 8.07 – 8.00 (m, 2H), 7.92 (d, $J = 8.1$ Hz, 1H), 7.89 – 7.83 (m, 3H), 7.57 (t, $J = 6.9$ Hz, 1H), 7.52 (t, $J = 7.0$ Hz, 1H), 7.19 – 7.06 (m, 5H), 7.03 (d, $J = 7.6$ Hz, 1H), 6.99 – 6.91 (m, 1H), 6.82 (s, 1H), 6.62 (t, $J = 2.3$ Hz, 1H), 5.51 (d, $J = 4.9$ Hz, 2H), 5.45 (d, $J = 13.7$ Hz, 1H), 5.36 (d, $J = 13.7$ Hz, 1H), 4.89 (d, $J = 10.0$ Hz, 1H), 4.80 (d, $J = 10.2$ Hz, 1H), 4.43 – 4.21 (m, 4H), 3.77 (s, 6H), 2.89 – 2.76 (m, 1H), 2.63 – 2.50 (m, 1H), 1.14 (d, $J = 6.7$ Hz, 3H), 1.05 (d, $J = 6.7$ Hz, 3H), 0.82 (d, $J = 6.5$ Hz, 3H), 0.77 (d, $J = 6.6$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 167.7, 167.3, 166.7, 165.9, 161.0, 138.4, 138.0, 135.8, 132.6, 131.6, 130.8, 129.6, 129.3, 128.6, 128.4, 128.0, 127.3, 127.1, 126.9, 126.9, 126.5, 125.5, 124.0, 107.5, 106.8, 71.3, 70.6, 58.4, 55.9, 54.1, 43.8, 43.4, 32.4, 32.3, 30.5, 29.9, 19.6, 19.5, 19.0; LC/MS mass calcd for $\text{C}_{44}\text{H}_{49}\text{O}_8\text{N}_8$ $[\text{M} + \text{H}]^+$ 817.4, found 817.3; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 95 %, retention time 7.23 min.

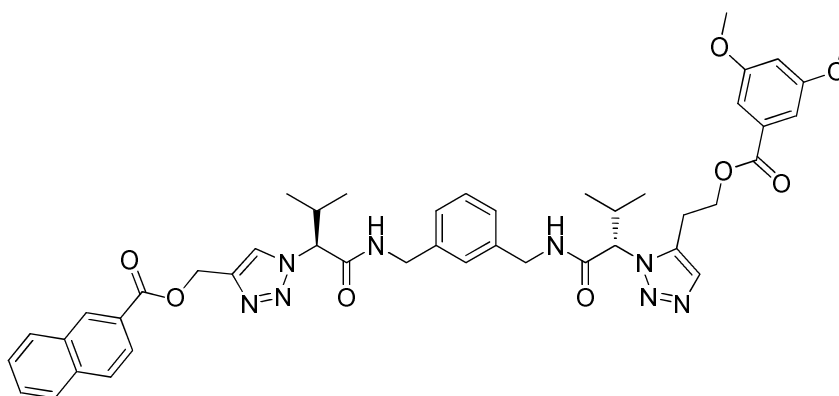
TZ04a-01-09a

TZ04a-01-09a was synthesized following general procedure C, starting from azide **TZ04a-01** and alkyne **ALK09**. The product was purified by flash column

chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 56 %. Mass = 39 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.55 (s, 1H), 8.11 (s, 1H), 8.00 – 7.95 (m, 1H), 7.92 (d, $J = 8.1$ Hz, 1H), 7.86 – 7.80 (m, 2H), 7.79 (s, 1H), 7.64 – 7.45 (m, 4H), 7.12 – 7.05 (m, 3H), 7.02 – 6.94 (m, 2H), 6.81 (s, 1H), 6.61 (t, $J = 2.3$ Hz, 1H), 5.42 (q, $J = 12.8$ Hz, 2H), 5.04 (d, $J = 10.1$ Hz, 1H), 4.99 (d, $J = 10.1$ Hz, 1H), 4.53 (t, $J = 6.8$ Hz, 2H), 4.34 – 4.16 (m, 4H), 3.78 (s, 6H), 3.14 (t, $J = 6.7$ Hz, 2H), 2.65 – 2.43 (m, 2H), 1.08 – 1.00 (m, 6H), 0.80 (d, $J = 6.6$ Hz, 3H), 0.73 (d, $J = 6.6$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 168.1, 167.9, 166.8, 166.4, 160.8, 138.1, 138.1, 135.8, 132.6, 132.0, 131.6, 129.6, 129.2, 128.6, 128.4, 128.0, 127.1, 127.1, 127.0, 126.9, 126.6, 125.4, 124.1, 122.1, 107.4, 105.9, 71.1, 71.0, 63.9, 58.3, 55.8, 43.6, 43.5, 32.2, 32.1, 25.7, 19.5, 19.0, 18.9; LC/MS mass calcd for $\text{C}_{45}\text{H}_{51}\text{O}_8\text{N}_8$ $[\text{M} + \text{H}]^+$ 831.4, found 831.3; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 95 %, retention time 7.24 min.

TZ04a-01-09s

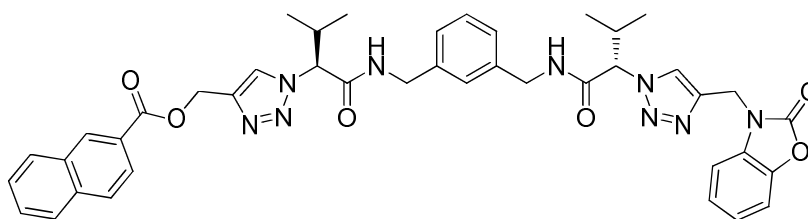


TZ04a-01-09s was synthesized following general procedure D, starting from azide **TZ04a-01** and alkyne **ALK09**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 22 %. Mass = 15 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.60 (s, 1H), 8.09 – 8.01 (m, 2H), 7.92 (d, $J = 8.1$ Hz, 1H), 7.86 (s, 1H), 7.83 (s, 1H), 7.60 – 7.49 (m, 4H), 7.30 (s, 1H), 7.12 (t, $J = 7.6$ Hz, 1H), 7.10 – 7.04 (m, 3H), 7.00 (d, $J = 7.6$ Hz, 1H), 6.88 (s, 1H), 6.59 (t, $J = 2.3$ Hz, 1H), 5.57 – 5.46 (m, 2H), 4.92 (d, $J = 10.0$ Hz, 1H), 4.58 (d, $J = 10.1$ Hz, 1H), 4.53 (t, $J = 6.6$ Hz, 2H), 4.44 – 4.20 (m, 4H), 3.78 (s, 6H), 3.26 – 3.07 (m, 2H),

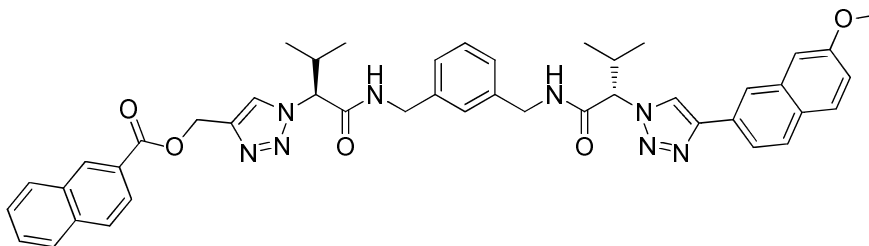
2.84 – 2.71 (m, 1H), 2.64 – 2.49 (m, 1H), 1.13 (d, $J = 6.7$ Hz, 3H), 1.05 (d, $J = 6.7$ Hz, 3H), 0.82 (d, $J = 6.5$ Hz, 3H), 0.76 (d, $J = 6.7$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 167.8, 167.5, 166.7, 166.3, 160.9, 138.4, 138.1, 135.8, 133.1, 132.6, 132.3, 132.2, 132.2, 131.6, 131.4, 129.6, 129.3, 128.6, 128.4, 128.0, 127.3, 127.2, 127.0, 126.9, 126.8, 125.5, 124.0, 107.5, 106.2, 71.2, 70.2, 67.3, 62.1, 58.4, 55.8, 43.8, 43.5, 32.5, 32.5, 23.4, 19.5, 19.0; LC/MS mass calcd for $\text{C}_{45}\text{H}_{51}\text{O}_8\text{N}_8$ $[\text{M} + \text{H}]^+$ 831.4, found 831.3; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 90 %, retention time 7.21 min.

TZ04a-01-10a



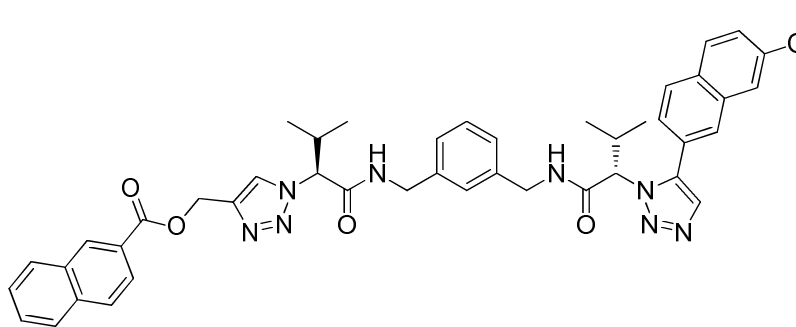
TZ04a-01-10a was synthesized following general procedure C, starting from azide **TZ04a-01** and alkyne **ALK10**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 74 %. Mass = 48 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.59 (s, 1H), 8.13 (s, 1H), 8.07 (s, 1H), 8.02 (dd, $J = 8.6, 1.6$ Hz, 1H), 7.94 (d, $J = 8.1$ Hz, 1H), 7.88 (t, $J = 7.9$ Hz, 2H), 7.66 – 7.58 (m, 1H), 7.58 – 7.52 (m, 2H), 7.38 (t, $J = 5.7$ Hz, 1H), 7.19 – 7.06 (m, 5H), 7.02 (d, $J = 7.7$ Hz, 1H), 6.96 (d, $J = 7.5$ Hz, 1H), 6.79 (s, 1H), 5.49 (q, $J = 12.9$ Hz, 2H), 5.08 (d, $J = 5.2$ Hz, 1H), 5.06 – 4.98 (m, 2H), 4.42 – 4.28 (m, 2H), 4.21 (td, $J = 14.3, 5.5$ Hz, 1H), 2.72 – 2.53 (m, 2H), 1.78 (s, 2H), 1.12 (d, $J = 6.6$ Hz, 3H), 1.07 (d, $J = 6.7$ Hz, 3H), 0.85 (d, $J = 6.6$ Hz, 3H), 0.78 (d, $J = 6.6$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 167.9, 167.7, 166.8, 154.6, 142.7, 138.2, 138.1, 135.8, 132.6, 131.6, 130.7, 129.6, 129.1, 128.7, 128.4, 128.0, 127.1, 127.1, 127.0, 127.0, 126.1, 125.4, 124.3, 124.1, 123.4, 123.0, 110.3, 109.4, 71.3, 71.2, 58.3, 43.5, 43.5, 37.7, 32.0, 32.0, 19.5, 19.5, 19.1, 19.0; LC/MS mass calcd for $\text{C}_{42}\text{H}_{44}\text{O}_6\text{N}_9$ $[\text{M} + \text{H}]^+$ 770.3, found 770.4; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 95 %, retention time 6.90 min.

TZ04a-01-11a

TZ04a-01-11a was synthesized following general procedure C, starting from azide **TZ04a-01** and alkyne **ALK11**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 76 %. Mass = 50 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.56 (s, 1H), 8.20 (s, 1H), 8.18 (s, 1H), 8.07 (s, 1H), 8.00 (dd, $J = 8.6, 1.6$ Hz, 1H), 7.87 (d, $J = 8.2$ Hz, 1H), 7.86 – 7.80 (m, 3H), 7.74 (t, $J = 8.4$ Hz, 2H), 7.59 – 7.54 (m, 1H), 7.34 (d, $J = 5.6$ Hz, 1H), 7.31 (s, 1H), 7.15 – 7.04 (m, 4H), 7.00 – 6.92 (m, 2H), 6.74 (s, 1H), 5.54 (q, $J = 12.8$ Hz, 2H), 5.04 (dd, $J = 26.4, 10.2$ Hz, 2H), 4.40 (ddd, $J = 28.8, 15.0, 6.3$ Hz, 2H), 4.23 – 4.08 (m, 2H), 3.90 (s, 3H), 2.74 – 2.50 (m, 2H), 1.16 (d, $J = 6.6$ Hz, 3H), 1.08 (d, $J = 6.6$ Hz, 3H), 0.91 (d, $J = 6.6$ Hz, 3H), 0.80 (d, $J = 6.6$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 168.1, 167.8, 166.8, 158.3, 148.5, 143.4, 138.1, 138.0, 135.8, 134.7, 132.6, 131.6, 129.9, 129.6, 129.2, 129.1, 128.6, 128.4, 128.0, 127.7, 127.2, 127.1, 126.9, 126.1, 125.6, 125.4, 124.6, 124.4, 124.2, 119.9, 119.6, 106.0, 71.4, 71.4, 58.4, 55.5, 43.6, 43.6, 32.0, 31.9, 30.5, 29.9, 19.6, 19.5, 19.1; LC/MS mass calcd for $\text{C}_{45}\text{H}_{47}\text{O}_5\text{N}_8$ $[\text{M} + \text{H}]^+$ 779.4, found 779.3; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 95 %, retention time 7.53 min.

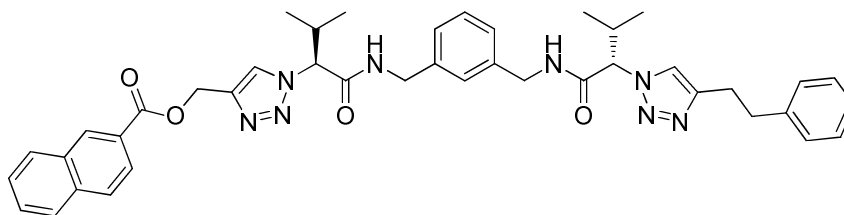
TZ04a-01-11s

TZ04a-01-11s was synthesized following general procedure D, starting from

azide **TZ04a-01** and alkyne **ALK11**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 35 %. Mass = 23 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.59 (s, 1H), 8.07 (s, 1H), 8.02 (dd, $J = 8.6, 1.7$ Hz, 1H), 7.94 (d, $J = 8.1$ Hz, 1H), 7.87 – 7.81 (m, 4H), 7.80 – 7.75 (m, 2H), 7.56 (ddd, $J = 8.2, 6.9, 1.3$ Hz, 1H), 7.53 – 7.47 (m, 2H), 7.38 (dd, $J = 8.4, 1.7$ Hz, 1H), 7.31 (t, $J = 5.6$ Hz, 1H), 7.23 – 7.18 (m, 2H), 7.16 (d, $J = 2.4$ Hz, 1H), 7.14 – 7.08 (m, 2H), 7.03 (s, 1H), 5.51 (d, $J = 3.5$ Hz, 2H), 4.95 (d, $J = 10.0$ Hz, 1H), 4.70 (d, $J = 10.1$ Hz, 1H), 4.52 – 4.22 (m, 4H), 3.93 (s, 3H), 2.80 – 2.66 (m, 1H), 2.64 – 2.49 (m, 1H), 1.07 (d, $J = 6.7$ Hz, 3H), 1.00 (d, $J = 6.8$ Hz, 3H), 0.79 (d, $J = 6.7$ Hz, 3H), 0.66 (d, $J = 6.5$ Hz, 3H); ^{13}C NMR (126 MHz, CDCl_3) δ 168.1, 167.8, 166.7, 159.1, 143.3, 140.7, 138.5, 138.1, 135.8, 135.2, 133.3, 132.6, 131.6, 130.2, 129.6, 129.4, 129.2, 128.7, 128.6, 128.4, 128.3, 128.0, 127.3, 127.1, 127.1, 127.0, 126.9, 126.7, 125.5, 124.0, 120.8, 120.4, 105.9, 71.2, 69.9, 58.4, 55.7, 43.9, 43.6, 33.5, 32.6, 19.5, 19.4, 19.3, 19.0; LC/MS mass calcd for $\text{C}_{45}\text{H}_{47}\text{O}_5\text{N}_8$ $[\text{M} + \text{H}]^+$ 779.4, found 779.3; HPLC (UV $\lambda = 254$ nm, Gradient B (section 5.2.2)) purity > 95 %, retention time 7.43 min.

TZ04a-01-18a



TZ04a-01-18a was synthesized following general procedure C, starting from azide **TZ04a-01** and alkyne **ALK18**. The product was purified by flash column chromatography, eluting with ethyl acetate/hexane (8:2) to yield a white solid. Yield = 63 %. Mass = 38 mg.

^1H NMR (500 MHz, CDCl_3) δ 8.56 (s, 1H), 8.10 (s, 1H), 7.99 (d, $J = 8.4$, 1H), 7.91 (d, $J = 8.2$ Hz, 1H), 7.83 (t, $J = 7.4$ Hz, 2H), 7.56 (t, $J = 7.1$ Hz, 1H), 7.53 (t, $J = 7.3$ Hz, 1H), 7.38 (s, 3H), 7.23 – 7.17 (m, 2H), 7.17 – 7.06 (m, 4H), 7.05 – 6.96 (m, 2H), 6.91 (s, 1H), 5.55 – 5.43 (m, 2H), 5.21 (s, 2H), 4.40 – 4.14 (m, 4H), 3.14 – 2.87 (m, 4H), 2.66 – 2.35 (m, 2H), 1.10 (dd, $J = 18.2, 6.5$ Hz, 6H), 0.83 – 0.61 (m, 6H); ^{13}C NMR (126 MHz, CDCl_3) δ 167.7, 166.8, 138.1, 135.8, 132.6, 131.6,

129.6, 129.2, 128.7, 128.6, 128.6, 128.4, 128.0, 127.2, 127.0, 126.9, 126.4, 125.4, 71.2, 58.3, 43.7, 43.6, 35.6, 32.3, 32.1, 30.5, 29.9, 27.7, 22.9, 19.5, 19.5, 19.0, 18.9, 14.3; LC/MS mass calcd for C₄₂H₄₇O₄N₈ [M + H]⁺ 727.4, found 727.3; HPLC (UV λ = 254 nm, Gradient B (section 5.2.2)) purity > 90 %, retention time 7.26 min.

5.2 Target-guided synthesis reactions

5.2.1 Dynamic combinatorial chemistry reactions

HPLC conditions for analysis of DCL libraries

Column, Luna 5 μ C18(2), 50 mm x 2.0 mm, and Luna 5 μ C18(2), 30 mm x 4.6 mm, in sequence; flow rate, 1 mL min⁻¹; wavelength, 254 nm; temperature, 30 °C.

Gradient A: H₂O/MeCN (0.01 % TFA) from 95 % H₂O to 80 % H₂O over 15 min, then to 60 % over a further 15 min, and finally to 5 % H₂O over 5 min.

DCL01 Blank equilibrium - Aniline catalysis of reversible hydrazone formation.

The four bishydrazide linkers **HZD01** to **HZD04** (4 x 9.0 μ L, 10 mM, DMSO), aldehyde **ALD01** (9.0 μ L, 10 mM, water) and aniline (75 μ L, 0.1 M, DMSO) were added to a mixture of DMSO (114 μ L, 15 % by vol.) and ammonium acetate buffer (1.3 mL, 100 mM, pH 6.4). The DCL was allowed to stand at room temperature with occasional shaking, and was monitored periodically by direct injection onto HPLC (Gradient A) to establish the blank composition until the relative populations of the hydrazones became constant.

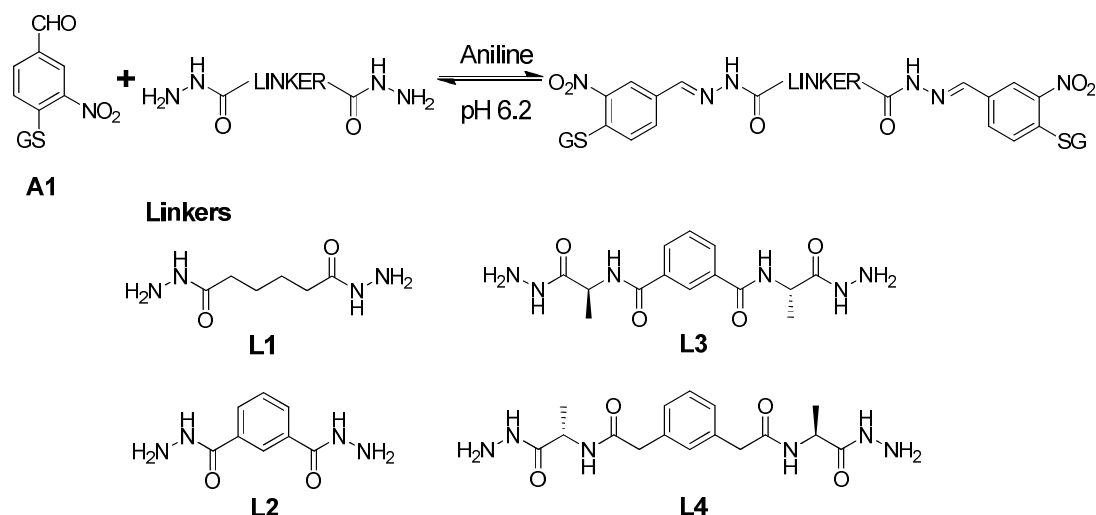
DCL01

Figure 1: DCL01 – library set up between aldehyde **ALD01** and hydrazide linkers **HZD01-HZD04**

DCL01 Blank equilibrium – Approaching equilibrium starting from one hydrazone and three hydrazides.

Three bishydrazide linkers **HZD01** to **HZD03** (3 x 9 μ L, 10 mM, DMSO), acylhydrazone **HZN01-04-01** (45 μ L, 2 mM, DMSO) and aniline (75 μ L, 0.1 M, DMSO) were added to a mixture of DMSO (78 μ L, 15 % by vol.) and ammonium acetate buffer (1.3 mL, 100 mM, pH 6.4). The DCL was allowed to stand at room temperature with occasional shaking, and was monitored periodically by HPLC (Gradient A) to establish the blank composition until the relative populations of the hydrazones became constant.

DCL01 Enzyme-templated DCL reaction – enzyme added initially

Protein, GST (to a final concentration of 20 μ M, for mGSTM1-1; 152.8 μ L, 1.1 mg mL⁻¹, in potassium phosphate buffer 0.1 M, pH 6.8) or BSA (to a final concentration of 20 μ M, 199.5 μ L, 2.0 mg mL⁻¹, in 0.9 % aqueous NaCl solution containing sodium azide), the four hydrazide linkers **HZD01** to **HZD04** (4 x 1.8 μ L, 10 mM, DMSO), aldehyde **ALD01** (1.8 μ L, 10 mM, water) and aniline (15 μ L, 0.1 M, DMSO) were added to a mixture of DMSO (22.8 μ L, 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μ L, for mGSTM1-1; 100.4 μ L, 100 mM, pH 6.4). The DCL was allowed to stand at room temperature, with

occasional shaking, for 48 hours. HPLC analysis (Gradient A) was performed and the traces were compared with the blank equilibrium.

DCL01 Enzyme-templated DCL reaction – enzyme added at equilibrium

The four hydrazide linkers **HZD01** to **HZD04** (4 x 1.8 μL , 10 mM, DMSO), aldehyde **ALD01** (1.8 μL , 10 mM, water) and aniline (15 μL , 0.1 M, DMSO) were added to a mixture of DMSO (22.8 μL , 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μL , for mGSTM1-1; 100.4 μL , 100 mM, pH 6.4). The DCL was allowed to stand at room temperature, with occasional shaking, for 24 hours. GST (to a final concentration of 20 μM , for mGSTM1-1; 152.8 μL , 1.1 mg mL⁻¹, in potassium phosphate buffer 0.1 M, pH 6.8) was added to the equilibrated DCL, which was left to stand at room temperature, with occasional agitation, for a further 48 hours. HPLC analysis (Gradient A) was performed and the traces were compared with the blank equilibrium.

DCL01 Enzyme-templated DCL reaction with inhibitors

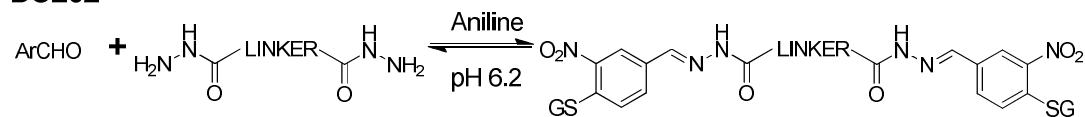
GST (to a final concentration of 20 μM , for mGSTM1-1; 152.8 μL , 1.1 mg mL⁻¹, in potassium phosphate buffer 0.1 M, pH 6.8), the four hydrazide linkers **HZD01** to **HZD04** (4 x 1.8 μL , 10 mM, DMSO), aldehyde **ALD01** (1.8 μL , 10 mM, water), aniline (15 μL , 0.1 M, DMSO) and one inhibitor (6 μL , 10mM, DMSO) were added to a mixture of DMSO (16.8 μL , 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μL , for mGSTM1-1; 100.4 μL , 100 mM, pH 6.4). The DCL was allowed to stand at room temperature, with occasional shaking, for 48 hours. HPLC analysis (Gradient A) was performed and the traces were compared with the blank equilibrium.

DCL01 Enzyme-templated DCL reaction – concentration study of inhibitor DNP GSH

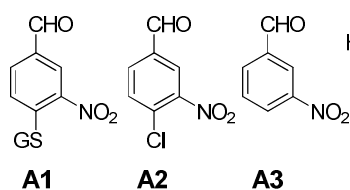
GST (to a final concentration of 20 μM , for mGSTM1-1; 152.8 μL , 1.1 mg mL⁻¹, in potassium phosphate buffer 0.1 M, pH 6.8), the four hydrazide linkers **HZD01** to **HZD04** (4 x 1.8 μL , 10 mM, DMSO), aldehyde **ALD01** (1.8 μL , 10 mM, water), aniline (15 μL , 0.1 M, DMSO) and GSH DNP (to a final concentration of 200, 400 and 600 μM , 6, 12 or 18 μL , 10mM, water) were added to

a mixture of DMSO (22.8 μL , 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μL , for mGSTM1-1; 94.4, 88.4 or 82.4 μL , 100 mM, pH 6.4). The DCL was allowed to stand at room temperature, with occasional shaking, for 48 hours. HPLC analysis (Gradient A) was performed and the traces were compared with the blank equilibrium.

DCL02



Aldehydes



Linkers

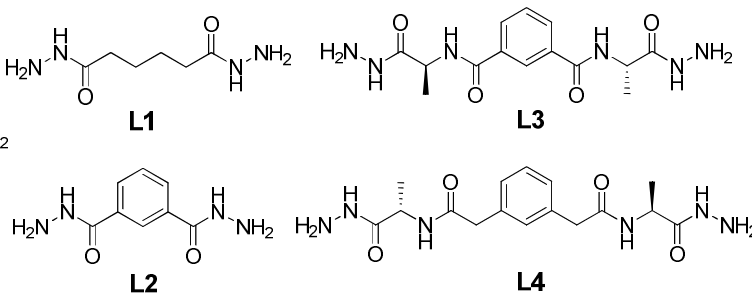


Figure 2: DCL02 – library set up between aldehydes **ALD01-ALD03** and hydrazide linkers **HZD01-HZD04**.

DCL02 Blank equilibrium - Aniline catalysis of reversible hydrazone formation.

The four bishydrazide linkers **HZD01** to **HZD04** (4 x 2.4 μL , 10 mM, DMSO), aldehyde **ALD01** (2.4 μL , 10 mM, water), aldehydes **ALD02** and **ALD03** (2 x 2.4 μL , 10 mM, DMSO) and aniline (15 μL , 0.1 M, DMSO) were added to a mixture of DMSO (15.6 μL , 15 % by vol.) and ammonium acetate buffer (252.6 μL , 100 mM, pH 6.4). The DCL was allowed to stand at room temperature with occasional shaking, and was monitored periodically by HPLC (Gradient A) to establish the blank composition until the relative populations of the hydrazones became constant.

DCL02 Enzyme-templated DCL reaction – enzyme added initially

Protein, GST (to a final concentration of 20 μM , for mGSTM1-1; 152.8 μL , 1.1 mg mL^{-1} , in potassium phosphate buffer 0.1 M, pH 6.8) or BSA (to a final concentration of 20 μM , 199.5 μL , 2.0 mg mL^{-1} , in 0.9 % aqueous NaCl solution

containing sodium azide), the four hydrazide linkers **HZD01** to **HZD04** (4 x 2.4 μL , 10 mM, DMSO), aldehyde **ALD01** (2.4 μL , 10 mM, water), aldehydes **ALD02** and **ALD03** (2 x 2.4 μL , 10 mM, DMSO) and aniline (15 μL , 0.1 M, DMSO) were added to a mixture of DMSO (15.6 μL , 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μL , for mGSTM1-1; 99.8 μL , 100 mM, pH 6.4). The DCL was allowed to stand at room temperature, with occasional shaking, for 48 hours. HPLC analysis (Gradient A) was performed and the traces were compared with the blank equilibrium.

DCL02 Enzyme-templated DCL reaction – enzyme added at equilibrium

The four hydrazide linkers **HZD01** to **HZD04** (4 x 2.4 μL , 10 mM, DMSO), aldehyde **ALD01** (2.4 μL , 10 mM, water), aldehydes **ALD02** and **ALD03** (2 x 2.4 μL , 10 mM, DMSO) and aniline (15 μL , 0.1 M, DMSO) were added to a mixture of DMSO (15.6 μL , 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μL , for mGSTM1-1; 99.8 μL , 100 mM, pH 6.4). The DCL was allowed to stand at room temperature, with occasional shaking, for 24 hours. GST (to a final concentration of 20 μM , for mGSTM1-1; 152.8 μL , 1.1 mg mL^{-1} , in potassium phosphate buffer 0.1 M, pH 6.8) was added to the equilibrated DCL, which was left to stand at room temperature, with occasional agitation, for a further 48 hours. HPLC analysis (Gradient A) was performed and the traces were compared with the blank equilibrium.

DCL02 Enzyme-templated DCL reaction – concentration study of inhibitor DNP GSH

GST (to a final concentration of 20 μM , for mGSTM1-1; 152.8 μL , 1.1 mg mL^{-1} , in potassium phosphate buffer 0.1 M, pH 6.8), the four hydrazide linkers **HZD01** to **HZD04** (4 x 2.4 μL , 10 mM, DMSO), aldehyde **ALD01** (2.4 μL , 10 mM, water), aldehydes **ALD02** and **ALD03** (2 x 2.4 μL , 10 mM, DMSO), aniline (15 μL , 0.1 M, DMSO) and GSH DNP (to a final concentration of 200, 400 and 600 μM , 6, 12 or 18 μL , 10mM, water) were added to a mixture of DMSO (15.6 μL , 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μL , for mGSTM1-1; 93.8, 87.8 or 81.8 μL , 100 mM, pH 6.4). The DCL was allowed to

stand at room temperature, with occasional shaking, for 48 hours. HPLC analysis (Gradient A) was performed and the traces were compared with the blank equilibrium.

5.2.2 *In situ* click chemistry reactions

LC-MS/SIM conditions for analysis of ISCC reactions

Column, Atlantis 3 μ C18, 20 mm x 3.0 mm ISTM column; flow rate, 0.9 mL min⁻¹; wavelength, 254 nm; temperature, 30 °C.

Gradient B: H₂O/MeCN (0.2 % Formic Acid) from 98 % H₂O to 2 % H₂O over 10 min, remaining at 2 % H₂O for 2 min.

Gradient C: H₂O/MeCN (0.2 % Formic Acid) from 98 % H₂O to 70 % H₂O over 5 min, then to 20 % H₂O over 10 min, and eventually to 2 % H₂O over 3 min, then remaining at 2 % H₂O for 2 min.

Table 1: The gradient used to monitor ISCC reactions with the following alkyne present in the reaction mixture

Gradient	Alkyne fragment in reaction
Gradient B	- ALK01 to ALK03 - ALK05 to ALK06 - ALK08 to ALK20
Gradient C	- ALK04 - ALK07

Measure of GST activity under *in situ* click chemistry conditions

L-Glutathione (reduced) (33.3 μ L, 3mM aqueous), CDNB (25 μ L, 4mM, ethanol) and GST (to a final concentration of 150 nM, for hGSTP1-1; 2.1 μ L, 2.0 mg mL⁻¹, stored in potassium phosphate buffer 0.1 M, pH 6.8) were added to ammonium acetate buffer (939.6 μ L 10mM, pH 7.2). Fractions of the reaction mixture (100 μ L) were removed and quenched in methanol (400 μ L) every ten minutes for one hour. The fraction was monitored for the formation of glutathione

conjugate by LC/MS (Gradient B) searching for m/z 472 in negative mode.

GST (for hGSTP1-1; 21 μL , 2.0 mg mL^{-1} , stored in potassium phosphate buffer 0.1 M, pH 6.8) was added to a mixture of DMSO (15 μL , 15 % by vol.) and ammonium acetate buffer (64 μL 10mM, pH 7.2). Samples of these mixtures were incubated at 25 °C over several days. The GST activity was then measured as described above.

Bisazide (AZ01 to AZ06) + Alkyne (ALK01 to ALK04) ISCC reaction – protein-templated reaction

One azide (7.5 μL , 10 mM, DMSO) and one alkyne (7.5 μL , 10 mM, DMSO) were added to a mixture of DMSO (30 μL , 15 % by vol.) and ammonium acetate buffer (for hGSTP1-1; 184.8 μL , for mGSTM1-1; 140 μL , for BSA; 105.6 μL , 10 mM, pH 7.0). Protein, GST (for hGSTP1-1; 70.2 μL , 1.8 mg mL^{-1} , for mGSTM1-1; 115 μL , 1.1 mg mL^{-1} , stored in potassium phosphate buffer 0.1 M, pH 6.8) or BSA (149.4 μL , 2.0 mg mL^{-1} , stored in 0.9 % aqueous NaCl solution containing sodium azide) was added and the reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μL) was removed and quenched in methanol (400 μL). LC/MS analysis (Gradient B or C, see Table 1) was performed in single ion monitoring (SIM) mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for monotriazole and bistriazole products.

Bisazide (AZ01) + Alkyne (ALK03/ALK04) ISCC reaction – optimisation of time of assay

Azide **AZ01** (62.5 μL , 10 mM, DMSO) and one alkyne **ALK03/ALK04** (62.5 μL , 10 mM, DMSO) were added to a mixture of DMSO (250 μL , 15 % by vol.) and ammonium acetate buffer (to a final volume of 2.125 mL, for hGSTP1-1; 1.75 mL, for background; 2.125 mL, 10 mM, pH 7.0). Reactions were set up in duplicate with one to measure the background reaction and to the second protein, hGSTP1-1 (to a final concentration of 15 μM , 375 μL , 2.8 mg mL^{-1} , stored in potassium phosphate buffer 0.1 M, pH 6.8) was added and the reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample

of the reaction mixture (200 μL) was removed over several time points and quenched in methanol (600 μL). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for monotriazole and bistriazole products.

Bisazide (AZ01) + Alkyne (ALK04) ISCC reaction – optimisation of fragment concentration

Azide **AZ01** (final concentration 50/100/250 μM , 2.5/5.0/12.5 μL , 10 mM, DMSO) and alkyne **ALK04** (final concentration 50/100/250 μM , 2.5/5.0/12.5 μL , 10 mM, DMSO) were added to a mixture of DMSO (to a total volume of 75 μL , 15 % by vol.) and ammonium acetate buffer (to a final volume of 425 μL , for hGSTP1-1; 140 μL , for background; 105.6 μL , 10 mM, pH 7.0). Reactions were set up in duplicate with one to measure the background reaction and to the second protein, hGSTP1-1 (to a final concentration of 15 μM , 105 μL , 2.0 mg mL^{-1} , stored in potassium phosphate buffer 0.1 M, pH 6.8) was added and the reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μL) was removed and quenched in methanol (400 μL). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for monotriazole and bistriazole products.

Bisazide (AZ01) + Alkyne (ALK03/ALK04) ISCC reaction – optimisation of GST concentration

Azide **AZ01** (20 μL , 10 mM, DMSO), one alkyne **ALK03/ALK04** (20 μL , 10 mM, DMSO) and protein, hGSTP1-1 (to a final concentration of 0/1/5/15/35/50 μM , 0/8/40/120/280/400 μL , 2.8 mg mL^{-1} , stored in potassium phosphate buffer 0.1 M, pH 6.8) were added to a mixture of DMSO (80 μL , 15 % by vol.) and ammonium acetate buffer (to a final volume of 680 μL , 680/672/640/560/400/280 μL , 10 mM, pH 7.0). The reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μL) was removed and quenched in methanol (600 μL). LC/MS analysis

(Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

Bisazide (AZ01) + Alkyne (ALK01 to ALK20) ISCC reaction – Background reaction

Azide **AZ01** (7.5 μ L, 10 mM, DMSO) and one alkyne (7.5 μ L, 10 mM, DMSO) were added to a mixture of DMSO (30 μ L, 15 % by vol.) and ammonium acetate buffer (255 μ L, 10 mM, pH 7.0). The reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μ L) was removed and quenched in methanol (400 μ L). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

Bisazide (AZ01) + Alkyne (ALK01 to ALK20) ISCC reaction – protein-templated reaction

Azide **AZ01** (7.5 μ L, 10 mM, DMSO) and one alkyne (7.5 μ L, 10 mM, DMSO) were added to a mixture of DMSO (30 μ L, 15 % by vol.) and ammonium acetate buffer (to a final volume of 255 μ L, for mGSTM1-1; 140 μ L, for BSA; 105.6 μ L, 10 mM, pH 7.0). Protein, GST (to a final concentration of 15 μ M, for mGSTM1-1; 115 μ L, 1.1 mg mL⁻¹, stored in potassium phosphate buffer 0.1 M, pH 6.8) or BSA (149.4 μ L, 2.0 mg mL⁻¹, stored in 0.9 % aqueous NaCl solution containing sodium azide) was added and the reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μ L) was removed and quenched in methanol (400 μ L). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[M + H]^+$ and $[M + Na]^+$ mass peaks for monotriazole and bistriazole products.

Bisazide (AZ01) + Alkyne (ALK01 to ALK20) ISCC reaction – protein-templated reaction + inhibitor

Azide **AZ01** (7.5 μ L, 10 mM, DMSO), one alkyne (7.5 μ L, 10 mM, DMSO) and one inhibitor (15 μ L, 10mM, DMSO) were added to a mixture of DMSO (15 μ L, 15 % by vol.) and ammonium acetate buffer (to a final volume of 255 μ L, for mGSTM1-1; 140 μ L, 10 mM, pH 7.0). GST (to a final concentration of 15 μ M, for

mGSTM1-1; 115 μL , 1.1 mg mL^{-1} , stored in potassium phosphate buffer 0.1 M, pH 6.8) was added and the reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μL) was removed and quenched in methanol (400 μL). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for monotriazole and bistriazole products.

Monoazide (TZ03a-01 or TZ04a-01) + Alkyne (ALK01 to ALK20) ISCC reaction – Background reaction

Azide **TZ03a-01** or **TZ04a-01** (1.5 μL , 10 mM, DMSO) and one alkyne (7.5 μL , 10 mM, DMSO) were added to a mixture of DMSO (36 μL , 15 % by vol.) and ammonium acetate buffer (255 μL , 10 mM, pH 7.0). The reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μL) was removed and quenched in methanol (400 μL). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for bistriazole product.

Monoazide (TZ03a-01 or TZ04a-01) + Alkyne (ALK01 to ALK20) ISCC reaction – protein-templated reaction

Azide **TZ03a-01** or **TZ04a-01** (1.5 μL , 10 mM, DMSO) and one alkyne (7.5 μL , 10 mM, DMSO) were added to a mixture of DMSO (36 μL , 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μL , for mGSTM1-1; 140 μL , 10 mM, pH 7.0). Protein, GST (final concentration 15 μM , for mGSTM1-1; 115 μL , 1.1 mg mL^{-1} , stored in potassium phosphate buffer 0.1 M, pH 6.8) or BSA (105.6 μL , 1.1 mg mL^{-1} , stored in 0.9 % aqueous NaCl solution containing sodium azide) was added and the reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μL) was removed and quenched in methanol (400 μL). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for bistriazole product.

Monoazide (TZ03a-01 or TZ04a-01) + Alkyne (ALK01 to ALK20) ISCC reaction – protein-templated reaction + inhibitor

Azide **TZ03a-01** or **TZ04a-01** (1.5 μL , 10 mM, DMSO), one alkyne (7.5 μL , 10 mM, DMSO) and one inhibitor (15 μL , 10mM, DMSO) were added to a mixture of DMSO (21 μL , 15 % by vol.) and ammonium acetate buffer (to a total volume of 255 μL , for mGSTM1-1; 140 μL , 10 mM, pH 7.0). GST (final concentration 15 μM , for mGSTM1-1; 115 μL , 1.1 mg mL⁻¹, stored in potassium phosphate buffer 0.1 M, pH 6.8) was added and the reaction mixture was allowed to stand at room temperature, with occasional shaking, for 7 days. A sample of the reaction mixture (200 μL) was removed and quenched in methanol (400 μL). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for bistriazole product.

Small scale thermal click reaction for reference

One alkyne (25 μL , 10 mM, acetonitrile) and one azide linker (25 μL , 10 mM, acetonitrile) were dissolved in acetonitrile (450 μL). The reaction was placed under a nitrogen atmosphere and heated to 70 °C. After 24 hours of heating the reaction mixture was filtered and injected directly onto the LC/MS column to analyse the product retention times (Gradient B or C, see Table 1).

LC/MS-SIM response to ISCC library fragments and products

Azide **AZ01**, alkynes **ALK04** and **ALK18** and triazole products **TZ04a-01**, **TZ04a-01-04a**, **TZ18a-01** and **TZ18a-01-18a** were monitored on the LC/MS-SIM over a range of concentrations. For each sample, compound (final concentrations 0.1, 1, 10 and 100 μM , 3 μL , from stock solution 0.01, 0.1, 1 and 10 mM, DMSO) was dissolved in DMSO (42 mL, 15 % by vol.) and ammonium acetate buffer (255 μL , 10 mM, pH 7.0). A sample of the mixture (200 μL) was removed and quenched in methanol (500 μL). LC/MS analysis (Gradient B or C, see Table 1) was performed in SIM mode monitoring for $[\text{M} + \text{H}]^+$ and $[\text{M} + \text{Na}]^+$ mass peaks for the product.

5.3 GST Binding assays

Inhibition assays were carried out on a Molecular Devices SpectraMax M5 multimode plate reader, using costar clear 96-well plates. Kinetic data was analysed using SigmaPlot®.

5.3.1 CDNB binding assay – to determine K_M and V_{MAX}

The K_M values for GSH and CDNB for each GST isoform in a DMSO concentration of 15 % were determined using the CDNB assay described by Habig and co-workers.^[7]

To determine K_M (CDNB): To a 360 μL well were added phosphate buffer (225 μL , 0.1 M, pH 6.8), DMSO (15 μL , 15 % by vol.), GST (15 μL , for mGSTM1-1 0.022 mg mL^{-1}) and GSH (15 μL , 40 mM stock, phosphate buffer - 0.1 M pH 6.8). The solution was mixed well and after incubation at 25 °C for 5 minutes, CDNB (15 μL , 12 samples to a final concentration of 20 μM – 2 mM, ethanol) was added quickly and mixed. Absorbance was measured at 340 nm, 25 °C for 3 minutes.

To determine K_M (GSH): To a 360 μL well were added phosphate buffer (225 μL , 0.1 M, pH 6.8), DMSO (15 μL , 15 % by vol.), GST (15 μL , 0.022 mg mL^{-1}) and CDNB (15 μL , 40 mM stock, ethanol). The solution was mixed well and after incubation at 25 °C for 5 minutes, GSH (15 μL , 12 samples to a final concentration of 20 μM – 2 mM, phosphate buffer - 0.1 M pH 6.8) was added quickly and mixed. Absorbance was measured at 340 nm, 25 °C for 3 minutes.

Table 2: K_M data for all four GST isoforms described in the presence of 15% DMSO.

GST Isoform	K_M^{CDNB} (mM)	K_M^{GSH} (mM)
mGSTM1-1	0.21 ± 0.03	0.55 ± 0.06
hGSTP1-1	0.72 ± 0.12	0.58 ± 0.18
SjGST	3.02 ± 0.81	0.55 ± 0.07
mGSTA4-4	0.99 ± 0.64	1.19 ± 0.29

5.3.2 CDNB binding assay – to determine IC₅₀ values

To a 360 μL well were added phosphate buffer (225 μL , 0.1 M, pH 6.8), DMSO (15 μL , 15 % by vol.), GST (15 μL , 22.0 $\mu\text{g mL}^{-1}$) and inhibitor (15 μL , final concentration 0.001 - 100 μM). The solution was mixed well and after incubation at 25 °C for 5 minutes, CDNB (15 μL , at the K_M value for each GST isoform) and GSH (15 μL , 40 mM, in large excess of the K_M) were added quickly and mixed. Absorbance was measured at 340 nm, 25 °C for 3 minutes.

Table 3: Raw data from CDNB binding assay for compounds from Chapter 2.

Compound	mGSTM1-1 logIC ₅₀	hGSTP1-1 logIC ₅₀	SjGST logIC ₅₀	mGSTA4-4 logIC ₅₀
ALD01	2.534 ± 0.344	> 2.7	2.424 ± 0.141	> 2.7
ALD03	> 2.7	-	-	-
HZD03	> 2.7	> 2.7	-	-
HZN01-01-01	0.082 ± 0.097	2.102 ± 0.422	0.534 ± 0.085	> 2
HZN01-02-01	-0.473 ± 0.081	1.072 ± 0.117	-0.599 ± 0.068	> 2
HZN01-03-01	-1.30 ± 0.26	1.129 ± 0.203	-0.005 ± 0.091	> 2
HZN01-04-01	-0.384 ± 0.086	-0.449 ± 0.172	0.255 ± 0.054	> 2
HZN02-02-02	0.305 ± 0.143	-	-	-
HZN02-03-02	0.970 ± 0.107	-	-	-
HZN02-04-02	1.698 ± 0.346	-	-	-
HZN03-02-03	0.528 ± 0.120	-	-	-
HZN03-03-03	2.027 ± 0.332	-	-	-
HZN03-04-03	> 2	-	-	-

Table 4: Raw data from CDNB binding assay for compounds from Chapter 3.

Compound	mGSTM1-1 logIC₅₀	hGSTP1-1 logIC₅₀	SjGST logIC₅₀	mGSTA4-4 logIC₅₀
ALK03	0.042 ± 0.038	1.434 ± 0.154	1.021 ± 0.064	> 2
ALK04	> 2	> 2	-	-
TZ03a-01	-0.860 ± 0.082	1.472 ± 0.117	0.902 ± 0.199	> 2
TZ04a-01	2.112 ± 0.433	> 2	> 2	> 2
TZ04s-01	> 2	-	-	-
TZ18a-01	> 2	-	-	-
TZ03a-01-03a	-2.427 ± 0.132	0.879 ± 0.173	-1.850 ± 0.097	> 2
TZ03a-01-04a	-1.259 ± 0.102	1.388 ± 0.127	1.117 ± 0.086	> 2
TZ04a-01-01a	0.889 ± 0.070	-	-	-
TZ04a-01-04a	1.119 ± 0.119	2.222 ± 0.392	1.639 ± 0.094	> 2
TZ04a-01-04s	1.229 ± 0.093	2.148 ± 0.285	-	-
TZ04a-01-07a	1.407 ± 0.133	1.742 ± 0.266	-	-
TZ04a-01-07s	0.583 ± 0.118	1.748 ± 0.204	-	-
TZ04a-01-08a	1.530 ± 0.130	-	-	-
TZ04a-01-08s	1.793 ± 0.236	-	-	-
TZ04a-01-09a	1.471 ± 0.143	-	-	-
TZ04a-01-09s	1.812 ± 0.284	-	-	-
TZ04a-01-10a	1.589 ± 0.108	1.565 ± 0.165	-	-
TZ04a-01-11a	0.986 ± 0.081	-	-	-
TZ04a-01-11s	1.545 ± 0.071	-	-	-
TZ04a-01-18a	1.484 ± 0.129	0.337 ± 0.143	-	-

5.3.3 CDNB competition assay – to determine K_i values

Four different inhibitor concentrations were paired with six substrate concentrations in order to determine K_i . To a 360 μL well were added phosphate buffer (225 μL , 0.1 M, pH 6.8), DMSO (15 μL , 15 % by vol.), GST (15 μL , 22.0 $\mu\text{g mL}^{-1}$) and inhibitor at four concentrations (15 μL , 0 - 10 μM). The solution was mixed well and after incubation at 25 °C for 5 minutes, CDNB (15 μL , 0.05 - 0.2 mM, at six concentrations around the K_M value) and GSH (15 μL , 40 mM, in large excess of the K_M) were added quickly and mixed. Absorbance was measured at 340 nm, 25 °C for 3 minutes.

5.3.4 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was carried out in collaboration with Dr Elizabeth Blackburn at the biophysical characterization facility in the centre for translational and chemical biology at the University of Edinburgh. The ITC measurements were performed on a MicroCal LLC Auto-ITC HT Microcalorimeter at 25 °C. mGSTM1-1 was dialysed against a 0.1 M potassium phosphate buffer, pH 6.8. The final dialysis buffer was used to make up the ligand solution as well as for instrument calibration and baseline controls. **HZN01-02-01** and **HZN01-03-01** were dissolved in the dialysis buffer to concentrations of 300 μM and 260 μM respectively. The concentration of mGSTM1-1 was determined by measuring the absorption at 280 nm to give a final concentration of 17.7 μM . The protein mGSTM1-1 was placed in the 2 mL sample chamber and compounds **HZN01-02-01** or **HZN01-03-01** in the syringe. A typical ITC measurement consisted of a first control injection of 1 μL followed by 29 successive injections of 10 μL for 20 seconds with a 3 minute interval between each injection. Raw data were collected, corrected for ligand heats dilution, and the peaks generated integrated using ORIGIN software (Microcal Inc) by plotting the values in microcalories against the molar ratio of injectant to reactant within the cell. Data was fitted using the one single-site binding model.^[8] From the dissociation constant K_d and the reaction enthalpy value ΔH , the change in free Gibbs energy (ΔG°) and entropy change (ΔS°) can be calculated using the equation $\Delta G^\circ = -RT \ln(1/K_d) = \Delta H - T\Delta S^\circ$ where R is the universal gas constant and T the absolute temperature.

5.4 Expression and purification of GST

Plasmid pET15-b-hGSTP1 was a kind gift from Dr. Sylvie Blond (University of Illinois, USA) to Dr Dominic Campopiano. All protein expression and purification was carried out in the lab of Dr Dominic Campopiano. All media were sterilized by autoclaving at 121 °C for 20 minutes at 15 psi prior to use.

5.4.1 Expression and purification of hGSTP1-1

Overexpression of His₆-Tagged hGSTP1-1 was carried out by transforming *E. coli* BL21 (DE3) cells with the pET15b/hGSTP1 vector. A single colony was added to 200 mL Luria Bertani (LB) medium (tryptone (10 g L⁻¹), yeast extract (5 g L⁻¹), sodium chloride (10 g L⁻¹); pH adjusted to 7.5 with sodium hydroxide) supplemented with ampicillin (100 µg mL⁻¹) and grown overnight at 37 °C and 250 rpm. This seed culture was then used to inoculate 3 L of fresh LB medium and grown to OD₆₀₀ of 0.6 before induction with isopropyl β-D-1-thiogalactopyranoside (IPTG, 1.0 mM final concentration) for 3 to 4 hours at 37 °C and 250 rpm. The cells were then harvested by centrifugation (4 000 x g for 15 min at 4 °C) and stored at -20 °C. Cells overexpressing His₆-hGSTP1-1 were resuspended in binding buffer (50 mM Tris- HCl, 0.5 M NaCl, 5 mM imidazole, pH 7, 4 mL per gram of wet cell paste) with one protease inhibitor cocktail tablet (EDTA-free) and disrupted by sonication (15 pulses of 30 s at 30 s intervals) at 4 °C. The cell debris were removed by centrifugation at 27 000 x g for 30 minutes at 4 °C, after which the supernatant was filtered through a 0.45 µm membrane prior to chromatography. The cell lysate was loaded onto a 5 mL HisTrap™ HP column (GE Healthcare) previously equilibrated with the same binding buffer. The column was then washed with 5 column volumes of binding buffer before the bound material was eluted using a linear gradient of 0-100 % elution buffer (50 mM Tris-HCl, 0.5 M NaCl, 0.5M imidazole, pH 8) over 20 column volumes at 4 °C. Fractions were analysed by SDS-PAGE and those containing hGSTP1-1 were pooled and dialysed against 2 L of a 0.1 M potassium buffer containing 150mM NaCl, pH 6.8 at 4 °C. The dialysed fractions were then applied onto a 320 mL HiPrep Sephacryl S-300 HR column (GE Healthcare) pre-equilibrated with a 0.1 M potassium phosphate buffer containing 150 mM NaCl, pH 6.8. The column was then eluted with one column volume of the same buffer. Fractions containing hGSTP1-1 were pooled and dialysed against 2 L of

a 0.1 M potassium phosphate buffer, pH 6.8 at 4 °C.

5.4.2 Polyacrylamide Gel Electrophoresis (PAGE)

SDS-PAGE was used to analyse the fractions collected during the purification of the protein.

SDS sample 2x loading buffer contained Tris/HCl (1.5 M, pH 8.0, 1 ml), glycerol (2 ml), bromophenol blue (0.05 %, 2 ml), SDS (10 %, 1.6 ml) and β -mercaptoethanol (0.4 ml). Samples for analysis by SDS-PAGE were prepared by addition of appropriate volume of this buffer, followed by heating (100 °C, 5 min) and centrifugation (13,000 rpm, 5 min).

A standard method was used for the production of acrylamide gels with a running gel containing 15 % acrylamide and a stacking gel of 4 % acrylamide. Briefly, the running gel (acrylamide/bis (29:1) 15 % w/v, SDS 0.1 % w/v, TEMED 0.15 % v/v and APS 0.1 % w/v in Tris buffer (375 mM, pH 8.8)) and was poured between glass plates, leveled and set at room temperature. The stacking gel (acrylamide/bis (29:1) 4% w/v, SDS 0.1 % w/v, TEMED 0.15 % v/v and APS 0.1 % w/v in Tris buffer (375 mM, pH 6.8)) was then added and set at room temperature using a mould to produce the wells in the finished gel. Gels were run in protein-running buffer (TGS buffer – 25 mM Tris, 192 mM glycine, 0.1 % (w/v) SDS, pH 8.3) at 150 V and then visualized using Coomassie Brilliant Blue R250, GelCode (Pierce) in accordance with manufacturer's instructions.

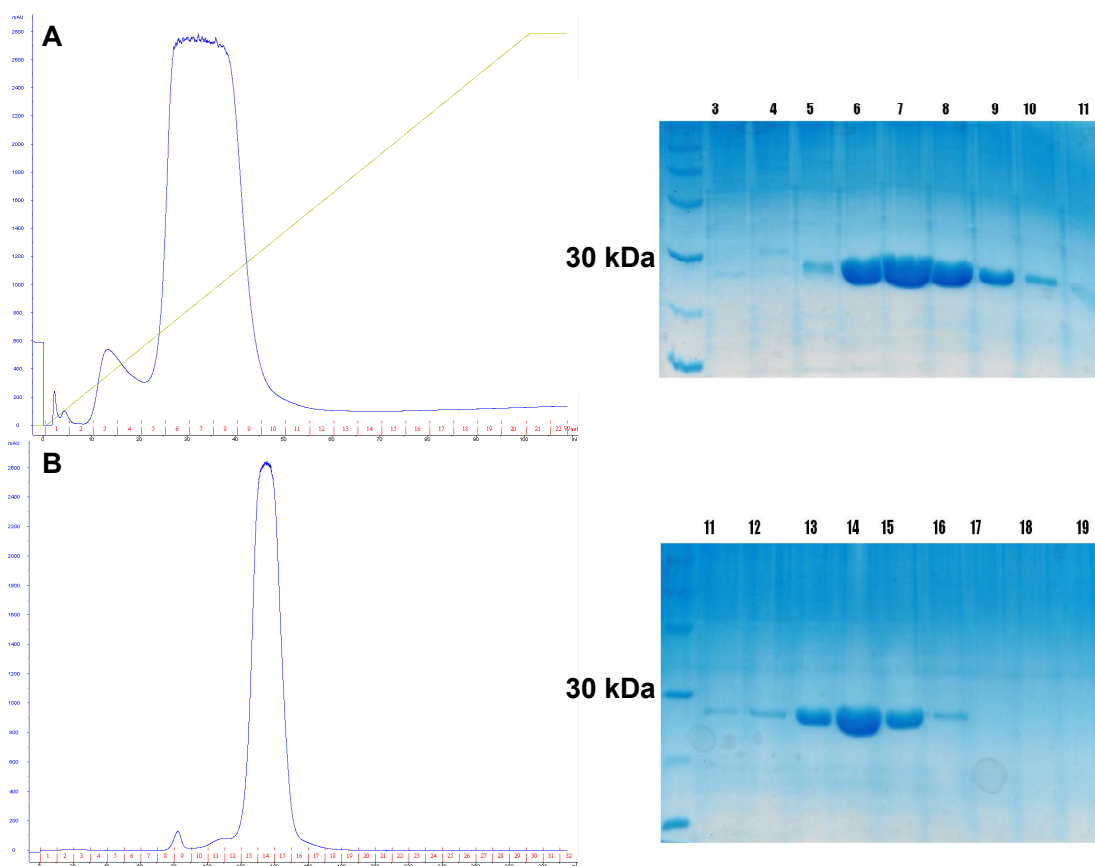


Figure 3: Elution profiles of His₆-hGST P1-1 from **A)** 5ml Ni HisTrap column and **B)** S200 size exclusion column showing the A₂₈₀ and SDS-PAGE analyses of the fractions collected.

5.4.3 Expression and purification of mGSTM1-1, SjGST and mGSTA4-4

The expression and purification of mGSTM1-1, SjGST and mGSTA4-4 was done by Anne Caniard.^[2, 9]

5.5 References

- [1] in *SBDSWeb*: <http://riodb01.ibase.aist.go.jp/sdbs/>, National Institute of Advanced Industrial Science and Technology, **Date of Access May 12, 2011**.
- [2] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat Chem* **2010**, *2*, 490.
- [3] E. D. Goddard-Borger, R. V. Stick, *Org. Lett.* **2007**, *9*, 3797.
- [4] J. T. Lundquist IV, J. C. Pelletier, *Org. Lett.* **2001**, *3*, 781.
- [5] C. Grandjean, A. Boutonnier, C. Guerreiro, J.-M. Fournier, L. A. Mulard, *J. Org. Chem.* **2005**, *70*, 7123.

- [6] P. van der Peet, C. T. Gannon, I. Walker, Z. Dinev, M. Angelin, S. Tam, J. E. Ralton, M. J. McConville, S. J. Williams, *ChemBioChem* **2006**, *7*, 1384.
- [7] W. H. Habig, M. J. Pabst, W. B. Jakoby, *J. Biol. Chem.* **1974**, *249*, 7130.
- [8] A. Arouri, P. Garidel, W. Kliche, A. Blume, *Eur. Biophys. J.* **2007**, *36*, 647.
- [9] A. M. Caniard, PhD thesis, University of Edinburgh (Edinburgh), **2010**.

Appendix

6.1 Identification of peaks observed in DCL01 and DCL02

6.1.1 Identification of peaks observed in DCL01

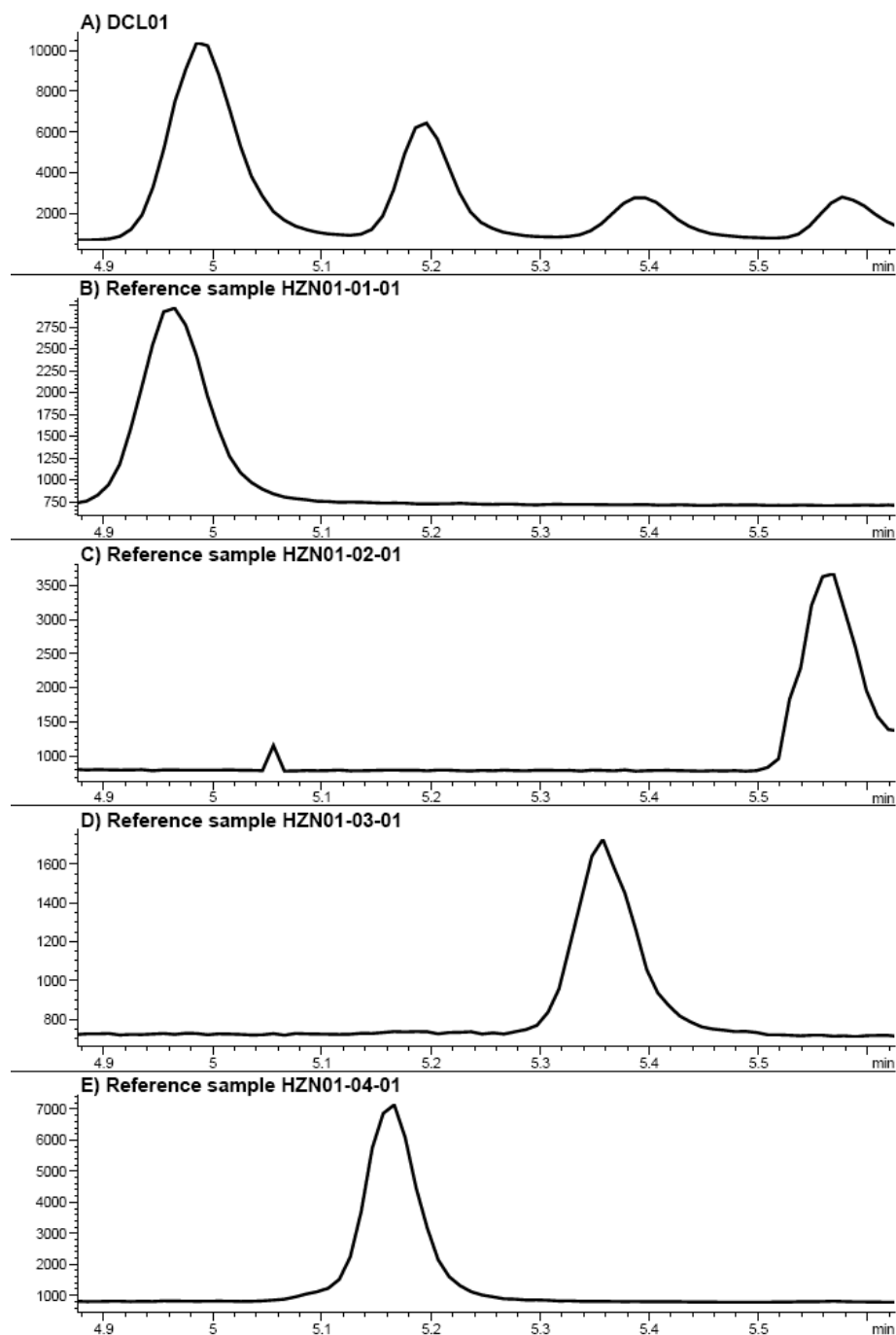


Figure 1: Identification of peaks in DCL01 using isolated reference samples. A) DCL01, B) HZN01-01-01, C) HZN01-02-01, D) HZN01-03-01, E) HZN01-04-01.

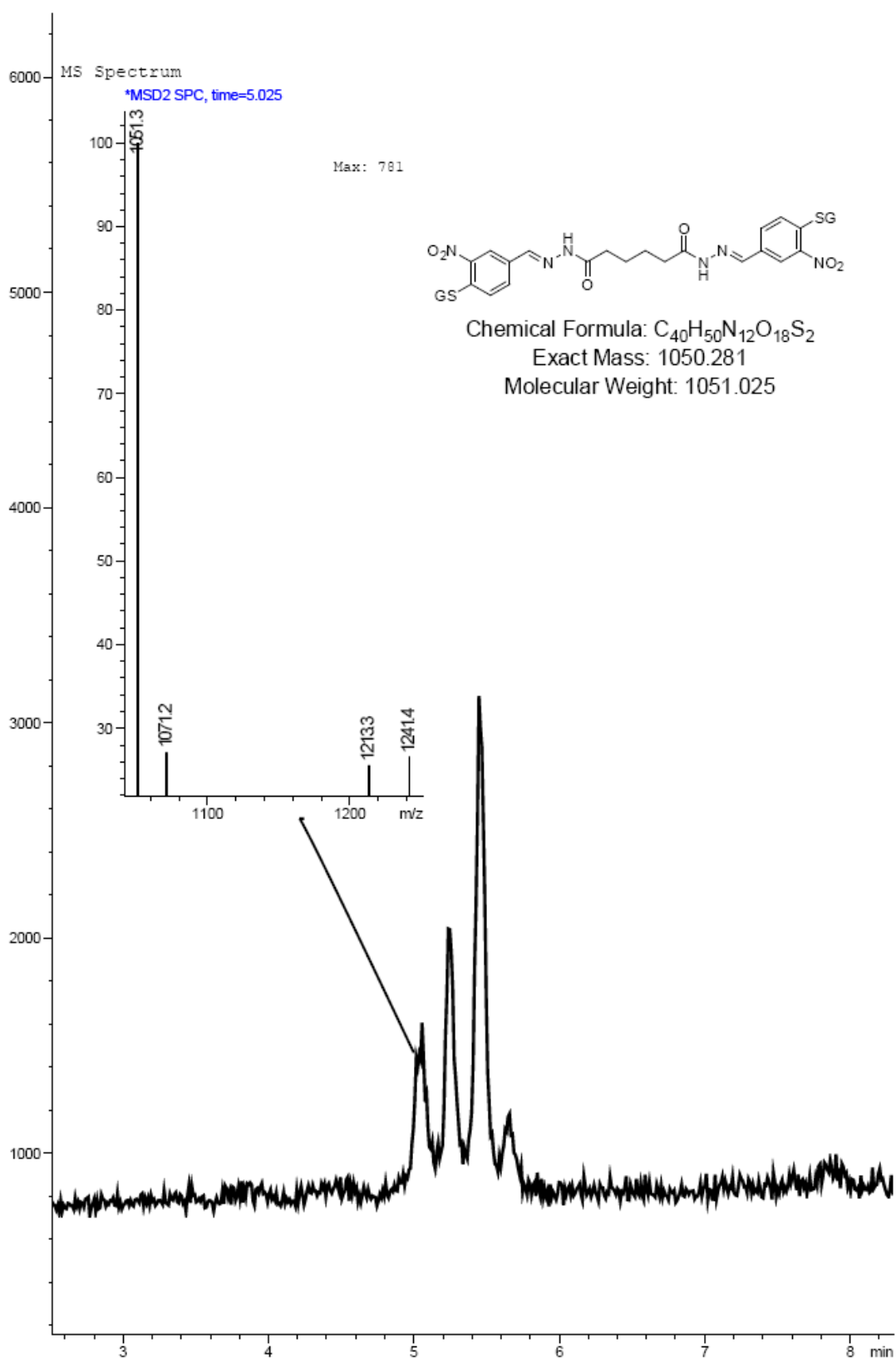


Figure 2: LC/MS-SIM trace of **DCL01** monitoring for mass of bis acylhydrazone products and MS spectrum for peak one identifying **HZN01-01-01** (mass = 1051.3).

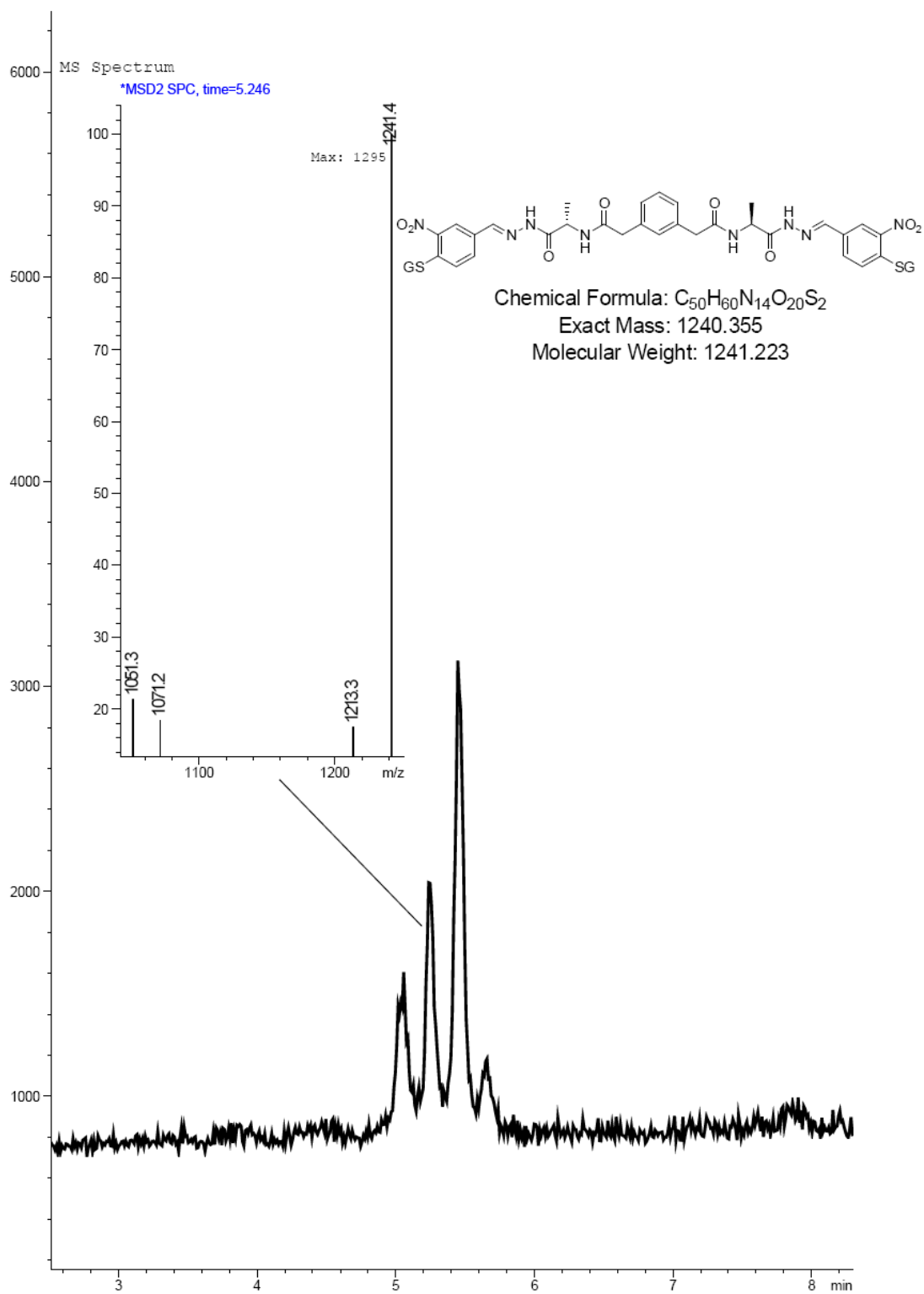


Figure 3: LC/MS-SIM trace of **DCL01** monitoring for mass of bis acylhydrazone products and MS spectrum for peak two identifying **HZN01-04-01** (mass = 1241.4).

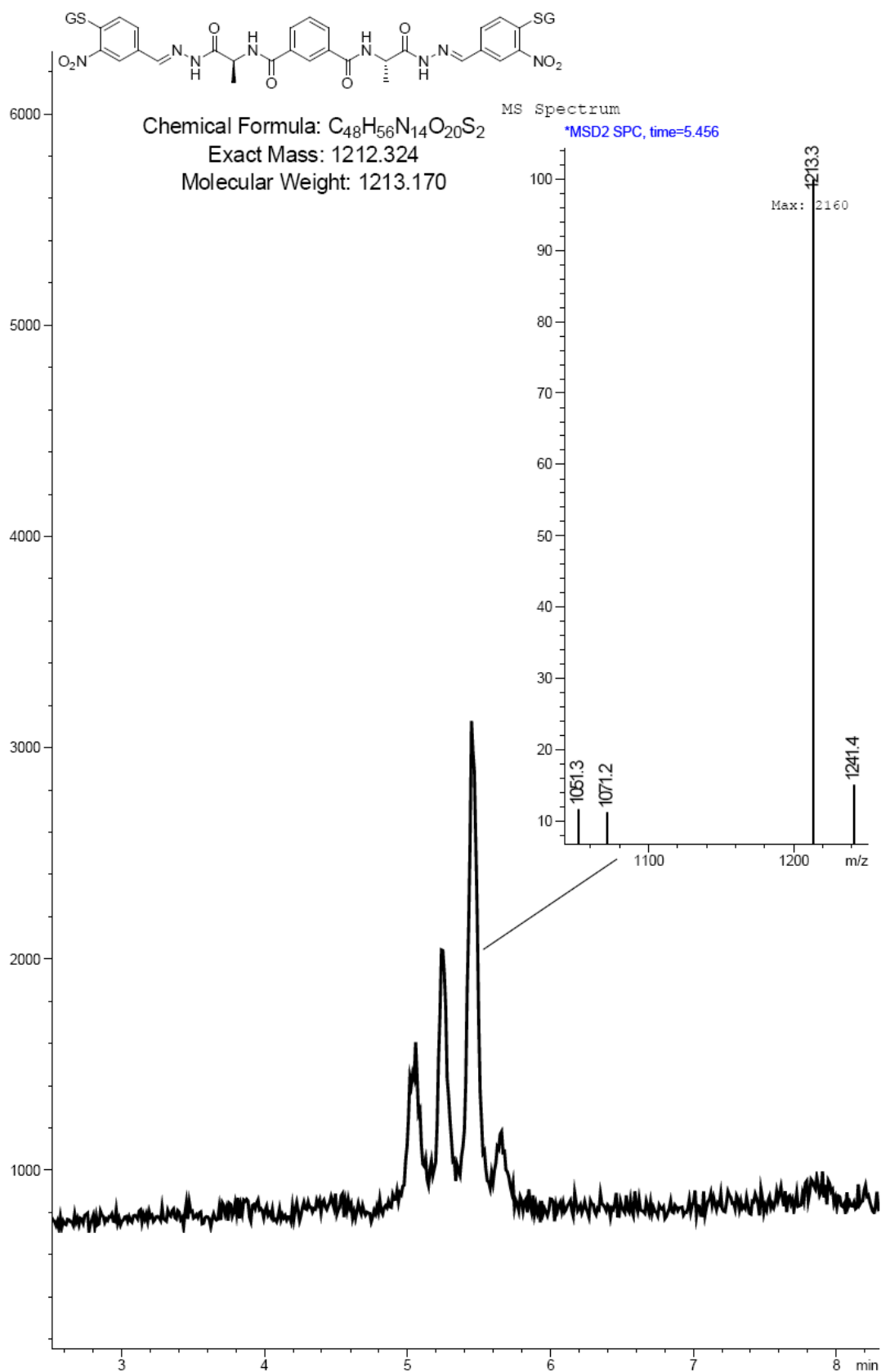


Figure 4: LC/MS-SIM trace of **DCL01** monitoring for mass of bis acylhydrazone products and MS spectrum for peak three identifying **HZN01-03-01** (mass = 1213.3).

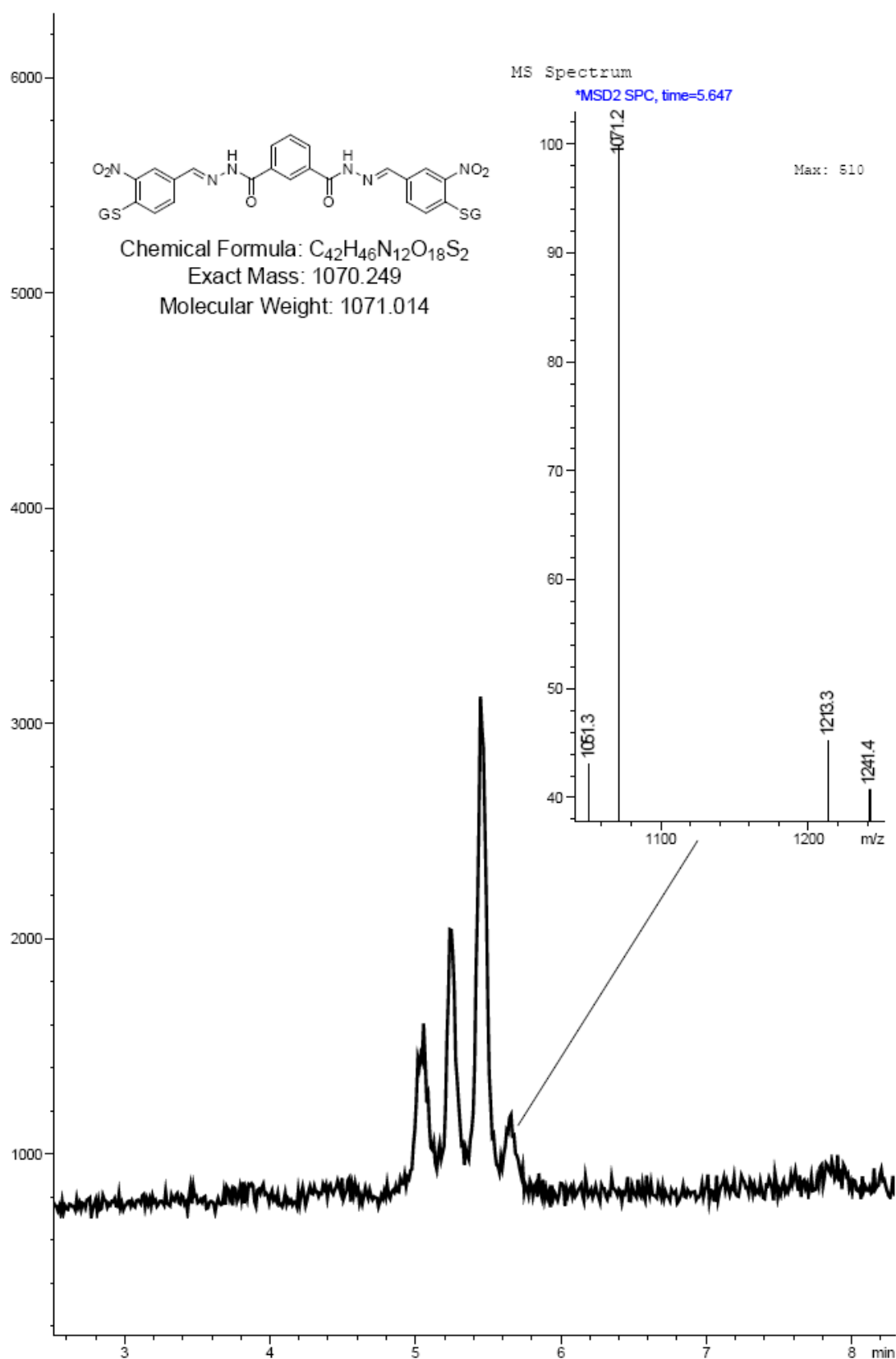


Figure 5: LC/MS-SIM trace of **DCL01** monitoring for mass of bis acylhydrazone products and MS spectrum for peak four identifying **HZN01-02-01** (mass = 1071.2).

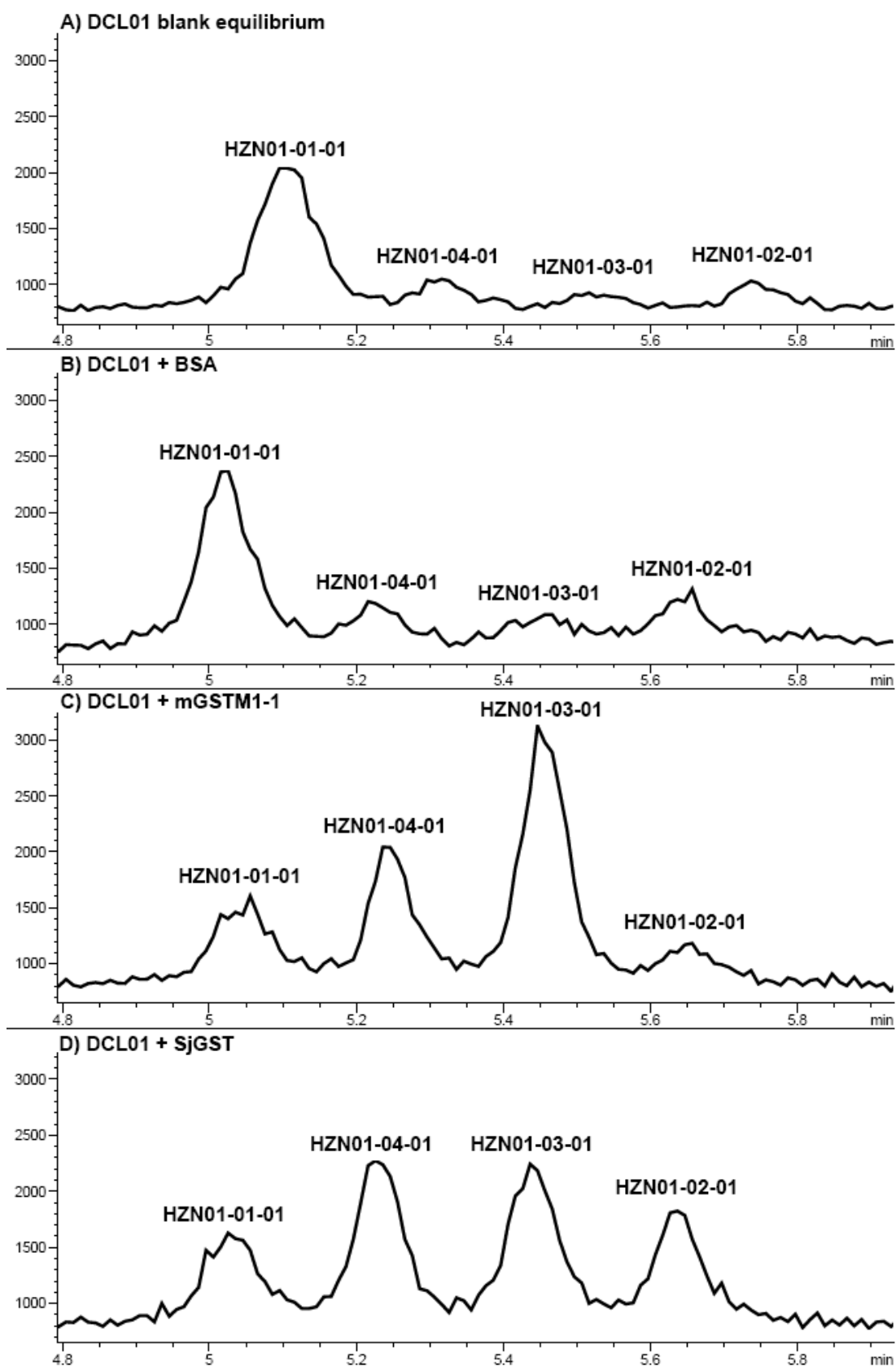


Figure 6: LC/MS-SIM traces of **DCL01** in the presence of GST, monitoring for mass of bis acylhydrazone products. A) blank equilibrium, B) BSA, C) mGSTM1-1, D) SjGST.

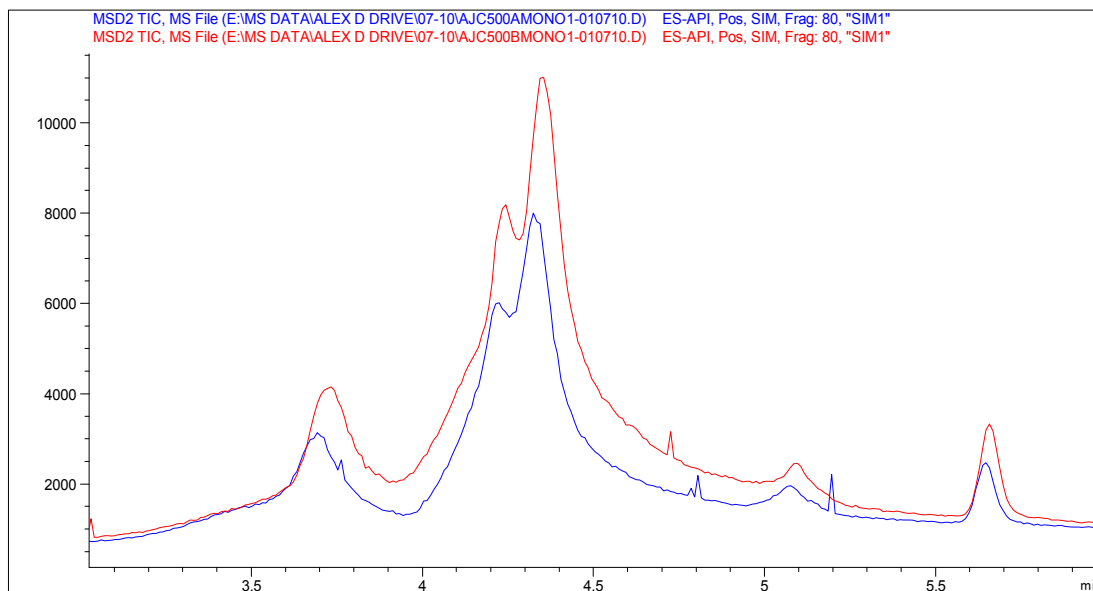


Figure 7: LC/MS-SIM trace monitoring for mass of four mono acylhydrazones for **DCL01** in the absence (blue) and presence (red) of hGSTP1-1.

6.1.2 Identification of peaks observed in DCL02

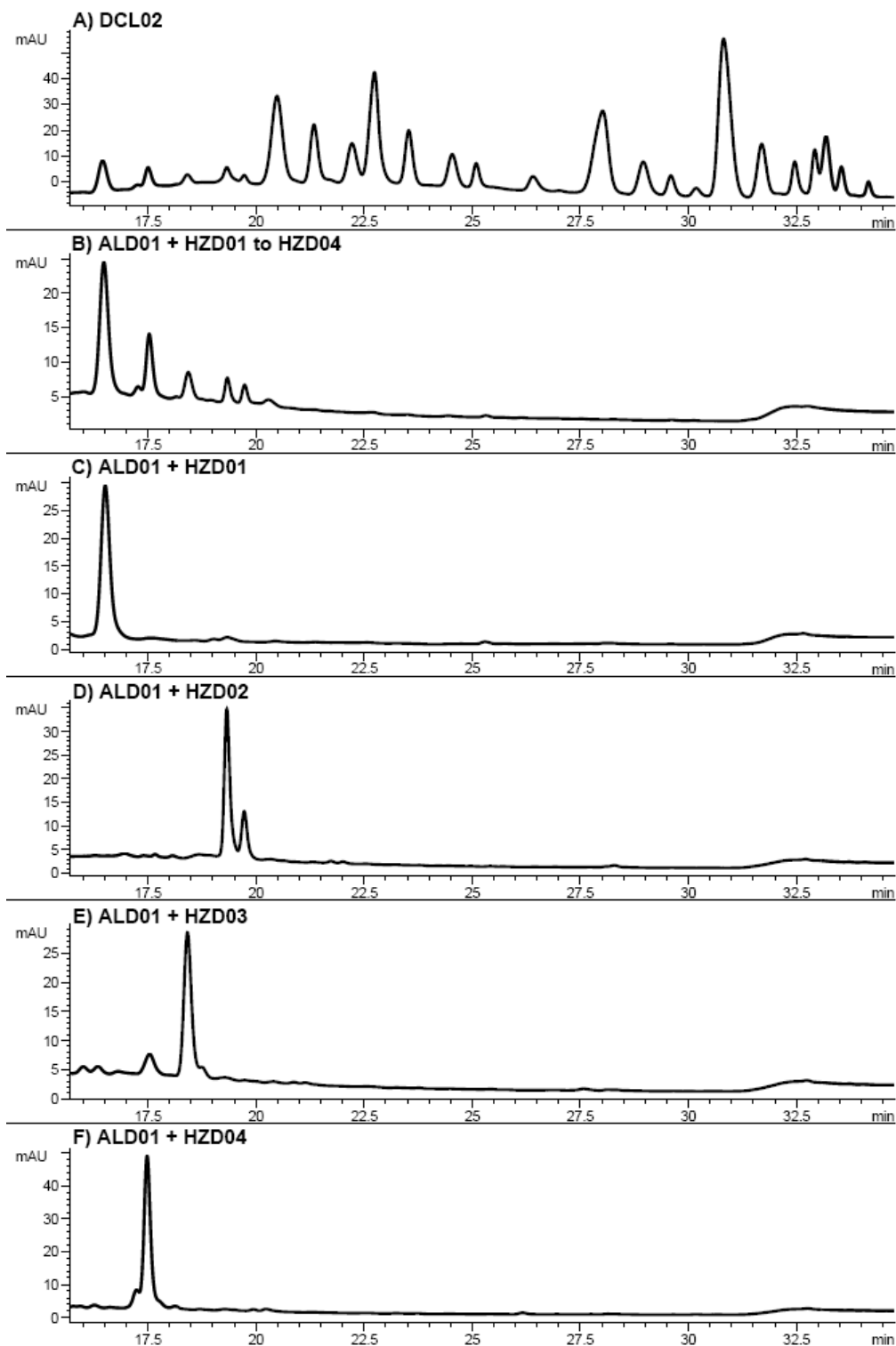


Figure 8: Deconvolution of DCL02 to assign peaks. A) DCL02, B) ALD01 + HZD01 to HZD04, C) ALD01 + HZD01, D) ALD01 + HZD02, E) ALD01 + HZD03, F) ALD01 + HZD04.

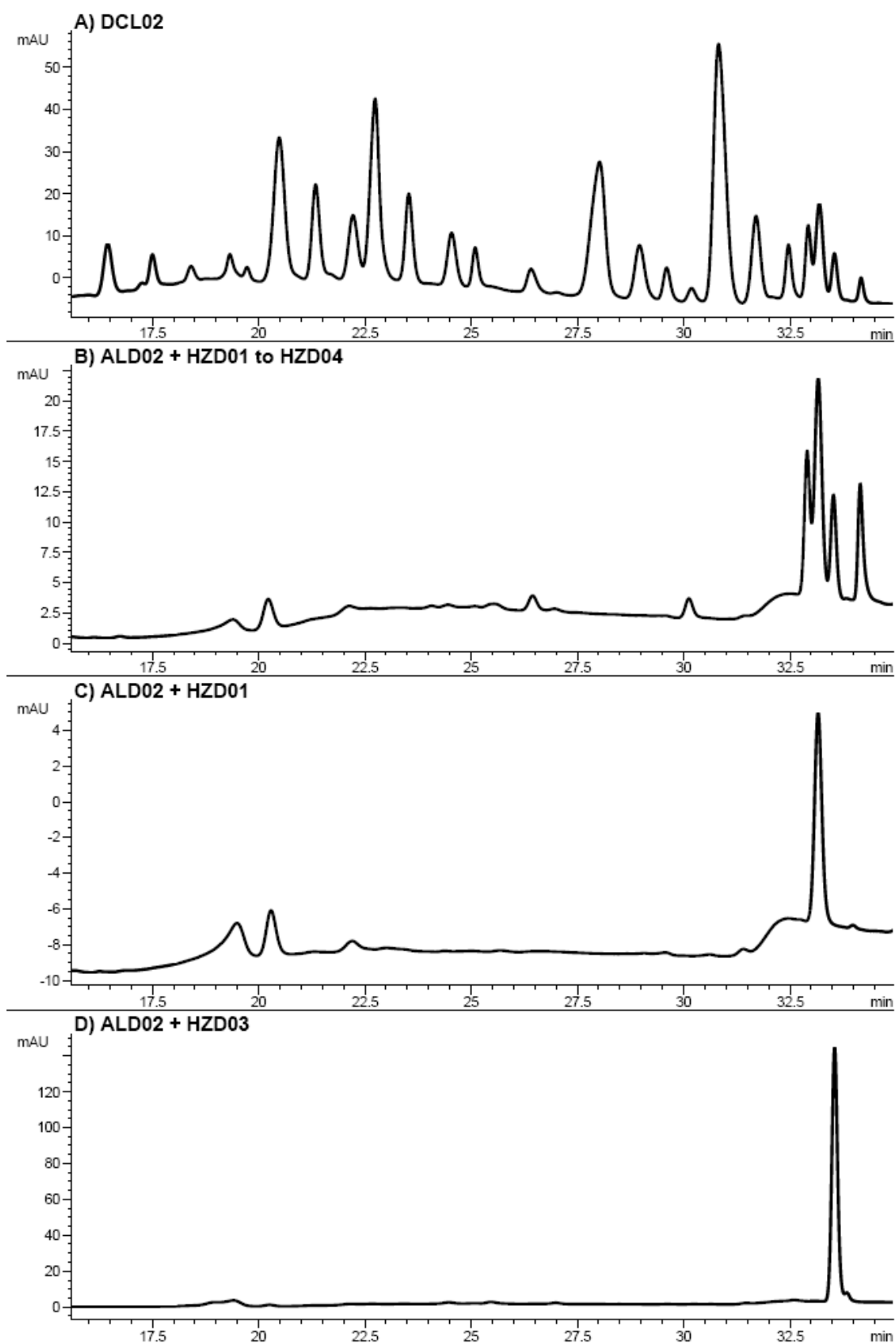


Figure 9: Deconvolution of **DCL02** to assign peaks. A) **DCL02**, B) **ALD02 + HZD01 to HZD04**, C) **ALD02 + HZD01**, D) **ALD02 + HZD03**.

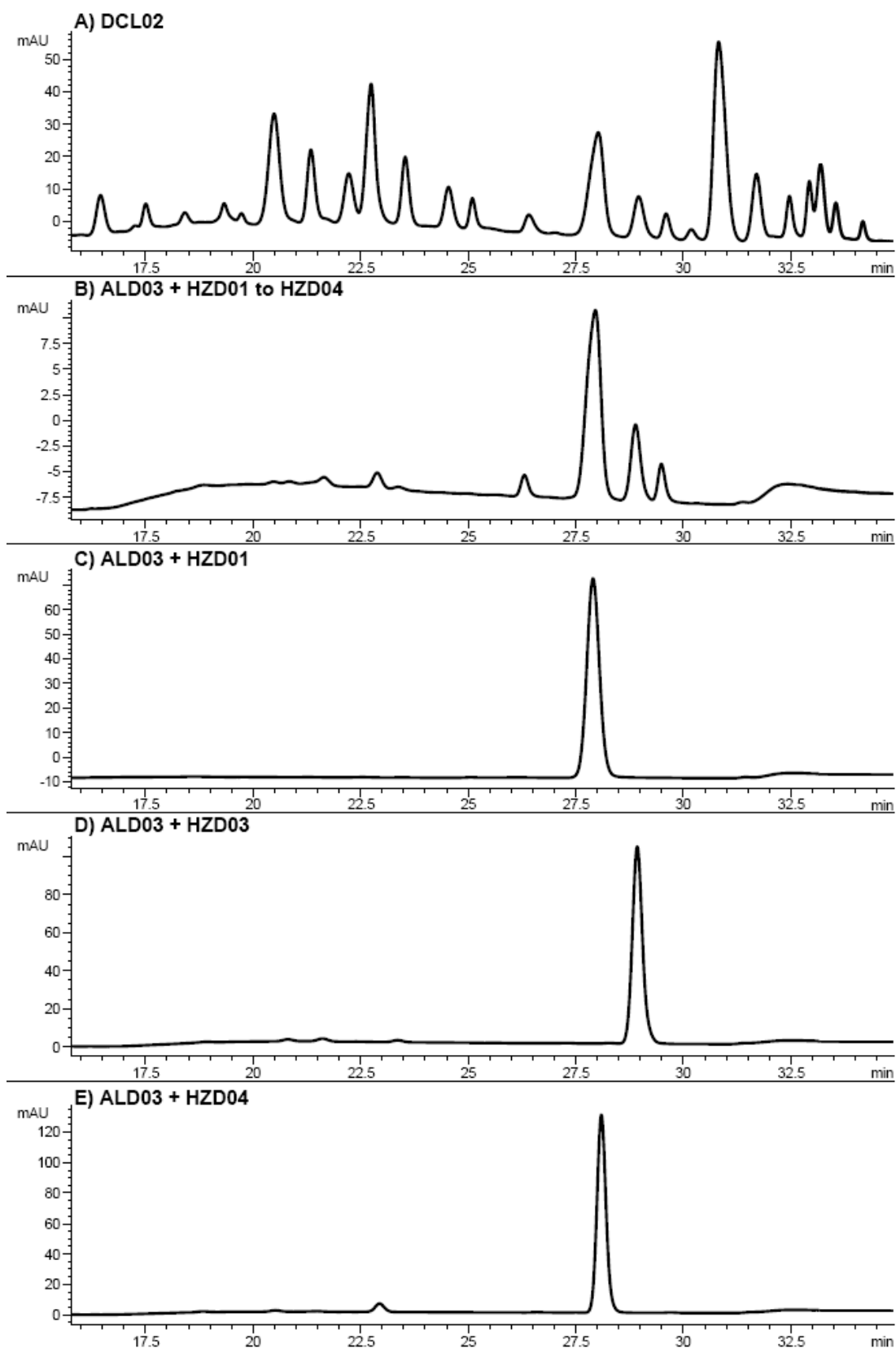


Figure 10: Deconvolution of **DCL02** to assign peaks. A) **DCL02**, B) **ALD03 + HZD01** to **HZD04**, C) **ALD03 + HZD01**, D) **ALD03 + HZD03**, E) **ALD03 + HZD04**.

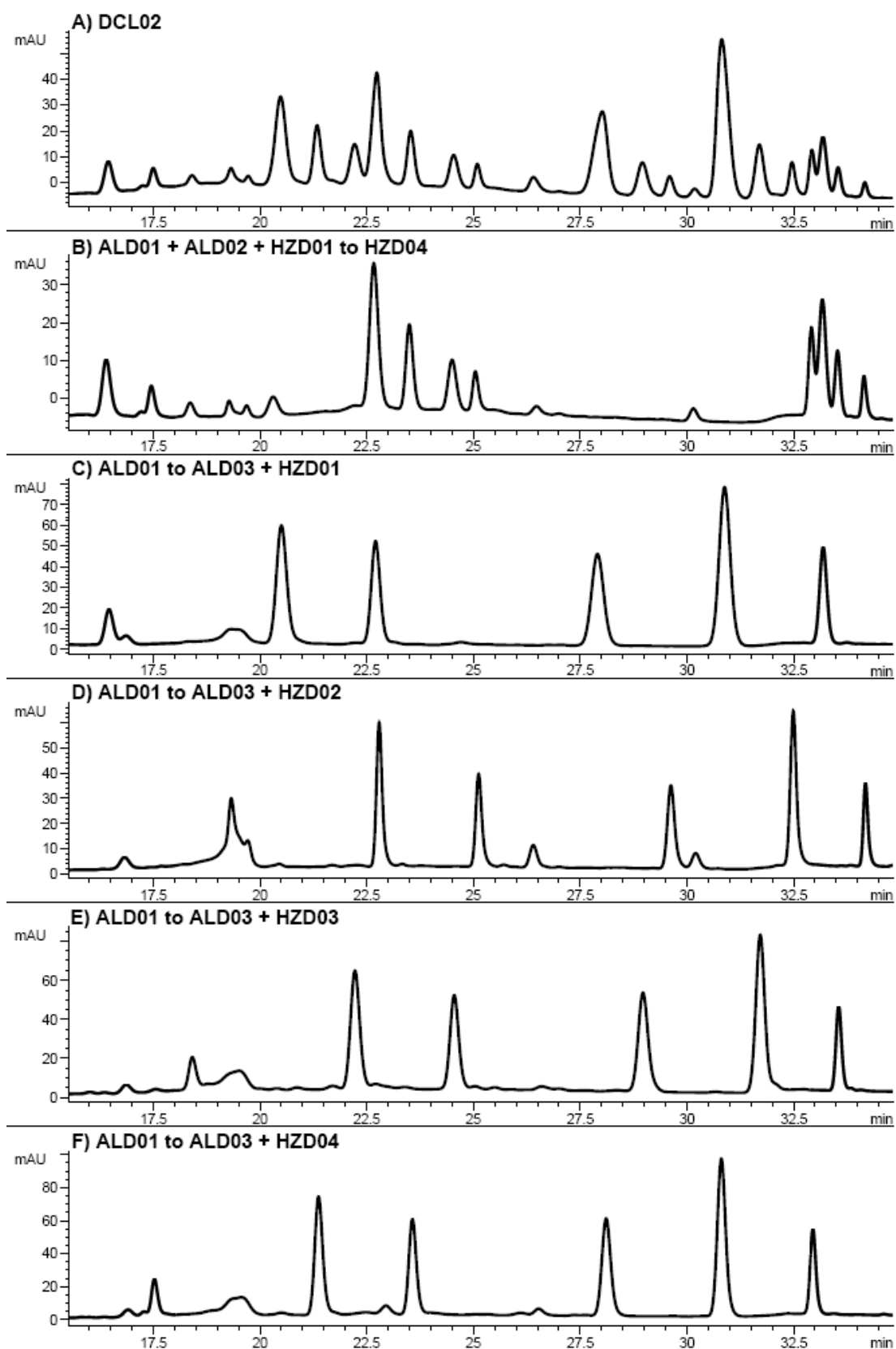


Figure 11: Deconvolution of DCL02 to assign peaks. A) DCL02, B) ALD01 + ALD02 + HZD01 to HZD04, C) ALD01 to ALD03 + HZD01, D) ALD01 to ALD03 + HZD02, E) ALD01 to ALD03 + HZD03, F) ALD01 to ALD03 + HZD04.

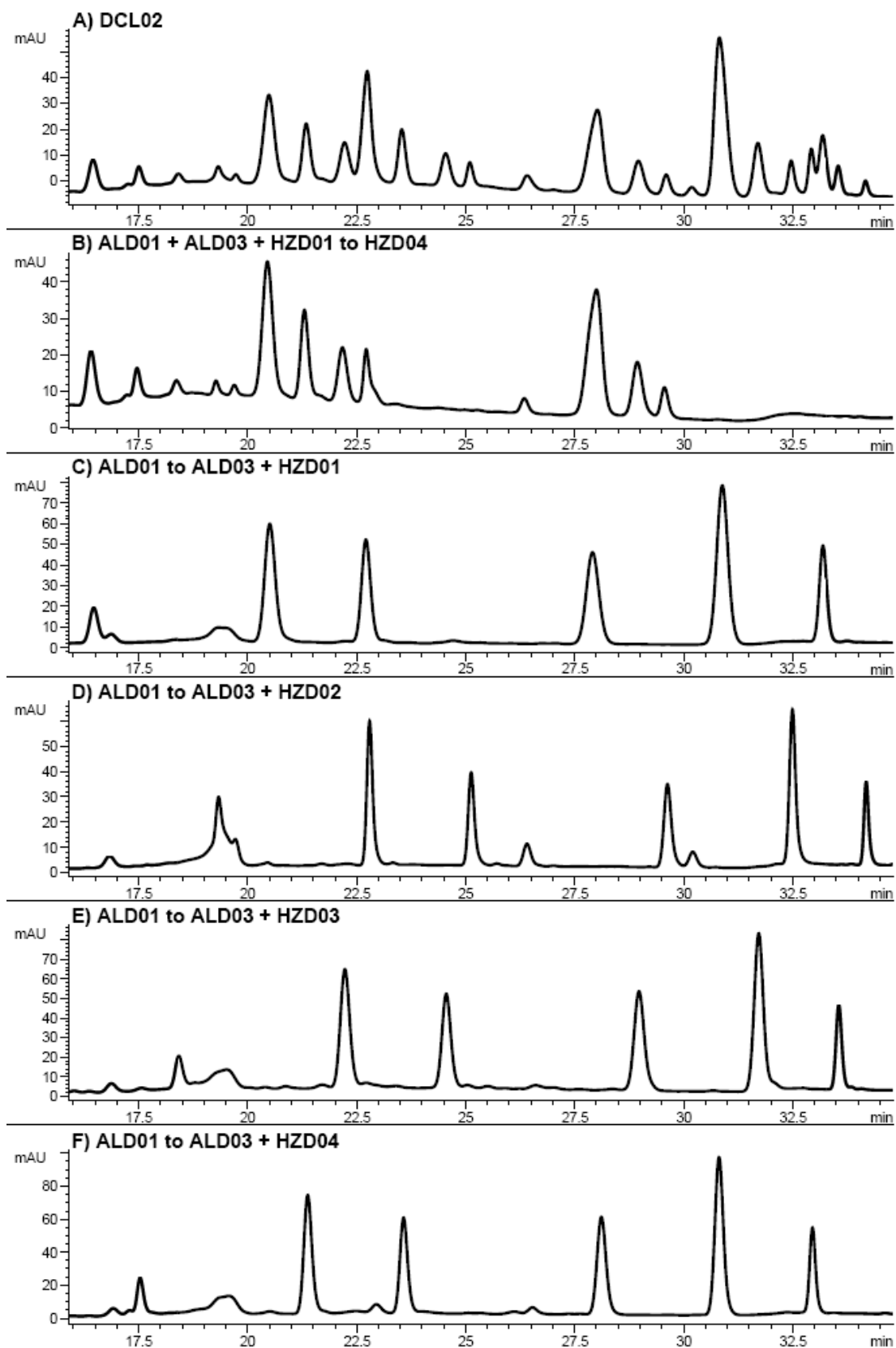


Figure 12: Deconvolution of **DCL02** to assign peaks. A) **DCL02**, B) **ALD01 + ALD03 + HZD01 to HZD04**, C) **ALD01 to ALD03 + HZD01**, D) **ALD01 to ALD03 + HZD02**, E) **ALD01 to ALD03 + HZD03**, F) **ALD01 to ALD03 + HZD04**.

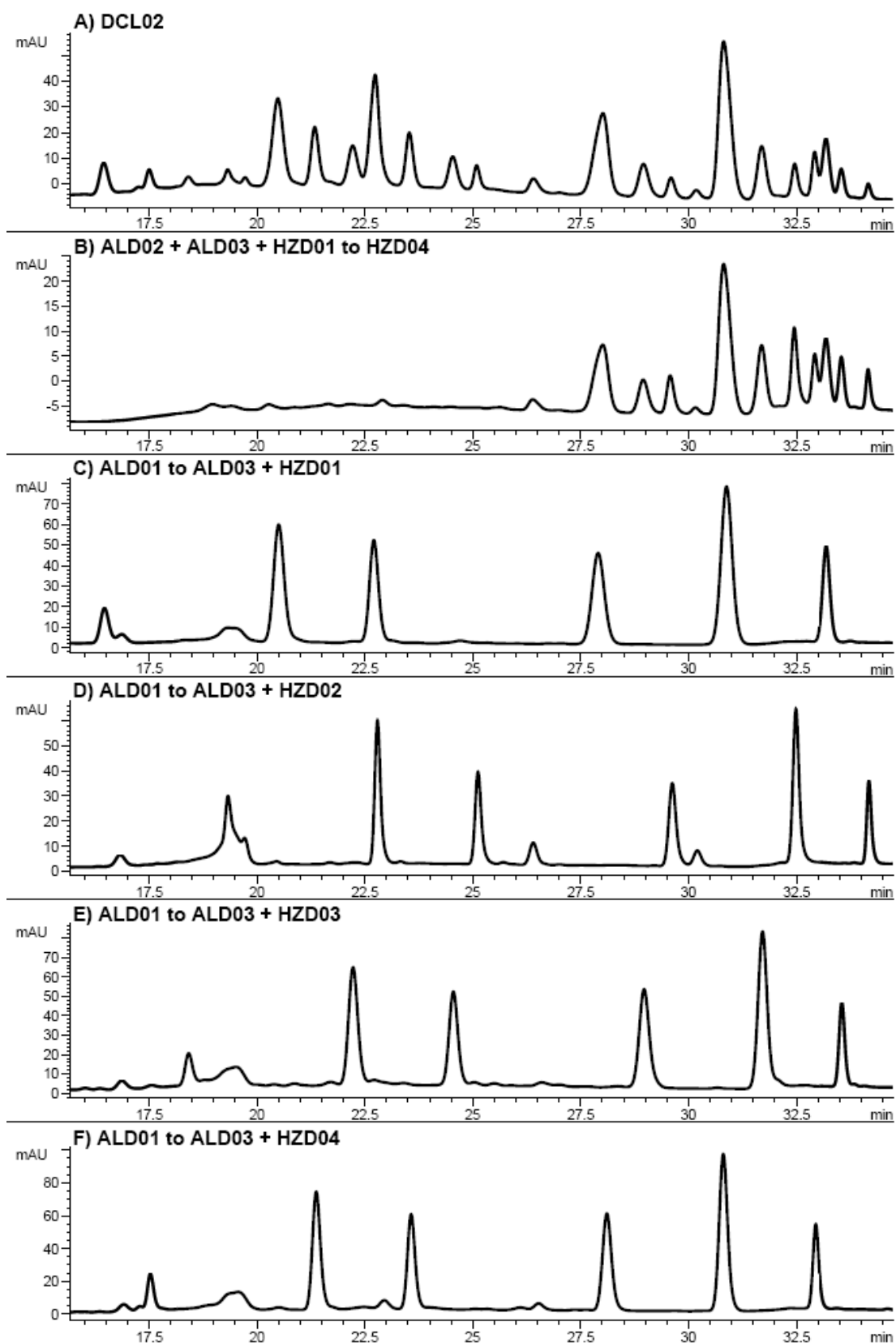


Figure 13: Deconvolution of DCL02 to assign peaks. A) DCL02, B) ALD02 + ALD03 + HZD01 to HZD04, C) ALD01 to ALD03 + HZD01, D) ALD01 to ALD03 + HZD02, E) ALD01 to ALD03 + HZD03, F) ALD01 to ALD03 + HZD04.

6.2 LC/MS-SIM traces from *in situ* click chemistry reactions

6.2.1 *In situ* click reactions between AZ01 and ALK04

Examples of the LC/MS-SIM traces for the *in situ* click chemistry reactions described in Chapter 3. The results for reactions containing the naphthalene-based alkyne **ALK04** have been shown.

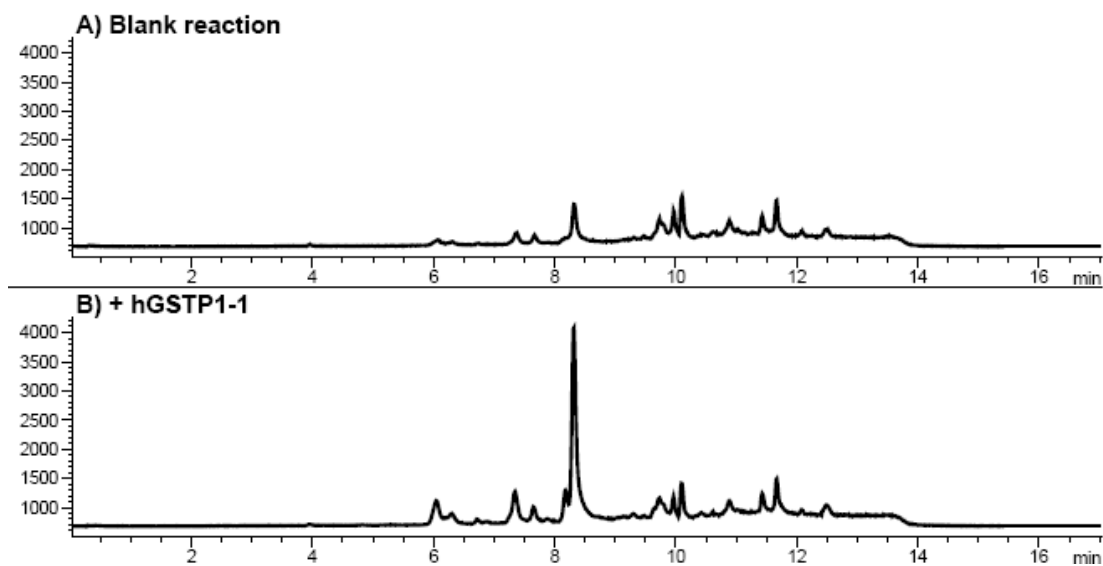


Figure 14: LC/MS-SIM traces for A) the background (blank) reaction and B) the hGSTP1-1-templated reaction between **AZ01** and **ALK04**.

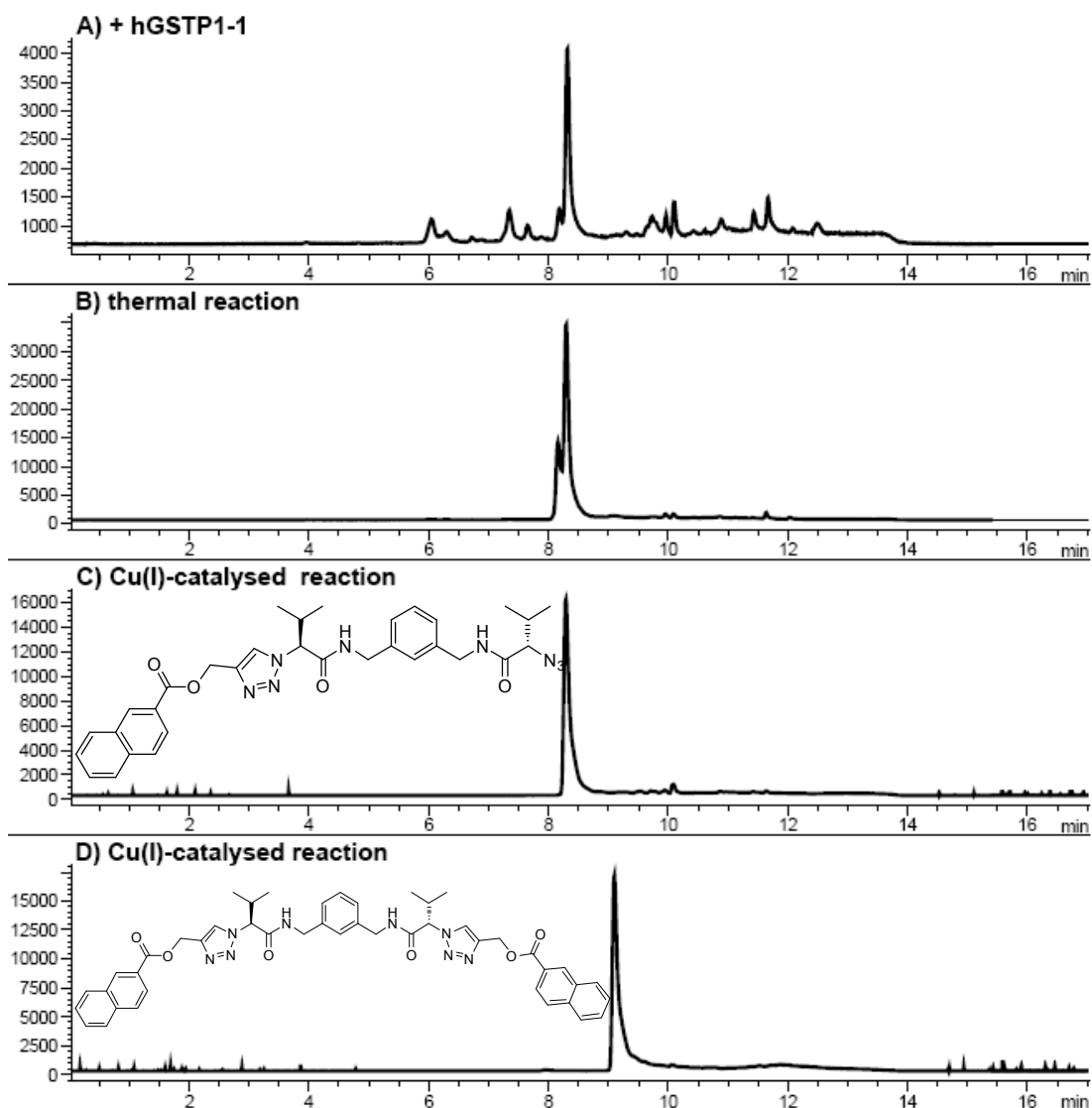


Figure 15: LC/MS-SIM traces for A) the hGSTP1-1-templated reaction between **AZ01** and **ALK04** and comparison with reference samples from B) the thermal reaction, C) Cu(I)-catalysed synthesis of monotriazole **TZ04a-01** and D) Cu(I)-catalysed synthesis of **TZ04a-01-04a**.

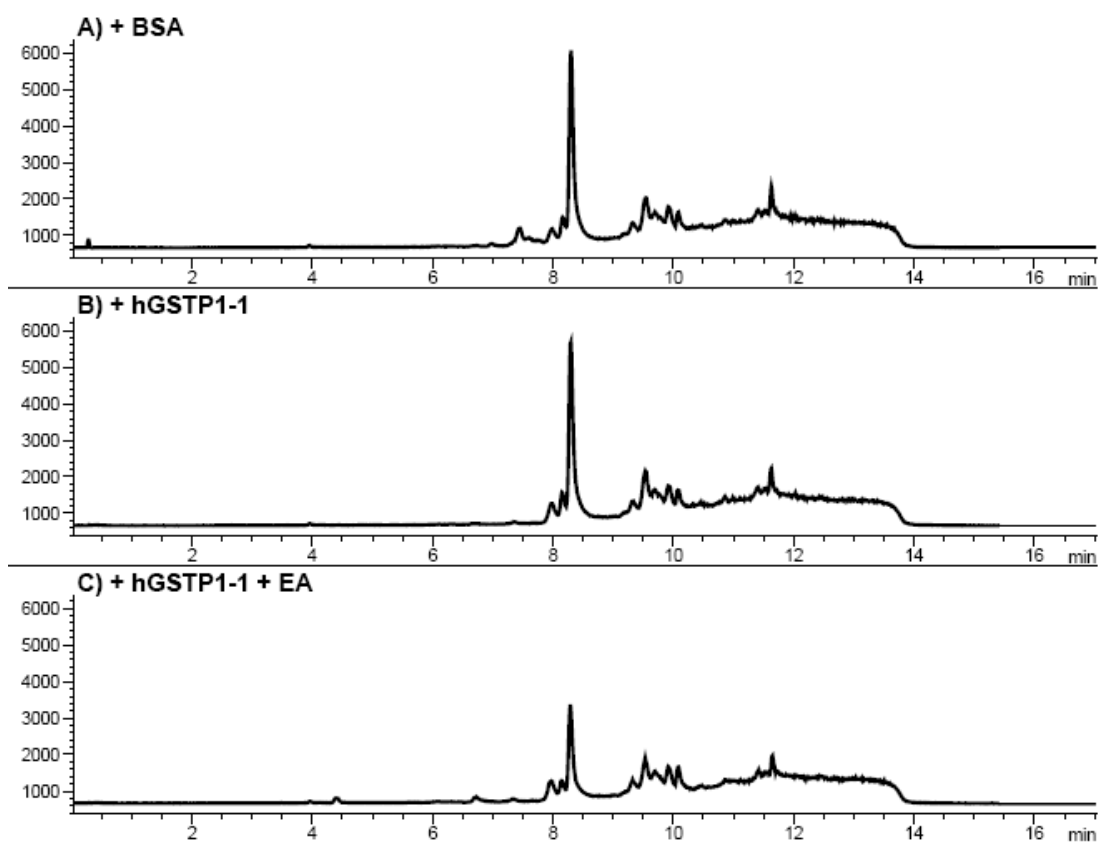


Figure 16: LC/MS-SIM traces for the reaction between **AZ01** and **ALK04**, A) in the presence of BSA, B) in the presence of hGSTP1-1 and C) in the presence of hGSTP1-1 with added inhibitor (ethacrynic acid).

6.2.2 *In situ* click reactions between TZ04a-01 and ALK04

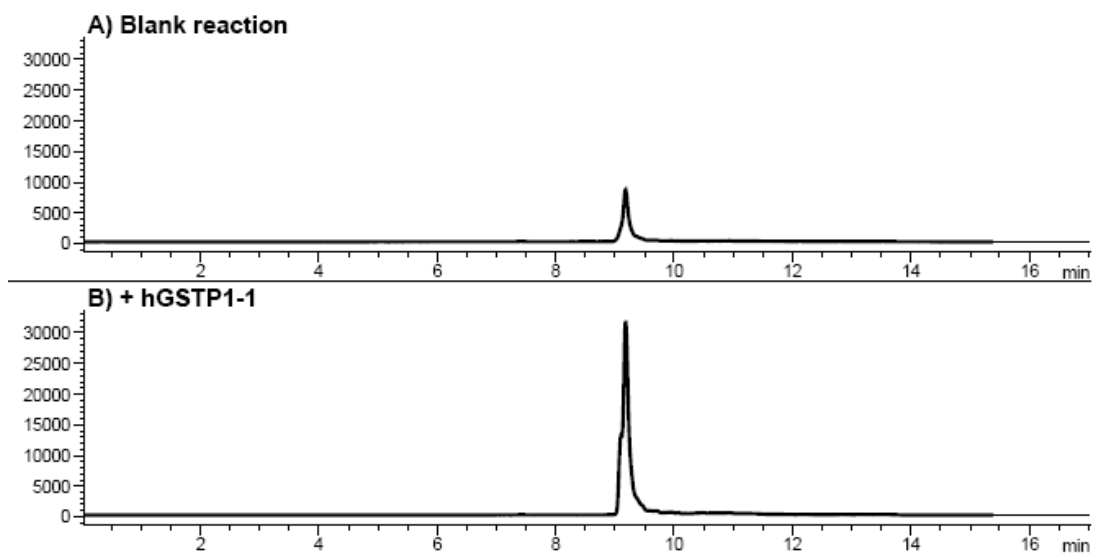


Figure 17: LC/MS-SIM traces for A) the background (blank) reaction and B) the hGSTP1-1-templated reaction between **TZ04a-01** and **ALK04**.

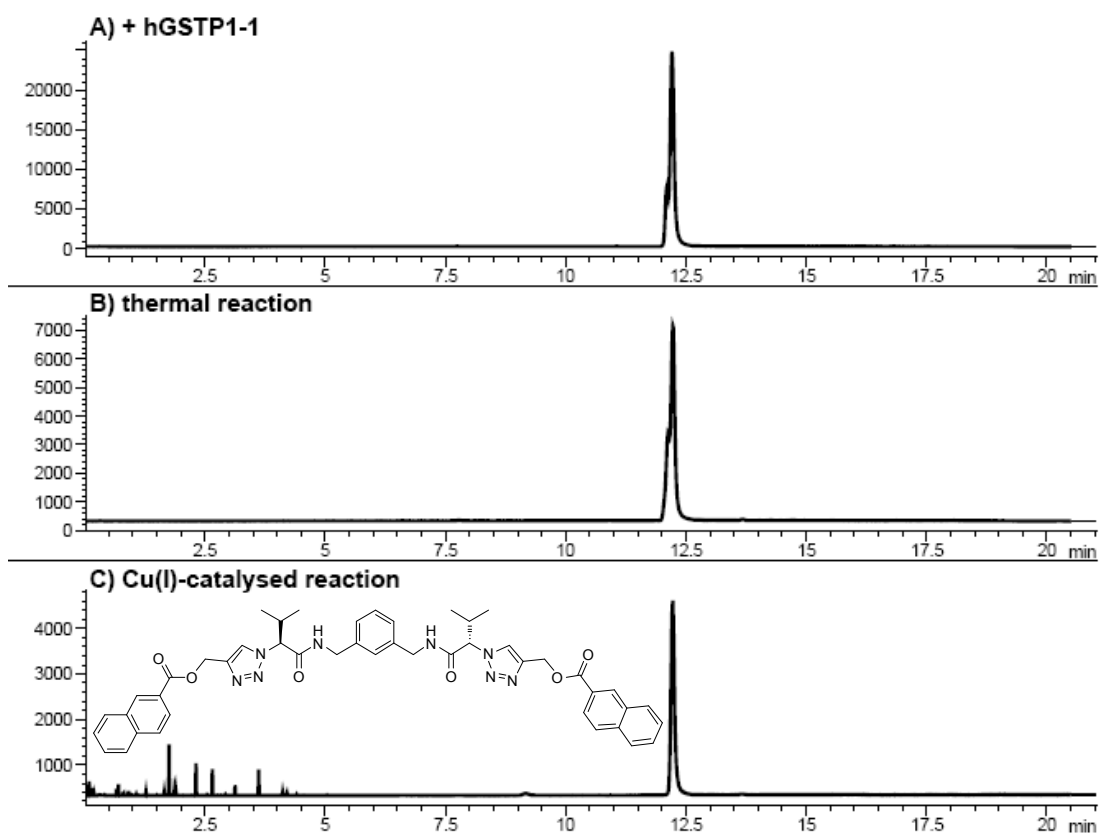


Figure 18: LC/MS-SIM traces for A) the hGSTP1-1-templated reaction between **TZ04a-01** and **ALK04** and comparison with reference samples from B) the thermal reaction and C) Cu(I)-catalysed synthesis of **TZ04a-01-04a**.

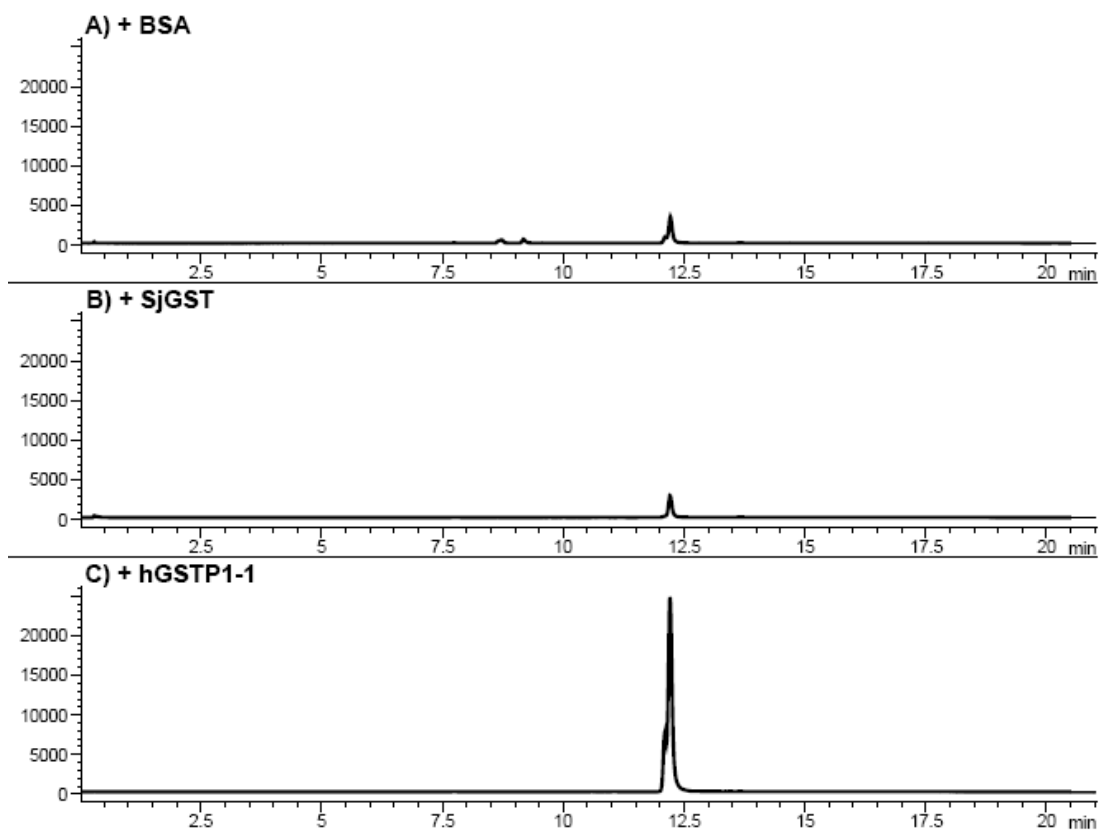


Figure 19: LC/MS-SIM traces for the reaction between **TZ04a-01** and **ALK04**, A) in the presence of BSA, B) in the presence of SjGST and C) in the presence of hGSTP1-1.

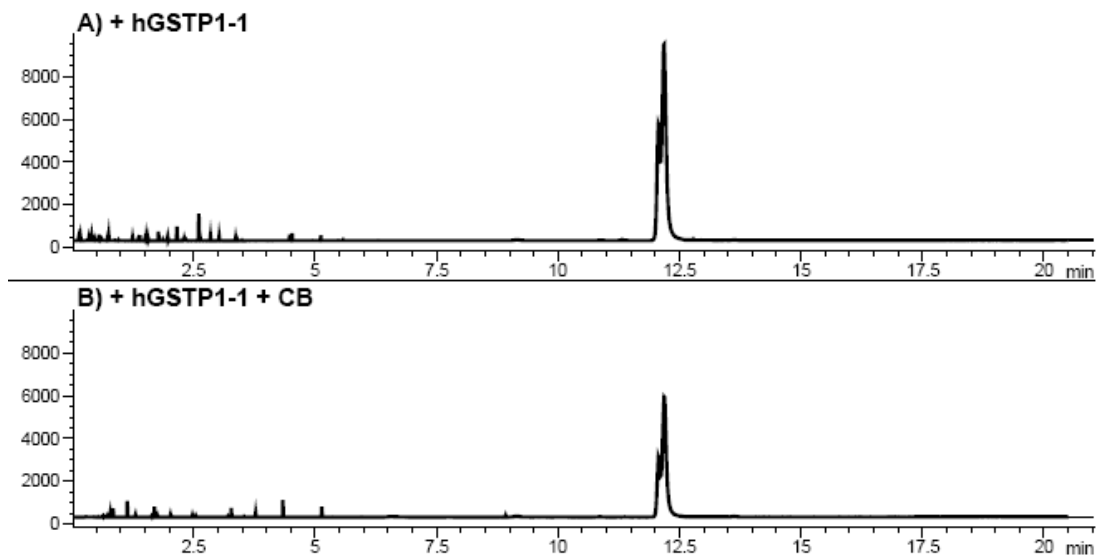


Figure 20: LC/MS-SIM traces for the reaction between **TZ04a-01** and **ALK04**, A) in the presence of hGSTP1-1 and B) in the presence of hGSTP1-1 with added inhibitor (cibacron blue).

6.2.3 Reference samples for *in situ* click reactions between azide AZ01 and alkyne library

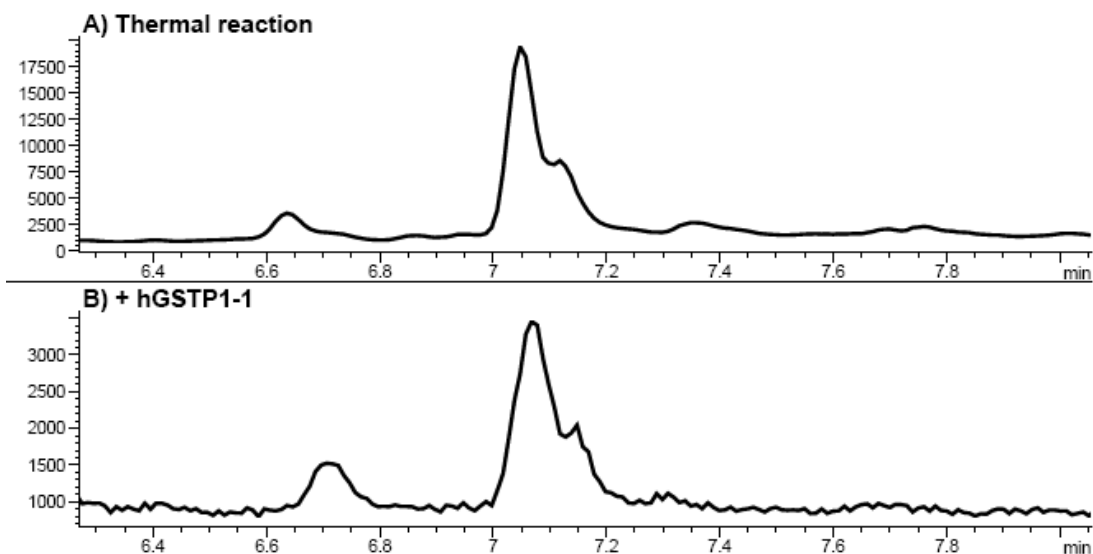


Figure 21: Reference for the reaction between **AZ01** and **ALK03**, SIM monitoring for mass of monotriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

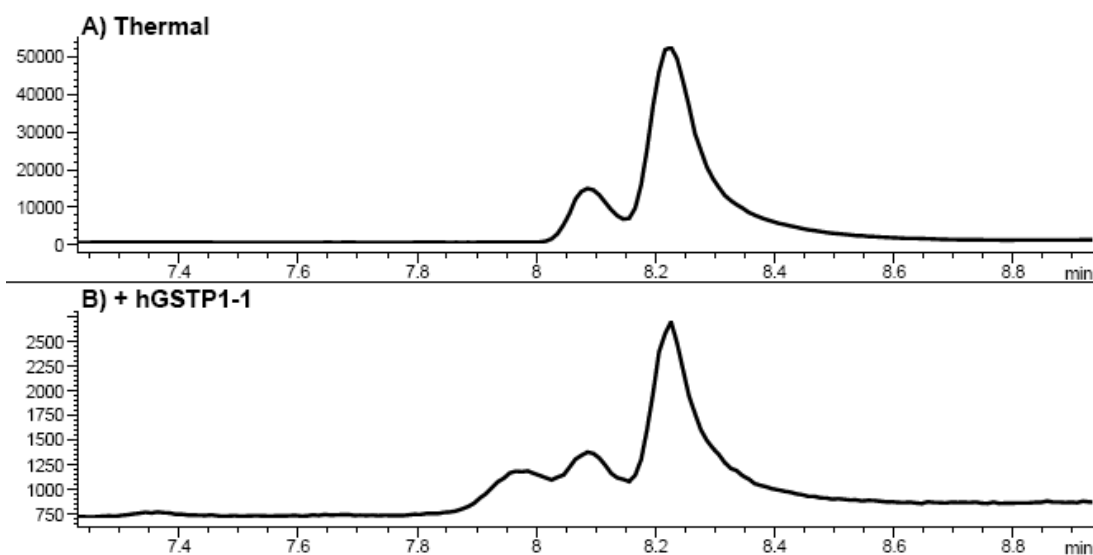


Figure 22: Reference for the reaction between **AZ01** and **ALK07**, SIM monitoring for mass of monotriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

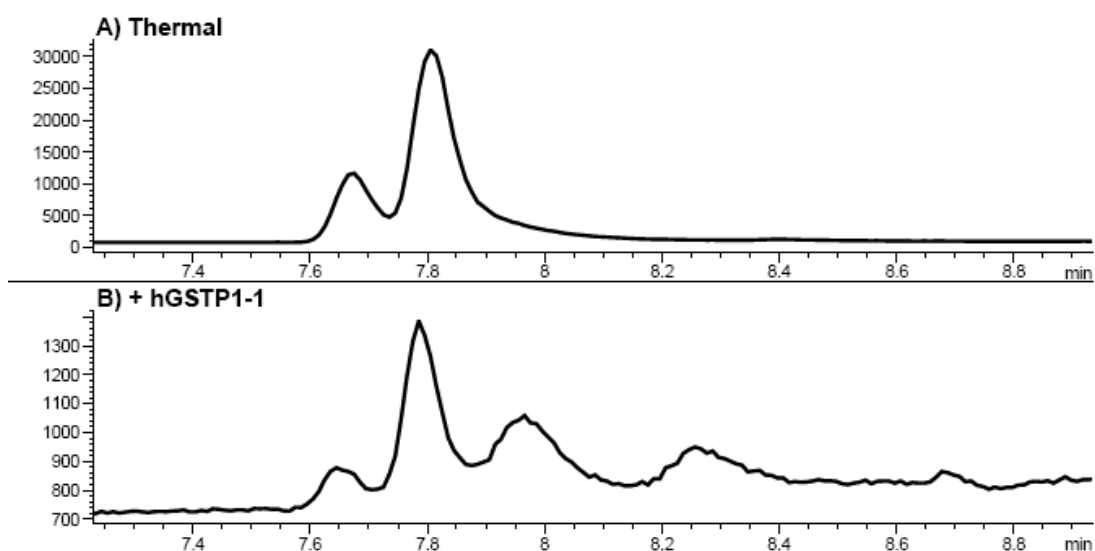


Figure 23: Reference for the reaction between **AZ01** and **ALK08**, SIM monitoring for mass of monotriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

6.2.4 Reference samples for *in situ* click reactions between azide **TZ04a-01** and alkyne library

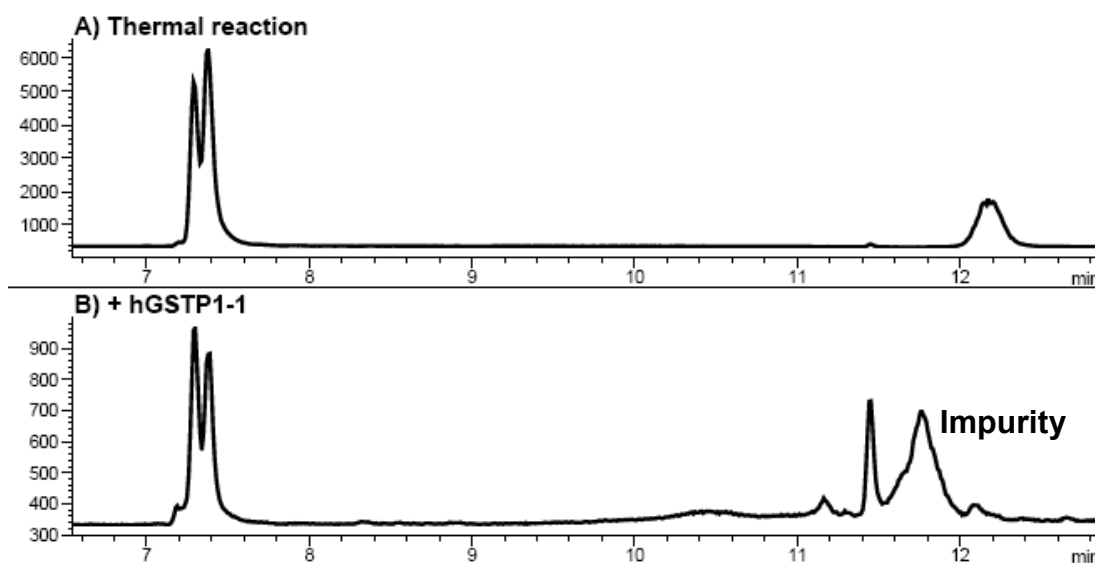


Figure 24: Reference for the reaction between **TZ04a-01** and **ALK01**, SIM monitoring for mass of ditriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

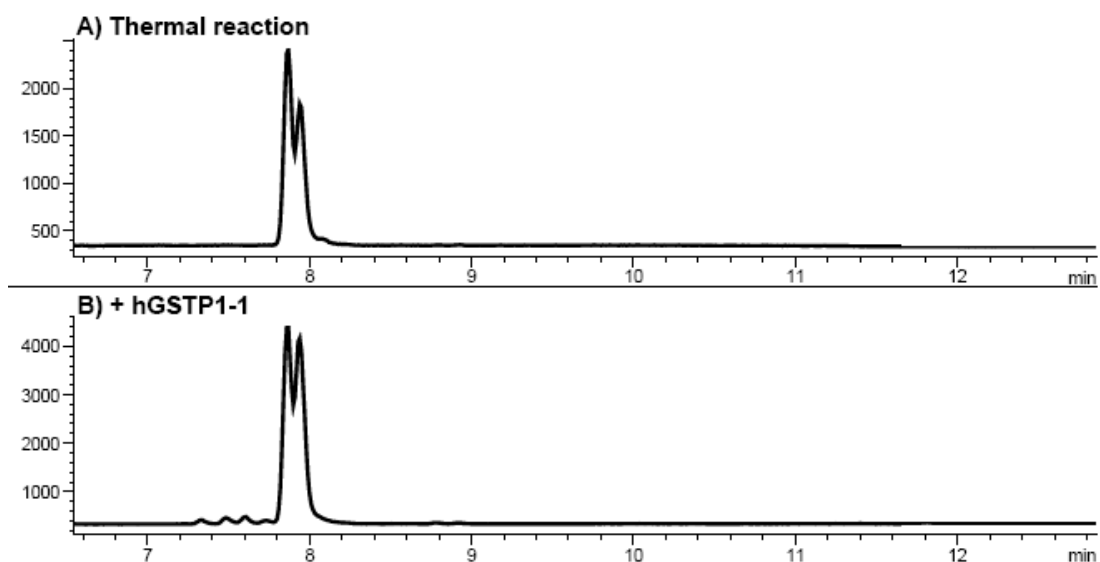


Figure 25: Reference for the reaction between **TZ04a-01** and **ALK03**, SIM monitoring for mass of ditriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

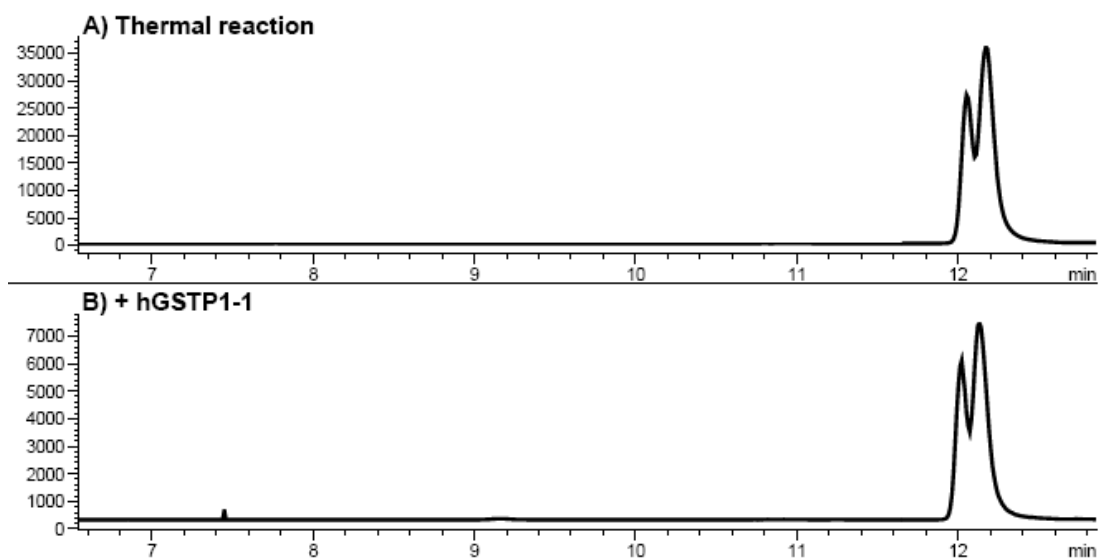


Figure 26: Reference for the reaction between **TZ04a-01** and **ALK07**, SIM monitoring for mass of ditriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

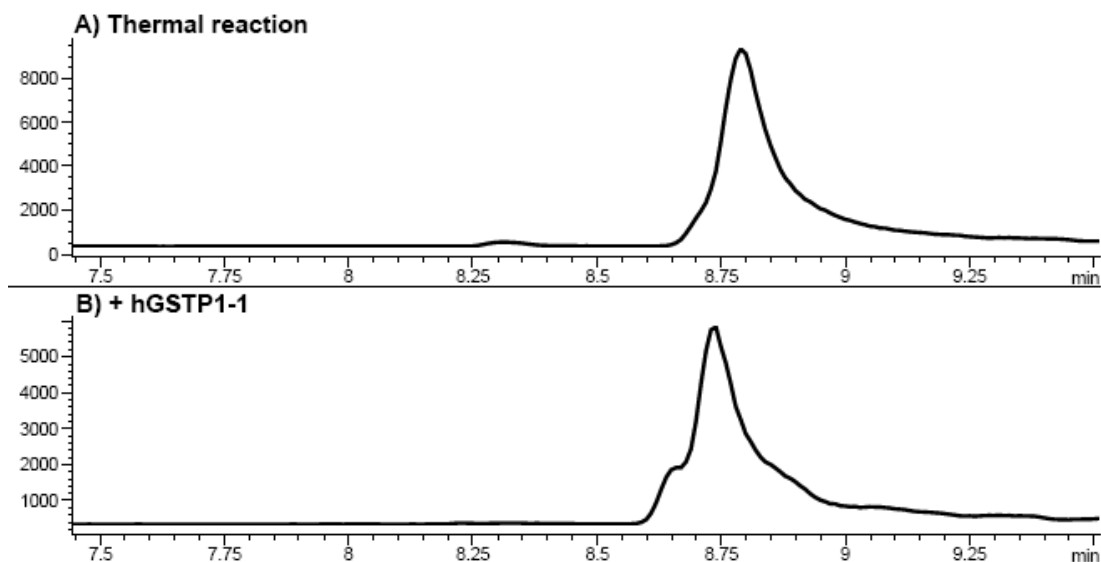


Figure 27: Reference for the reaction between **TZ04a-01** and **ALK08**, SIM monitoring for mass of ditriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

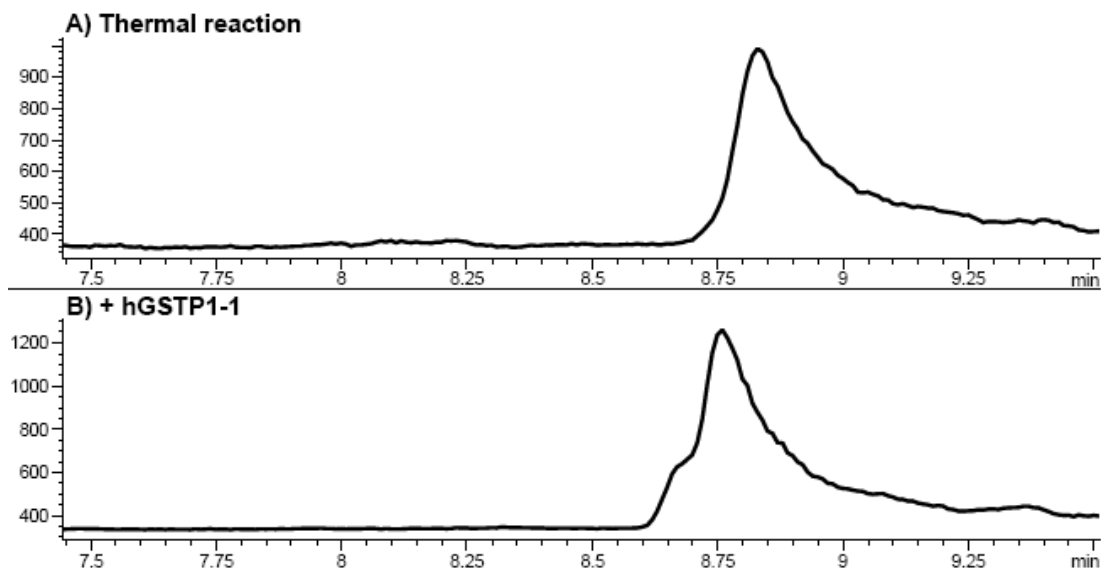


Figure 28: Reference for the reaction between **TZ04a-01** and **ALK09**, SIM monitoring for mass of ditriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

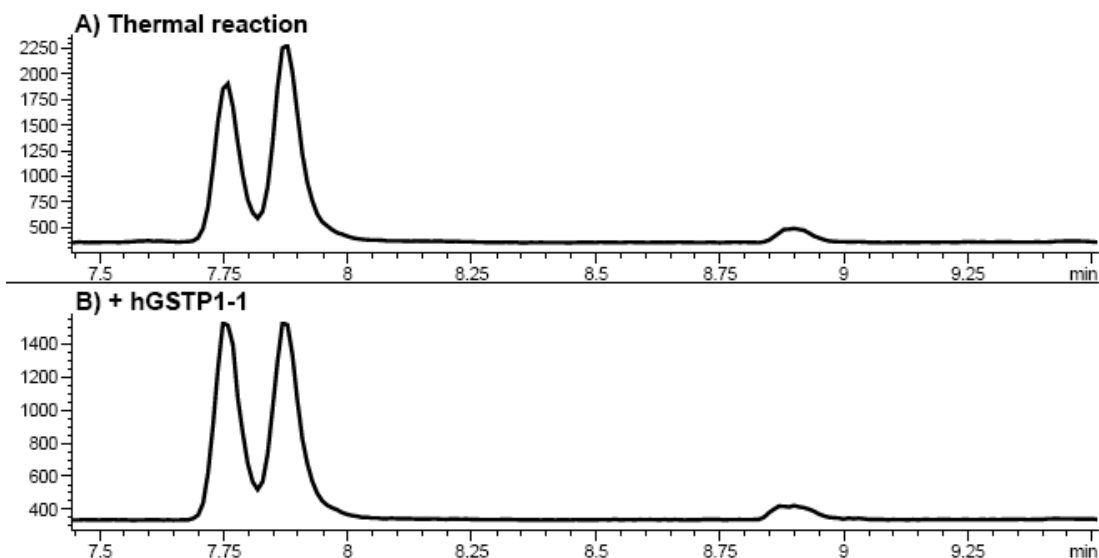


Figure 29: Reference for the reaction between **TZ04a-01** and **ALK11**, SIM monitoring for mass of ditriazole. A) product from the thermal reaction, B) product from the *in situ* reaction in the presence of hGSTP1-1.

6.3 Results from *in situ* click reactions between TZ04a-01 and alkyne hits using hGSTP1-1

The following figures are the results from the ISCC screen of the reaction between TZ04a-01 and the alkyne library using hGSTP1-1 as the biological template, as summarised in Chapter 3, Section 3.21.

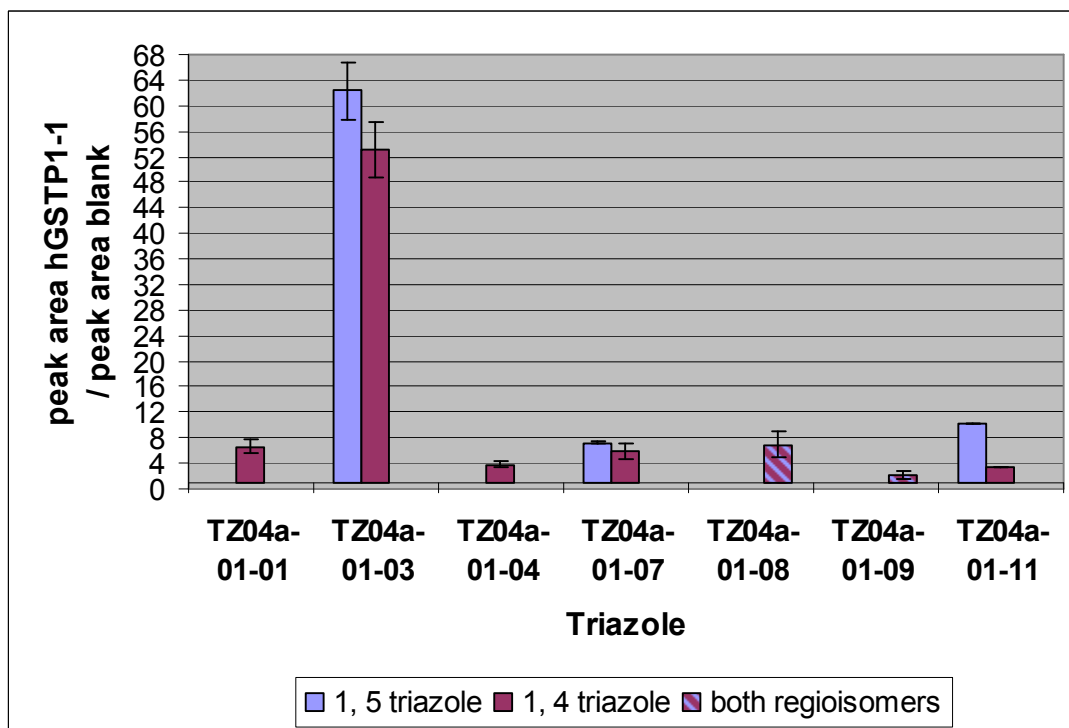


Figure 30: ISCC reaction between **TZ04a-01** and alkyne 'hits' in the presence of hGSTP1-1. Graph shows the ratio of triazole observed in the presence of GST over the background reaction. Error bars represent standard deviation.

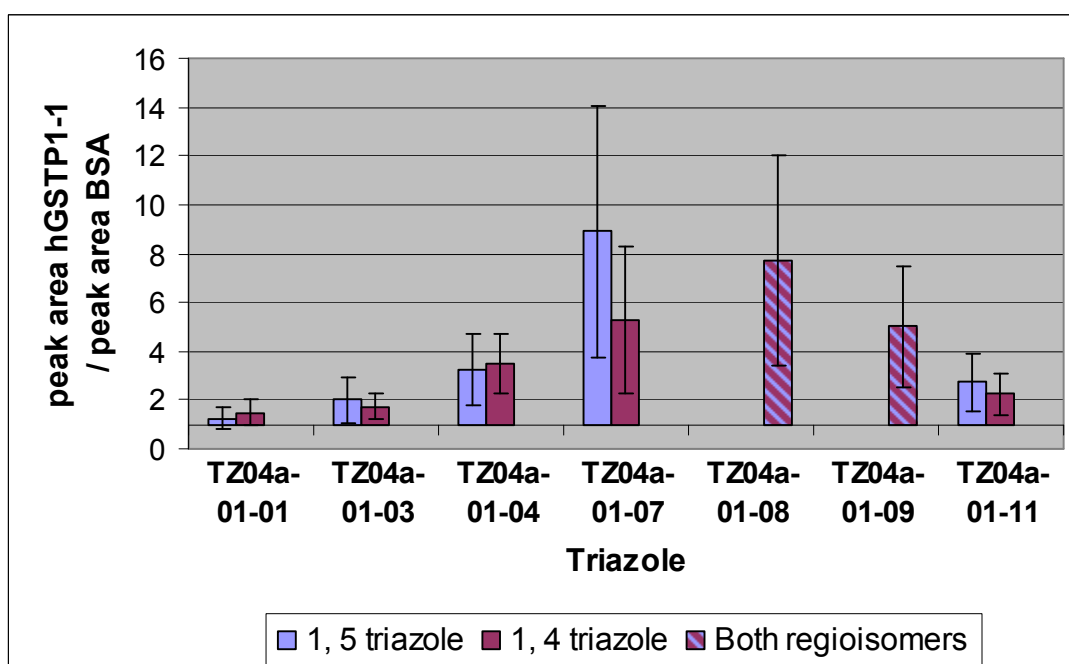


Figure 31: ISCC reaction between **TZ04a-01** and alkyne 'hits' in the presence of hGSTP1-1. Graph shows the ratio of triazole observed in the presence of GST over the reaction in the presence of BSA. Error bars represent standard deviation.

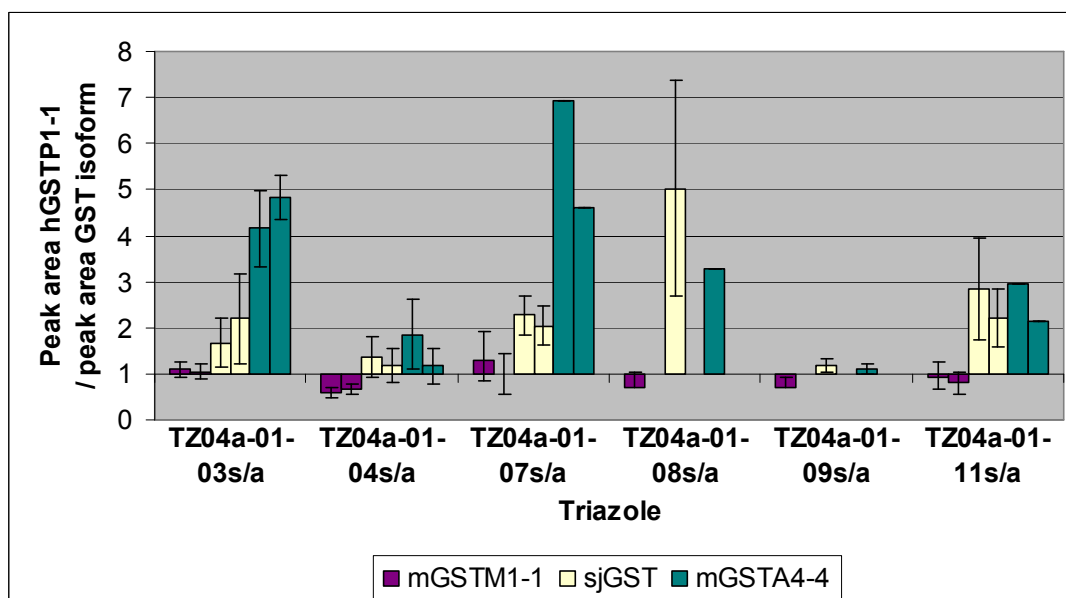


Figure 32: ISCC reaction between **TZ04a-01** and alkyne 'hits' in the presence of hGSTP1-1. Graph shows the ratio of triazole observed in the presence of hGSTP1-1 over the reaction in the presence of other GST isoforms. Error bars represent standard deviation.

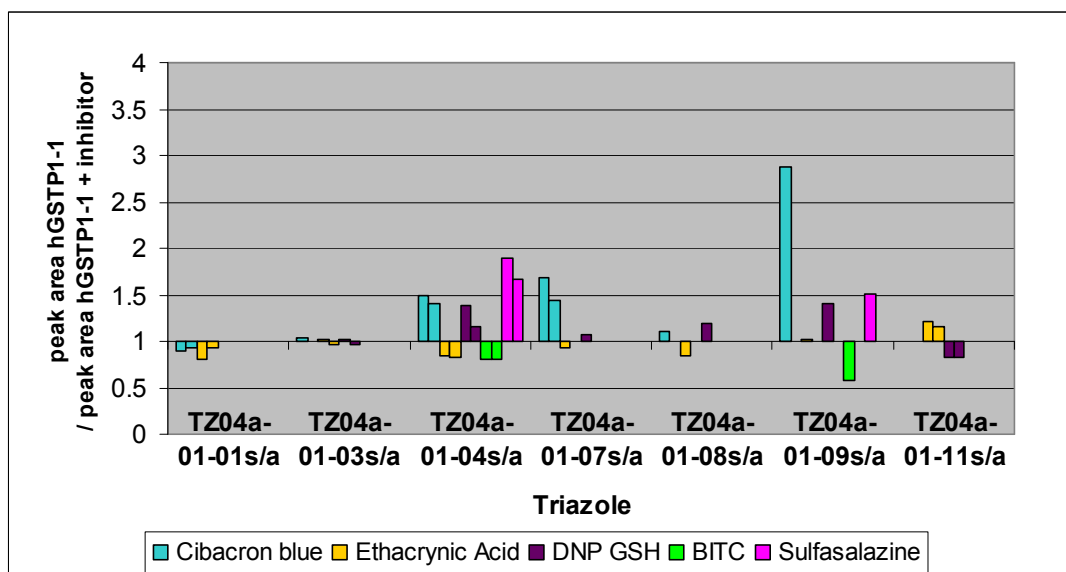


Figure 33: ISCC reaction between **TZ04a-01** and alkyne 'hits' in the presence of hGSTP1-1. Graph shows the ratio of triazole observed in the presence of hGSTP1-1 over the reaction in the presence of hGSTP1-1 plus GST inhibitor. Error bars represent standard deviation.