

# Towards Visible Light Switching of Peptide-DNA and Peptide-Protein Interactions

Ryan David Dean

A thesis submitted to  
Cardiff University  
for the degree of  
DOCTOR OF PHILOSOPHY



School of Chemistry  
Cardiff University  
September 2016

**DECLARATION**

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ..... (candidate)      Date .....

**STATEMENT 1**

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed ..... (candidate)      Date .....

**STATEMENT 2**

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ..... (candidate)      Date .....

**STATEMENT 3**

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate)      Date .....

**STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS**

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ..... (candidate)      Date .....

## Summary

Peptides derived from DNA-binding zinc finger proteins were synthesised with pairs of cysteine residues with  $i,i+7$  and  $i,i+11$  relative spacings introduced into their sequence. The sidechains of these cysteine residues were then alkylated with the well-known water soluble photochrome 3,3'-bis(sulfo)-4,4'-bis(chloroacetamino) azobenzene (BSBCA). The change of shape of the azobenzene dye in these peptide-dye conjugates allowed photocontrol of peptide structure and thus peptide-DNA interactions. For a single zinc finger helix, UV irradiation yielded a peptide conjugate with a dissociation constant with respect to its cognate DNA sequence of 100 nM with no binding apparent prior to irradiation. However, the relatively short half-life of BSBCA proved a stumbling block, particularly in the control of larger peptides using multiple azobenzenes to control several  $\alpha$ -helical structural elements within large peptides. In addition to the short half-life of *cis*-BSBCA under physiological conditions, multiple BSBCA switches attached to the same peptide were shown not to relax independently of each other.

These results led to the design, synthesis and examination of novel photo switches sensitive to visible, rather than UV light, with improved light state half-lives and bidirectional optical switching. Initial studies on thioindigo-based switches proved that molecules sufficiently polar to be water soluble were inaccessible by concise synthetic routes. Attention was then turned to the synthesis of *ortho*-halogen substituted azobenzenes and investigation of several new conjugation strategies for linking these photosensitive molecules to peptides.

Subsequent refinements to the design of the tetra-*ortho*-halogen substituted azobenzenes altered the position of UV/visible absorbance bands of the *cis* and *trans* isomers to create a 47 nm separation in the wavelengths of the  $n-\pi^*$  absorbances of the isomers to allow effective bidirectional switching. These changes also improved the half-life of the *cis* state from 24 minutes at 20 °C to 3,256 minutes at 60 °C.

One of these new azobenzenes was reacted with apoptosis-inducing Bak peptides with different cysteine spacings ( $i,i+7$  and  $i,i+11$ ). Less stringent control over the binding of these peptides to Bcl-X<sub>L</sub> was observed than with BSBCA, perhaps due to the more flexible nature of the new crosslinker, but the optical properties of this class of molecules suggest a little further development will yield much improved photoswitches.

## Abbreviations

BSBCA	3,3'-bis(sulfo)-4,4'-bis(chloroacetamino) azobenzene
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
mRNA	messenger ribonucleic acid
UV	ultraviolet
HGH	human growth hormone
LOV	light, oxygen voltage
AsLOV	<i>Avena sativa</i>
CD	circular dichroism
ATP	adenosine triphosphate
LOV	light oxygen voltage
AsLOV	<i>Avena sativa</i> light oxygen voltage
FMN	flavin mononucleotide
LUMO	lowest unoccupied orbital
HOMO	highest unoccupied molecular orbital
CAP	catabolite activator protein
<i>E. coli</i>	<i>Escherichia coli</i>
Cys	cysteine
His	histidine
ZF-TFs	zinc finger transcription factors
ZFN	zinc finger nucleases
NHEJ	non-homologous end joining
DSB	double strand breaks
EMT	epithelial to mesenchymal transitions
Fmoc	fluorenylmethyloxycarbonyl
TFA	trifluoro acetic acid

MALDI-TOF	matrix-assisted laser desorption ionisation -time of flight
SD	standard deviation
DMSO	dimethyl sulfoxide
EMSA	electronic mobility gel shift assays
ITC	isothermal titration calorimetry
nm	nano meter
HCl	hydrochloric acid
NBS	<i>N</i> -bromosuccinimide
TBDMS	tert-butyldimethylsilyl ethers
NMR	nuclear magnetic resonance
MS	mass spectrometry
LC-MS	liquid chromatography mass spectrometry
THF	tetrahydrofuran
DMF	dimethylformamide
HPLC	high performance liquid chromatography
DIBAL	di- <i>iso</i> -butyl aluminium hydride
IR	infra-red

## Acknowledgements

I would like to thank Rudolf Allemann for the opportunity to carry out this Phd.

With special thanks to Robert Mart for all his help knowledge and his never ending patience with the writing of this thesis. I would also like to thank Roger Glanville, Sarah Adams and Chris Jones for their help and support during my phd mainly in the pub.

---

## Table of Contents

1	Introduction .....	1
1.1	Targeting protein-protein interactions .....	1
1.2	Methods for photo control of peptides.....	4
1.2.1	Photocages .....	4
1.3	Reversible photocontrol.....	7
1.3.1	Photoresponsive proteins.....	7
1.3.2	Spiropyrans .....	8
1.3.3	Diarylethenes .....	10
1.3.4	Thioindigo.....	11
1.4	Azobenzene .....	13
1.4.1	Mechanism of <i>trans</i> to <i>cis</i> conversion .....	15
1.4.2	Solvent effects on azobenzene spectra .....	16
1.4.3	Tuning the absorbance spectrum of azobenzene.....	16
1.5	Effects of substitution on azobenzene.....	17
1.5.1	Substitution of azobenzenes with heteroatoms at the <i>ortho</i> position .....	17
1.5.2	Tetra- <i>ortho</i> -haloazobenzenes .....	21
1.5.3	Additional effects of <i>para</i> substituents .....	24
1.5.4	Push-pull azobenzenes.....	27
1.5.5	Glutathione reduction .....	27
1.6	Azobenzene to control biological activity .....	27
1.6.1	Photopharmacology.....	27
1.6.2	Azobenzene as a steric switch.....	28
1.6.3	Genetically encoded incorporation of azobenzene .....	30
1.6.4	Conjugating azobenzenes to pairs of cysteine sidechains .....	32
1.7	Examples of photocontrol of transcription using azobenzenes .....	37
2	Zinc finger peptides.....	40
2.1	DNA transcription .....	40
2.1.1	Prokaryotic transcription control .....	41
2.1.2	Eukaryotic transcription control.....	42
2.2	Zinc finger transcription factors .....	44

---

2.2.1	Designed zinc finger proteins .....	45
2.3	Photo activated zinc finger transcription factors.....	46
2.3.1	Slug and snail .....	46
2.4	Existing photo activated zinc finger peptides .....	46
2.4.1	Design of the Slug finger peptides .....	47
2.4.2	Solid state peptide synthesis .....	49
2.4.3	Synthesis of sulfonated azobenzene crosslinker .....	50
2.4.4	Purification of slug peptides.....	57
2.5	Circular dichroism.....	65
2.6	Fluorescence anisotropy.....	68
2.7	Conclusions.....	71
3	Visible light photoswitches.....	74
3.1	Introduction.....	74
3.2	Thioindigo.....	74
3.2.1	Retrosynthesis of thioindigo derivatives.....	77
3.3	Trifluoromethyl derived photo switches .....	82
3.3.1	Retrosynthetic analysis and synthesis .....	83
3.3.2	Spectra of halotrifluoromethyl azobenzene photoswitches .....	86
3.4	Functionalising azobenzene to attach to a peptide .....	90
3.5	Alternatives to haloacetamides.....	95
3.5.1	Sulfur substitution attempts .....	98
3.5.2	Oxygen substitution .....	100
4	Refining photoswitch properties .....	106
4.1	Tetra- <i>ortho</i> -chloro-di- <i>para</i> -hydroxymethylenylazobenzene .....	106
4.2	Synthesis of tetra- <i>ortho</i> -haloazobenzenes .....	114
4.3	UV/visible spectra of tetra halogenated azobenzenes.....	116
4.4	Di- <i>ortho</i> -haloazobenzenes.....	120
4.5	Mixed halogens.....	126
4.6	Attaching to a peptide.....	132
4.7	Preventing hydrolysis.....	134
4.8	Attachment to cysteine .....	141
5	Conclusions .....	144
6	Methods.....	147



6.1	General chemical methods.....	147
6.1.1	Half-life Calculations.....	147
6.1.2	Crosslinking of Slug.....	149
6.1.3	HPLC.....	149
6.1.4	Solid Phase Synthesis.....	149
6.1.5	Circular Dichroism Spectroscopy.....	150
6.1.6	Fluorescence Anisotropy.....	150
6.1.7	Binding Calculations.....	150
6.1.8	Azobenzene (27).....	152
6.1.9	Bis-4,4'-chloroacetamidoazobenzene (49).....	152
6.1.10	4-Chloro-3-nitrobenzene sulfonate (109).....	153
6.1.11	4-Chloro-3-nitrobenzenesulfonamide (113).....	154
6.1.12	4[(3-Chloro-4-nitrophenol) sulphonyl] morpholine (114 and 115).....	154
6.1.13	Methyl 2-((5-(morpholinosulfonyl)-2-nitrophenyl)thio)acetate (116) ..	156
6.1.14	Methyl 2-((5-(morpholinosulfonyl)-2-nitrophenyl)thio)acetic acid (117)	157
6.1.15	2,4-Dibromo-6-trifluoromethylaniline (133).....	157
6.1.16	Tetra-2,2',4,4'-bromo-6,6'-di(trifluoromethyl)-azobenzene (125).....	158
6.1.17	4-Bromo-2-trifluoromethylaniline (129).....	159
6.1.18	Di-4,4'-bromo-2,2'-di(trifluoromethyl)-azobenzene (121).....	159
6.1.19	4-Bromo-2-chloro-6-trifluoromethylaniline (130).....	160
6.1.20	4,4'-Dibromo-2,2'-dichloro-6,6'-di(trifluoromethyl)-azobenzene (126)	160
6.1.21	2-Chloro-6-trifluoromethylaniline (128).....	161
6.1.22	4,4'-Dichloro-2,2'-di(trifluoromethyl)-azobenzene (124).....	161
6.1.23	2,4-Dichloro-6-trifluoromethylaniline (134).....	162
6.1.24	2-Bromo-4-chloro-6-trifluoromethylaniline (131).....	163
6.1.25	2,2'-Dibromo-4,4'-dichloro-6,6'-di(trifluoromethyl)-azobenzene (123)	163
6.1.26	2,2',4,4'-Tetrachloro-6,6'-di(trifluoromethyl)-azobenzene (122).....	164
6.1.27	4-Bromo-2,6-difluoroaniline (136).....	164
6.1.28	1,2-Bis(4-bromo-2,6-difluorophenyl)diazene (137).....	165
6.1.29	(E)-2,2'-((diazene-1,2-diylbis(3,5-difluoro-4,1-phenylene))bis(sulfanediyl))bis(ethan-1-ol) (138).....	166
6.1.30	2-((Tert-butyldimethylsilyl)oxy)ethane-1-thiol (140).....	167

6.1.31	(1,2-Bis(4-((2-((tert-butyldimethylsilyl)oxy)ethyl)thio)-2,6-difluorophenyl)diazene (141).....	168
6.1.32	3-Bromo-4-(phenylthio)tetrahydrothiophene 1,1-dioxide (144) .....	169
6.1.33	3-(Phenylthio)-2,5-dihydrothiophene 1,1-dioxide (145) .....	169
6.1.34	4-Bromo-3-(phenylthio)-2,3-dihydrothiophene 1,1-dioxide (146) and 3-bromo-4-(phenylthio)-2,5-dihydrothiophene 1,1-dioxide (147).....	170
6.1.35	4-Bromo-3-(phenylthio)-2,3-dihydrothiophene 1,1-dioxide (147) .....	170
6.1.36	3-Bromo-4-(phenylthio)-2,5-dihydrothiophene 1,1-dioxide (146) .....	171
6.1.37	3-(Phenylthio)-2,3-dihydrothiophene 1,1-dioxide (142) .....	171
6.1.38	3-Bromo-4-hydroxytetrahydrothiophene 1,1-dioxide (152).....	172
6.1.39	3-Hydroxy-2,3-dihydrothiophene 1,1-dioxide (153) .....	172
6.1.40	3,4-Dibromotetrahydrothiophene 1,1-dioxide (160).....	173
6.1.41	3-Bromo-2,3-dihydrothiophene 1,1-dioxide (157) .....	173
6.1.42	3-(Benzylamino)-2,3-dihydrothiophene 1,1-dioxide (159).....	174
6.1.43	( <i>E</i> )-4,4'-(Diazene-1,2-diyl)bis(3,5-dibromobenzonitrile) (164) .....	175
6.1.44	( <i>E</i> )-4,4'-(Diazene-1,2-diyl)bis(3,5-dichlorobenzonitrile) (168).....	175
6.1.45	( <i>E</i> )-4,4'-(Diazene-1,2-diyl)bis(3,5-dichlorobenzoic acid) (169) .....	176
6.1.46	( <i>E</i> )-(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))dimethanol (170) 176	
6.1.47	( <i>E</i> )-(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))bis(methylene) bis(2-chloroacetate) (171) .....	177
6.1.48	4-Amino-3,5-dichlorobenzoic acid (168).....	178
6.1.49	Ethyl-(4-amino-3,5-chloro)-benzoate (173) .....	178
6.1.50	Diethyl-4,4'-(diazene-1,2-diyl)-(E)-bis(3,5-dichlorobenzoate) (172) .....	179
6.1.51	(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))dimethanol (170) ....	180
6.1.52	( <i>E</i> )-(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))-bis-(methylene)- bis-(2-chloroacetate) (171) .....	180
6.1.53	4-Amino-3,5-dibromobenzoic acid (175).....	181
6.1.54	Ethyl 4-amino-3,5-dibromobenzoate (176) .....	182
6.1.55	Diethyl 4,4'-(diazene-1,2-diyl)bis(3,5-dibromo)benzoate (177) .....	182
6.1.56	( <i>E</i> )-(Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene))dimethanol (178) 183	
6.1.57	( <i>E</i> )-(Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene))bis(methylene) bis(2-chloroacetate) (179).....	184
6.1.58	4-Bromo-2,6-difluoroaniline (136).....	184

6.1.59	4-Amino-3,5-difluorobenzonitrile (180X) .....	185
6.1.60	4-Amino-3,5-difluorobenzoic acid (181) .....	186
6.1.61	Ethyl(4-amino-3,5-fluoro)benzoate (182).....	186
6.1.62	Diethyl 4,4'-(diazene-1,2-diyl)bis(3,5-difluorobenzoate) (183).....	187
6.1.63	(E)-(Diazene-1,2-diylbis(3,5-difluoro-4,1-phenylene))dimethanol (183) 188	
6.1.64	(E)-(Diazene-1,2-diylbis(3,5-difluoro-4,1-phenylene))bis(methylene)-bis- (2-chloroacetate) (185) .....	188
6.1.65	4-Amino-3-benzoic acid (187) .....	189
6.1.66	Ethyl (4-amino-benzoate) (188) .....	190
6.1.67	Diethyl-4,4'-(diazene-1,2-diylbenzoate) (189).....	190
6.1.68	(E)-(Diazene-1,2-diylbis(4,1-phenylene))dimethanol (190) .....	191
6.1.69	(E)-(Diazene-1,2-diylbis(4,1-phenylene))bis(methylene)-bis(2- chloroacetate) (191) .....	192
6.1.70	2-Bromo-4-aminobenzonitrile (194) .....	192
6.1.71	4-Amino-3-chlorobenzonitrile (193) .....	193
6.2	General method for the synthesis of 4-amino-3-halobenzoic acids.....	194
	4-Amino-3-bromobenzonitrile (197) .....	194
	4-Amino-3-chlorobenzoic acid .....	194
	4-Amino-3-fluorobenzoic acid.....	195
6.3	General method for the synthesis of ethyl(4-amino-3 halobenzoates).....	195
6.3.1	(4-Amino-3-bromo)benzoate (200) .....	195
6.3.2	(4-Amino-3-chloro)benzoate (199) .....	196
6.3.3	Ethyl(4-amino-3,5-fluoro)benzoate (198) .....	196
6.4	General method for the synthesis of haloazobenzene ethyl esters .....	196
6.4.1	Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-bromobenzoate) (203).....	197
6.4.2	Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-chlorobenzoate) (202) .....	197
6.4.3	Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-fluorobenzoate) (201) .....	197
6.5	General method for the reduction of haloazobenzene diethyl esters to diols 198	
6.5.1	(E)-(diazene-1,2-diylbis(3-bromo-4,1-phenylene))dimethanol (206) .....	198
6.5.2	(E)-(diazene-1,2-diylbis(3-chloro-4,1-phenylene))dimethanol (205) ....	198
6.5.3	(E)-(Diazene-1,2-diylbis(3-fluoro-4,1-phenylene))dimethanol (204).....	199
6.6	General method for the chloroacetylation of azobenzene diols.....	199

6.6.1	(E)-(Diazene-1,2-diylbis(3-bromo-4,1-phenylene))bis(methylene)-bis-(2-chloroacetate) (209).....	200
6.6.2	(E)-(Diazene-1,2-diylbis(3-chloro-4,1-phenylene))bis(methylene)-bis-(2-chloroacetate) (208).....	200
6.6.3	(E)-(Diazene-1,2-diylbis(3-fluoro-4,1-phenylene))bis(methylene) bis(2-fluoroacetate) (207).....	200
6.6.4	4-Amino-3-bromo-5-fluorobenzonitrile (216) .....	201
6.6.5	4-Amino-3-bromo-5-fluorobenzoic acid (217) .....	201
6.6.6	Ethyl 4-amino-3-bromo-5-fluorobenzoate (218).....	202
6.6.7	Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-bromo-5-fluorobenzoate) (219) 203	
6.6.8	(E)-(Diazene-1,2-diylbis(3-bromo-5-fluoro-4,1-phenylene))dimethanol (220) 203	
6.6.9	(E)-(Diazene-1,2-diylbis(3-bromo-5-fluoro-4,1-phenylene))bis(methylene)-bis-(2-chloroacetate) (221).....	204
6.6.10	4-Amino-3-chloro-5-fluorobenzonitrile (210).....	205
6.6.11	4-Amino-3-chloro-5-fluorobenzoic acid (211) .....	206
6.6.12	Ethyl 4-amino-3-chloro-5-fluorobenzoate (212) .....	206
6.6.13	Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-chloro-5-fluorobenzoate) (213) 207	
6.6.14	(E)-(Diazene-1,2-diylbis(3-chloro-5-fluoro-4,1-phenylene))dimethanol (214) 208	
6.6.15	(E)-(Diazene-1,2-diylbis(4,1-phenylene))bis(methylene) bis(2-chloroacetate) (215) .....	209
6.6.16	Benzyl bromide (224).....	209
6.6.17	Benzyl azide (225) .....	210
6.6.18	Benzyl azide (225) .....	210
6.6.19	N-Benzyl-1,1,1-triphenyl-15-phosphanimine (227).....	211
6.6.20	N-Benzyl chloroacetamide (228).....	211
6.6.21	(E)-1,2-bis(2,6-Dibromo-4-(bromomethyl)phenyl)diazene (229).....	212
6.6.22	(E)-2,2'-(((Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene)) bis(methylene)) bis(azanediy)) bis(ethan-1-ol) (234) .....	213
6.6.23	(E)-N,N'-((Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene)) bis(methylene)) bis(2-chloro-N-(2-hydroxyethyl)acetamide) (235) .....	214
7	Appendix.....	216
7.1	Slug peptide masses.....	216

---

7.2	Crystal structure data .....	220
7.2.1	( <i>E</i> )-(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))bis(methylene) bis(2-chloroacetate) (171).....	220
7.2.2	( <i>E</i> )-(Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene))bis(methylene) bis(2-chloroacetate) (179).....	225
7.2.3	( <i>E</i> )-(Diazene-1,2-diylbis(3,5-difluoro-4,1-phenylene))bis(methylene)-bis(2-chloroacetate) (185) .....	229
7.2.4	( <i>E</i> )-(Diazene-1,2-diylbis(4,1-phenylene))-bis-(methylene)-bis(2-chloroacetate) (191) .....	233
7.2.5	( <i>E</i> )-(Diazene-1,2-diylbis(3-bromo-4,1-phenylene))-bis-(methylene)-bis(2-chloroacetate) (209).....	240
7.2.6	( <i>E</i> )-(Diazene-1,2-diylbis(3-bromo-5-fluoro-4,1-phenylene))-bis-(methylene)-bis(2-chloroacetate) (221).....	244
8	References.....	249
9	List of Figures .....	255
10	List of Tables.....	262

# Introduction

## 1 Introduction

### 1.1 Targeting protein-protein interactions

Protein-protein interactions are involved in nearly all cellular processes and understanding how these protein-protein interactions influence cellular behaviour is crucial for understanding the molecular basis of life. Mapping the networks of interactions will allow the identification of the ideal points to specifically intervene in disease states. These points could then be targeted with therapeutic agents to block specific interactions or mimic others to correct disease states or cause affected cells to undergo programmed cell death. Protein-protein interactions can be difficult to mimic because they often occur over large areas of their surfaces; typically 750 to 1500 Å<sup>2</sup> but up to 4600 Å<sup>2</sup> in some cases.<sup>1</sup> Despite the large contact areas much of the binding energy can result from a relatively small area within this larger region. For example, in a complex between human growth hormone (HGH) and its first bound receptor the side chains of thirty amino acids make contact with the receptor, but a pair of tryptophan residues in a central hydrophobic region are responsible for three quarters of the binding energy.<sup>2</sup> Such sites are referred to as hot spots and are possible targets for small molecule inhibitors, which can be used to sterically and electrostatically mimic the amino acid sidechains at these positions.<sup>3,4</sup> However, if no hot spots are present it can be difficult to engineer small molecules to mimic multiple lower-affinity interactions as the binding energy is more evenly distributed over the interacting surface. The protein surface might also be dynamic or flexible and be capable of multiple interactions. This can be difficult to detect by examining protein crystal structures as the average solution conformation may not resemble the conformation found in crystal structures.<sup>5</sup>

Another way to influence protein-protein interactions is to alter the concentration of one of the proteins or increase the concentration of another protein that competes for binding sites or blocks the protein-protein interaction partners. A number of methods have been described that can be used to alter the frequency of gene transcription and hence translation and the concentration of a particular protein in a cell. If the sequence of the target gene is known, the gene may be deleted preventing it from ever being transcribed. This can be achieved by removal of its start codon or by replacing the DNA encoding the protein of interest with a drug resistance gene as a selectable marker for

generating knock outs.<sup>6</sup> This technique is particularly effective when used to target regulatory proteins that affect the concentration of multiple downstream proteins.

Gene editing can be used more subtly to introduce mutations into the genome of an organism. Contemporary techniques allow a single site on the genome to be targeted to allow the effect of a single nucleotide polymorphisms like that responsible for sickle cell anemia to be investigated.<sup>6</sup> These small changes can have dramatic effects; sickle cell anaemia results from a single point mutation of an adenosine to a thymine in the *HBA1* gene which changes a single amino acid of  $\alpha$ -globin protein from a glutamic acid to valine. This does not effect the fold of the globin, but alters its propensity to oligomerise in red blood cells and can change the shape of the cell from a round disc to a sickle shape. This reduces the efficiency of oxygen transportation throughout the body.

Gene editing is often achieved using homologous recombination. A vector is constructed that contains a gene encoding a protein that confers resistance to a cytotoxic agent. This allows cells that do not take up the vector to be eliminated as only the cells that contain the new gene survive subsequent treatment with the cytotoxic agent. This gene also contains a negative selection marker and a gene encoding the modified protein to be inserted to replace an existing problematic gene or to introduce a certain characteristic into the organism. These genes are placed between two regions homologous to the position in the genome of interest. Cells are treated with this plasmid and the cytotoxic agent is applied to select cells containing the plasmid, killing cells without this newly inserted resistance. In a fraction of these cells, the complementary flanking sequences will align with chromosomal DNA so these will overlap with the existing DNA sequence and hybridise resulting in incorporation of the modified gene at the specified site on the genome. The negative selection marker is then used to screen out non-homologous insertions of the plasmid DNA into other loci.<sup>7</sup> Gene knockouts can also be created by gene editing using designer or custom made endonucleases that create double-stranded breaks in DNA at precisely defined locations in the genome. Specificity is achieved by using a combination of modular DNA recognition proteins (zinc fingers or transcription activation like domains) that bind to specific sequences on the genome to specify the target sequence and a non-specific catalytic domain which is only active as a dimer so two different recognition elements can combine to specify a non-palindromic site.<sup>8</sup> Theses nucleases can cut the



DNA causing double strand breaks. Machinery in the cell then repairs these breaks either by homology directed repair, where a template guides the repair towards the correct sequence the repair is not error prone as the repair follows the sequence of the complimentary DNA strand or by non-homologous end joining where there is no strand to follow resulting in a repair which is error prone.<sup>8,9</sup> Recently, gene editing systems based on proteins that interact with clustered regularly interspaced short palindromic repeats in the bacterial genome to form an adaptive defence against foreign DNA (CRISPR/CAS9) have become very widely used. These make use of CAS9 nuclease and short guide RNA to direct the nuclease to cut at a desired position. The guide RNA binds to CAS9 and to the DNA at the site of interest and guides the CAS9 locus. Addition of a DNA template that contains a sequence complimentary to the short guide RNA during repair allows genes to be removed and exchanged for the desired sequence.

These approaches are only feasible if the target organism can survive with the gene modified or turned off , as genes cannot easily be subsequently reactivated. This method is therefore unsuitable for the study of cell development or differentiation. Temporary control of the expression of some genes can be achieved by dosing cells with inhibitors that interfere with the action of an appropriate transcription factor by interrupting either its binding to DNA, or its interaction with the RNA polymerase complex.<sup>10</sup> This approach can provide varying degrees of spatial and temporal control, but relies on cells recovering from invasive direct injection procedures or rapid diffusion/uptake and activity kinetics of the inhibitor.

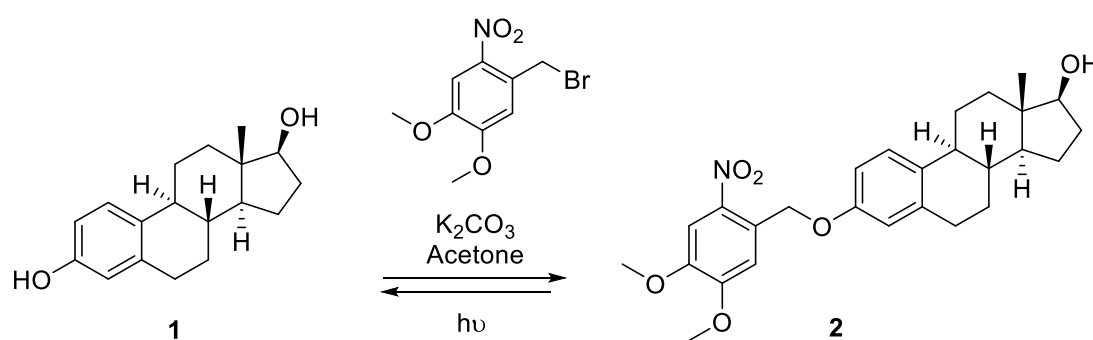
Protein-protein interactions can also be perturbed using peptides derived from the interacting portions of proteins. Although short peptides rarely adopt the same conformation as they do as part of a larger protein, their conformations can be fixed by chemical modification. For example, a peptide derived from the  $\alpha$ -helical Bcl-2 homology region of the proapoptotic protein Bid was constrained into a more helical conformation by a hydrocarbon staple.<sup>11</sup> This peptide was shown to bind to an interaction partner of native Bid, Bcl-x<sub>L</sub>, with an affinity greater than the non-crosslinked peptide. Furthermore, a related Bak peptide was shown to bind to the same pocket at that of the wild type Bak and Bid BH<sub>3</sub> regions when constrained by a photoresponsive staple,<sup>12</sup> which can moderate the conformation and therefore binding affinity of the peptide in response to light. The development of a photocontrolled

peptide for controlling gene expression in a cell would create a powerful tool for investigating protein-protein interactions. Molecules introduced into cells in an inert form could be activated with precisely directed light after the cells have had chance to recover from any perturbation caused its introduction. This approach requires that the activity of the active and inactive forms are sufficiently different to the effects of the inactive species present before and after irradiation, and that the photoswitch is stable and functional under physiological conditions.<sup>13</sup>

## 1.2 Methods for photo control of peptides

### 1.2.1 Photocages

Photocontrol can be exerted in numerous ways, whilst most reported methods have been designed for very specific purposes, some have more general applicability. Research into drugs whose activity can be controlled with light (termed photopharmacology<sup>13</sup>) is particularly promising because one of the major problems in the pharmaceutical industry is achieving sufficient specificity. Up to 87% of small molecule drug candidates are discarded, many due to low selectivity.<sup>14</sup> As photoactivation can potentially be used to select single cells, the concentration of an active species can be very precisely controlled making otherwise unselective drugs useful. Photocaging is ideal for this application as reduced activity of a photocaged prodrug allows higher concentrations of the prodrug to be used, effectively widening the therapeutic window.<sup>15</sup>

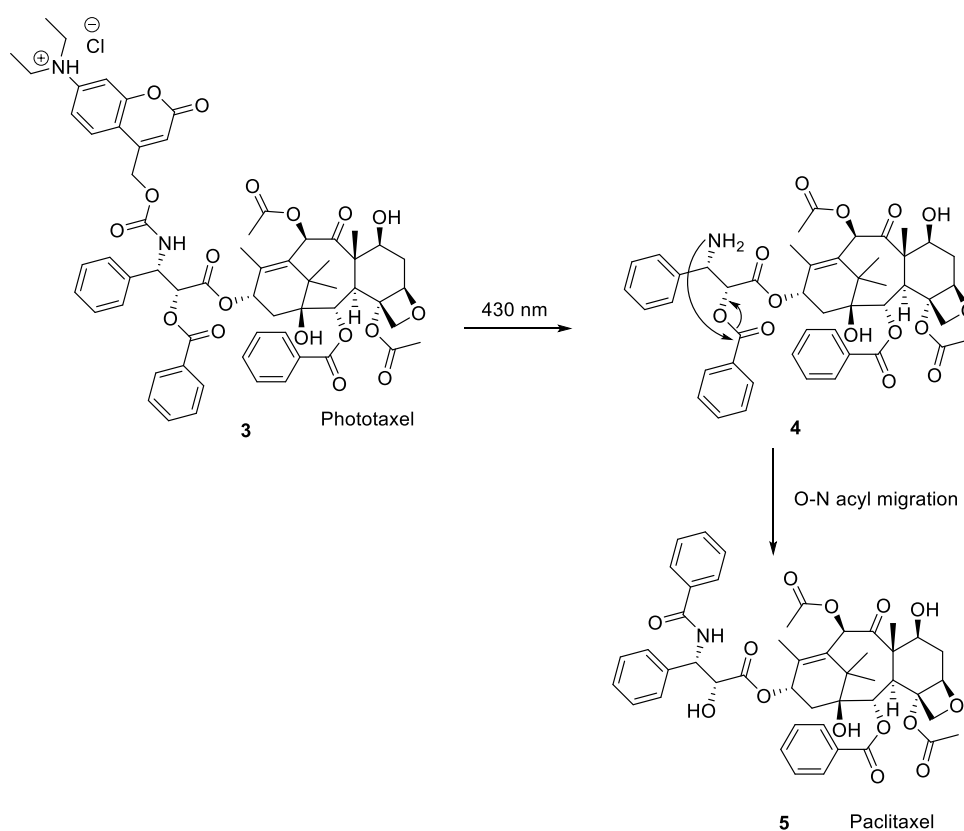


**Figure 1.** Estradiol protection and photoactivation.

Photocaging can also be used to activate gene transcription and therefore protein production by caging hormones such as estradiol (Figure 1).<sup>16</sup> When estradiol binds to the estrogen receptor it causes a conformational change, which then leads to gene

expression.<sup>17</sup> UV light degraded 86% of the prodrug **2** in cells, resulting in increased gene transcription.<sup>16</sup>

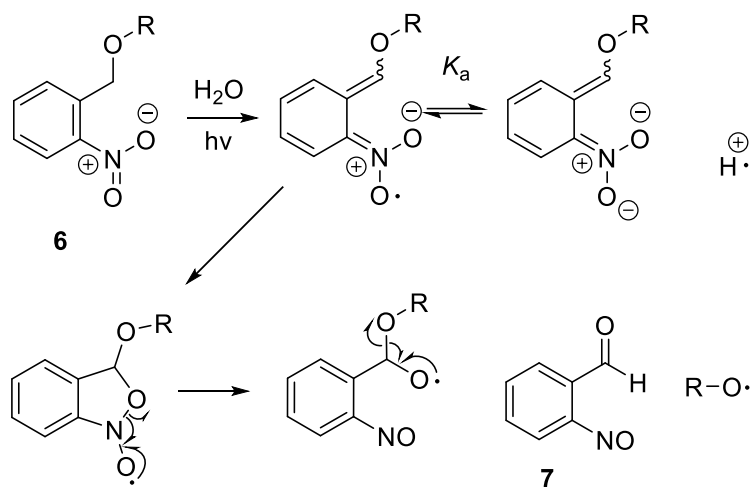
Most anti-cancer drugs act by preventing cell division, but this can dramatically affect naturally rapidly replicating cells in the body such as those of the hair and stomach lining as well as those forming tumours. Activating prodrugs only at tumour sites avoids these issues.<sup>15</sup> This approach is illustrated by phototaxel (**3**) which is activated by irradiation with 430 nm light, with 69% paclitaxel (**5**) formed after 30 minutes via an acyl shift through intermediate **4** (Figure 2).<sup>15</sup>



**Figure 2.** Phototaxel and photo deprotection to paclitaxel.

Photocages can also be attached to proteins, peptides or other bio-active molecules and then removed by irradiation with light. Decaging typically proceeds by a radical mechanism and reaction times from the initial pulse of light to the release of the photocage can be quite lengthy, with some taking up to 24 hours to reach completion.<sup>16</sup> The rate of release of the active drug can be further hampered by decaging mechanisms that require proton transfers making reaction rates dependant on the surrounding

environment. Additionally, the longer the half-life of the reactive intermediate the greater the likelihood of unwanted side reactions.<sup>18</sup>



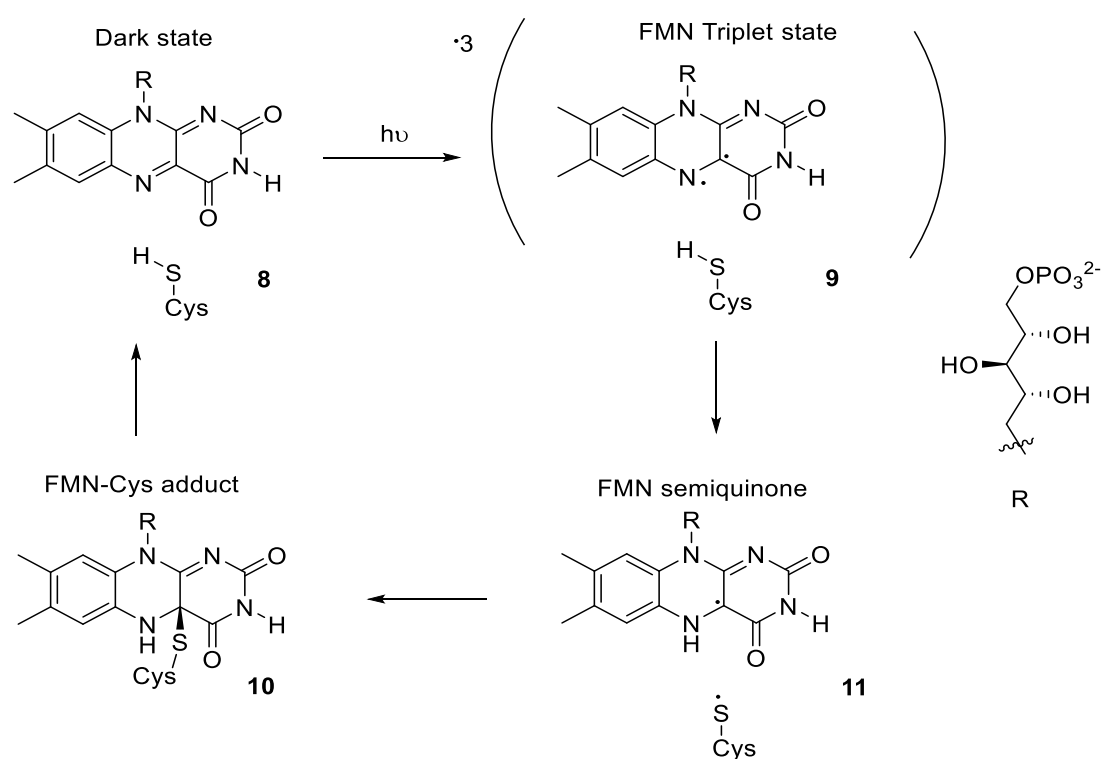
**Figure 3.** Radical photo cleavage of *ortho*-nitrobenzyl ether protecting groups.

One commonly used class of photocages are based on 2-nitrobenzyl ethers (Figure 3). This group has been known since the 1970s and is used as a photo-labile protecting group in organic chemistry. Adoption has been limited by the relative expense of its precursors and the radical-based deprotection reaction, as the presence of any radical acceptors in the molecule can lead to undesirable side reactions.<sup>18,19</sup> This approach was transferred to biology by Kaplan *et. al.*,<sup>20</sup> using a nitrobenzyl group to photocage, then release ATP upon irradiation. Photocaged ATP was introduced into ghost cells, eosinophilic epithelial cells without a nucleolus, and was shown not to be an inhibitor or substrate for Na, K - ATPase. Photodecaging allowed the concentration of ATP to be varied by varying the intensity of the incident light. The same nitrobenzyl group was used to control the binding of zinc finger transcription factors by Chou *et. al.*,<sup>21</sup> a combination that has been commercialised and was available in over 40 different derivatives by 2002. However, the wide availability has shown that decaging these molecules can release toxic by-products and that these by-products are strongly UV absorbing, reducing release of the desired active agent. The final drawback of photocaged molecules is that they are single use, once activated they remain active and the concentration is then dependent on the biological half-life of the active agent.

### 1.3 Reversible photocontrol

#### 1.3.1 Photoresponsive proteins

Photoresponsiveness is vital to plants, controlling a variety of processes from seed germination to the direction of growth and flowering. There are four major classes of photoreceptors in plants, covering different wavelengths of light.<sup>22</sup> Photochromes sense red and far red light (600-750 nm) and several classes of proteins respond to blue light (320-500 nm) including cytochromes, and LOV (light, oxygen voltage) proteins.<sup>22,23</sup> These proteins contain cofactors that can absorb longer wavelengths of light than the protein itself; for example LOV domains contain a flavin chromophore.<sup>24</sup>

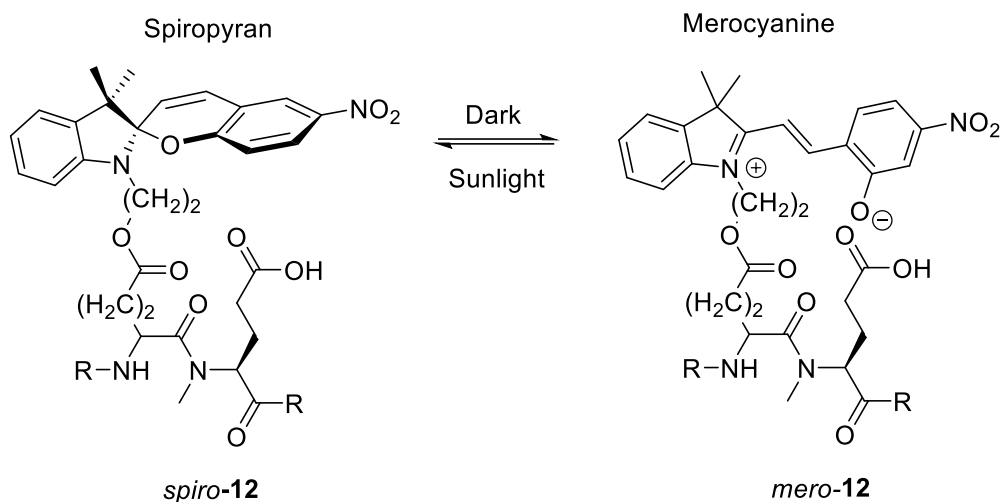


**Figure 4.** Photocycle of flavin mononucleotide (FMN) in LOV domains.

Blue light causes a reaction with the sidechain of a cysteine residue, leading to conformational changes. In the case of *Avena sativa* LOV<sub>2</sub> domain (AsLOV<sub>2</sub>) these changes cause a C-terminal  $\alpha$ -helix that is tightly bound to the  $\beta$  sheet of the LOV domain in the dark state to be released.<sup>25</sup> Release of this  $J_\alpha$  helix activates the adjacent downstream phototropin kinase domain.<sup>26</sup>

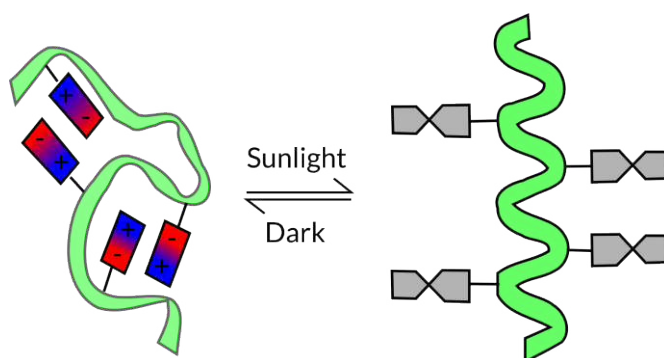
### 1.3.2 Spiropyrans

Spiropyrans are an example of reversible photo switches that can be used to regulate the conformation of peptides.



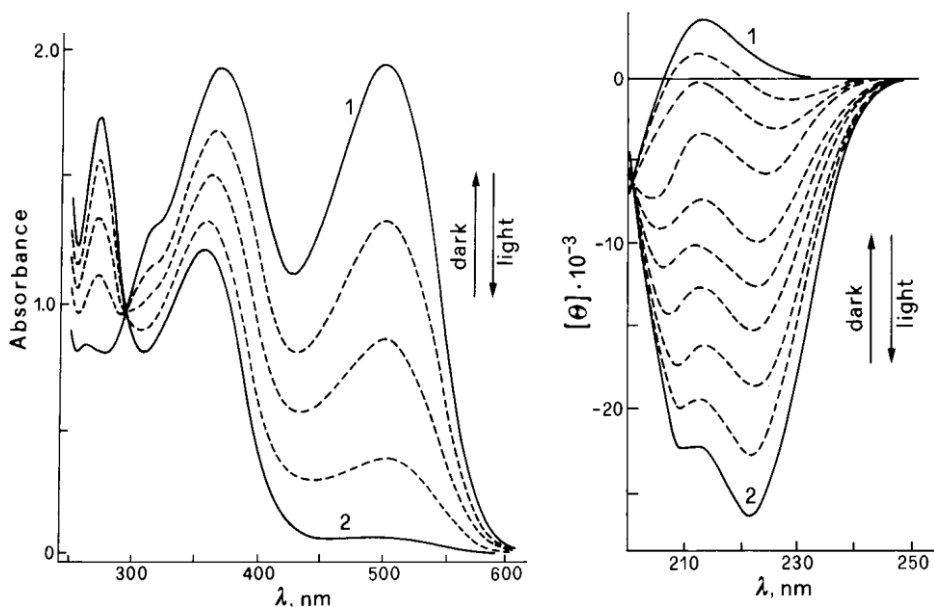
**Figure 5.** Spiropyrans are photochromic molecules that change structure when exposed to different wavelengths of light. (R is poly-L-glutamic acid).<sup>27</sup>

The merocyanine form is more thermodynamically stable than the spiropyran form, but the energy barrier between the two states is very low. On exposure to daylight the zwitterionic merocyanine will form the spiropyran,<sup>27,28</sup> and when conjugated to a peptide by, eg. pair of amides, the change can influence the conformation of the polyglutamate in **12** or can be extended to control the protein it is attached to (Figure 6).



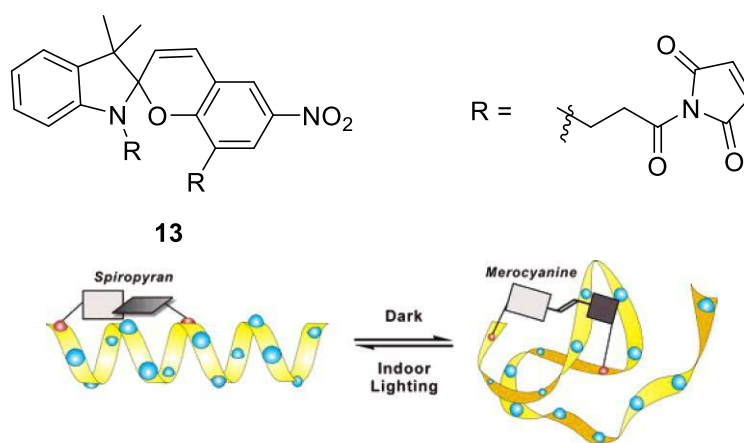
**Figure 6.** Conformational change brought about in poly-L-glutamic acid by the different states of spiropyran (Adapted from reference 27).

This conformational change in **12** is clearly visible in UV spectra, showing the switch has changed states, and circular dichroism (CD) spectra which show a significant increase in  $\alpha$ -helical contribution to its structure.

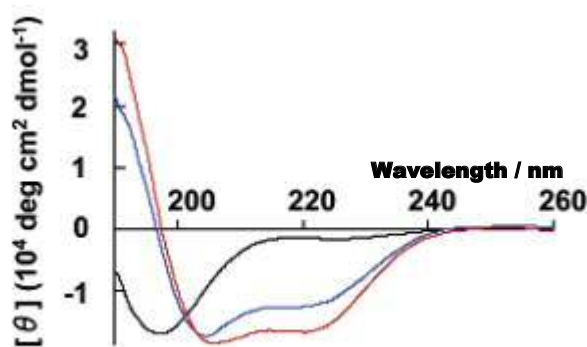


**Figure 7.** UV (left) and CD (right) absorbance spectra for poly-L-glutamic acid with spiropyran attached showing the light (1) and dark states (2) with dashed lines showing intermediate relaxation spectra.<sup>27</sup>

Control over secondary structure has also been achieved by attaching the spiropyran to two specific amino acid sidechains with an  $i, i+7$  spacing.<sup>29</sup>

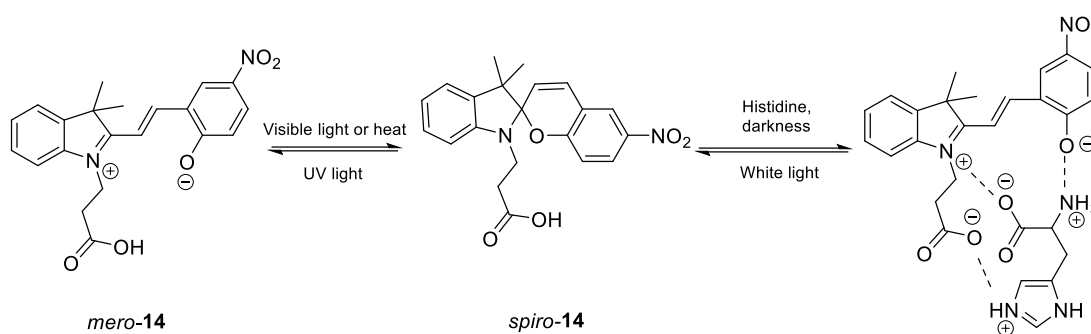


**Figure 8.** Spiropyran attached a peptide at two points to control peptide  $\alpha$ -helical conformation (adapted from reference 29).



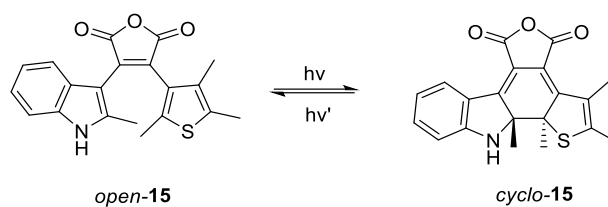
**Figure 9.** CD spectra of native peptide (black) and crosslinked peptide in light (red) and dark (blue) states in phosphate buffer (100 mM, pH 6.6) at 25 °C.<sup>29</sup>

The change in the CD spectrum brought about by switching of the peptide is small, but there is some control of the conformation of the peptide. One of the major problems with this photochrome arises from the zwitterionic nature of its dark state, which means it can interact with other ions in the solution such as histidine (Figure 10).<sup>28</sup>



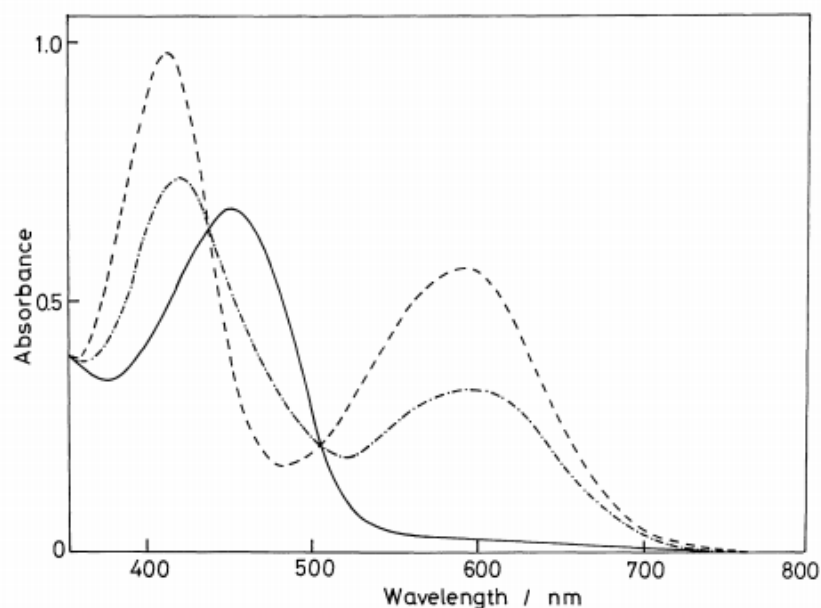
**Figure 10.** Interaction of spiropyran **14** and histidine in solution.<sup>28</sup>

### 1.3.3 Diarylethenes



**Figure 11.** Diarylethenes can cyclise to hexadienes.<sup>30</sup>

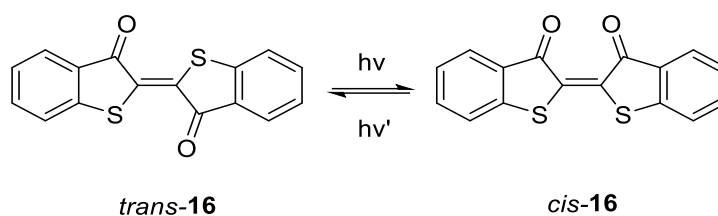




**Figure 12.** Absorbance spectra of **15** in diarylethene (open, —) and hexatriene (closed, -----) states and the photo stationary state after irradiation at 491 nm (-·-·-) in benzene.<sup>30</sup>

Diarylethenes contain a hexatriene motif; three consecutive alkenes. This motif can undergo a photo-induced cyclisation that drastically changes the shape of the molecule and its rigidity.<sup>31</sup> The open and closed isomers are both thermodynamically stable and are readily interconverted using different wavelengths of light.<sup>30</sup> The lifetimes and extents of conversion of these states are dependent upon the substitution of the diarylethenes, with reported half-lives ranging from minutes to molecules that are thermodynamically stable in the closed form.<sup>31</sup> Diarylethenes can be attached to DNA as a structural constraint and have been shown to be able to control transcription *in vitro*.<sup>32</sup>

#### 1.3.4 Thioindigo

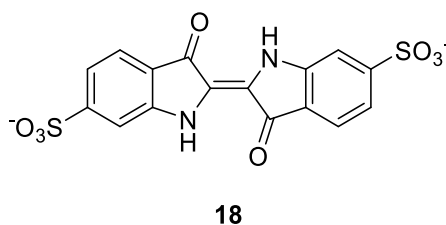


**Figure 13.** Thioindigo undergoes photoinduced *cis/trans* isomerisation.

Thioindigo is a modified version of indigo, a commonly used dye in the clothing industry. The reduced *leuco* "white indigo" form of the commercial dye is very soluble

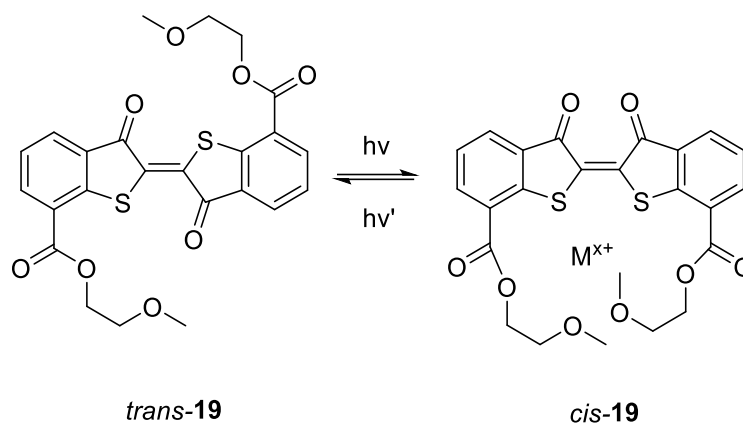
allowing it to be easily administered to fabrics, but as it dries it oxidises back to the highly insoluble parent form. Unlike, indigo whose *Z*-isomer cannot be detected due to its short lifetime caused by additional hydrogen bonding capabilities, the isomers of thioindigo have more evenly matched energy levels and are readily separable.<sup>33</sup> Thioindigo is an excellent example of a reversible photo switch with many favourable properties; the barrier to rotation is so great that it shows almost no reversion of the light state (*cis* excited state) to the dark state (the *trans* ground state) and excitation occurs at visible wavelengths ( $> 450$  nm).<sup>34</sup>

*Trans* to *cis* isomerisation is promoted by irradiation at 450 nm light, yielding a 70% *cis* photo stationary state. The reverse *cis* to *trans* isomerisation is promoted by irradiation with 510 nm light.<sup>35</sup> The major disadvantages of these molecules are their typically very poor solubility in aqueous solution, once attached to peptides they can result in an insoluble conjugate.<sup>36</sup> Several attempts have been made to increase thioindigo solubility, as the result would be extremely useful for the dyestuff industry. Many early modifications aimed at increasing the polarity of thioindigo,<sup>37</sup> following the example of the highly soluble indigo carmine (**18**) shown (Figure 14).



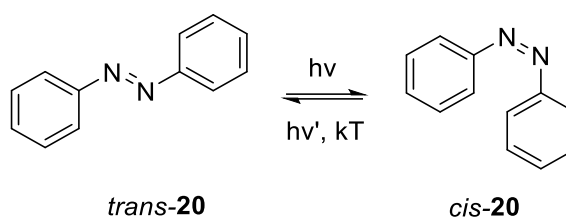
**Figure 14.** Indigo carmine, a water soluble indigo.

Another approach involved the addition of bulky substituents to the ring system to reduce the  $\pi$ -stacking that favours self-association over dissolution. This was achieved by the addition of ethylenedioxy side groups at the 7 and 7' positions (Figure 15). The resulting thioindigo was used as a photoreponsive chelating agent where the *cis* form can chelate metal ions and reversion to the *trans* form releases the trapped ions.<sup>38</sup>



**Figure 15.** Thioindigo molecular tweezers.

#### 1.4 Azobenzene

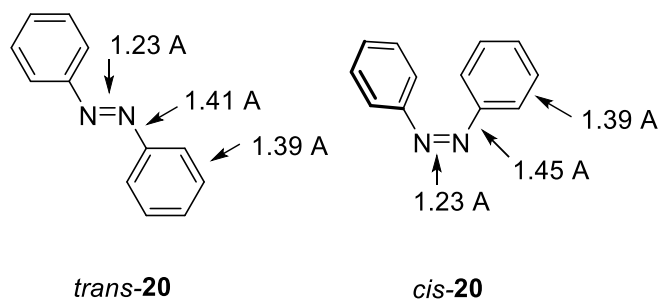


**Figure 16.** Azobenzene in the *trans* conformation (left) and the *cis* conformation (right).

Azobenzene is one of the most versatile and widely used photoswitches with a wide range of applications reported. Azobenzene is a photochromic molecule (Figure 16) consisting of two phenyl rings that are connected by an azo group and is part of a large class of azo-derived organic dyes. The *trans* form of azobenzene was discovered in 1843 by Eilhard Mitscherlich<sup>39</sup> who described very little except that its properties deviated from stilbene. The crystal structure of *trans* azobenzene was solved by Robertson in 1935.<sup>40</sup> The *cis* form of azobenzene was first documented by Hartley in 1937,<sup>41</sup> who discovered that the absorbance of a solution of azobenzene was not reproducible. This effect persisted even after recrystallization and eventually the two isomers were separated by their slightly different solubility. This is probably due to a combination of the increase in the dipole moment due to the nitrogen lone pairs being more exposed in the *cis* state and the reduced planarity of the *cis* state improving solvation and discouraging  $\pi$ -stacking between azobenzene molecules.

The newly discovered *cis* state could be converted back to the *trans* state by heating the azobenzene in solution. These experiments were complicated by the need to exclude light as it was observed that exposure to different types of light caused changes in the spectra brought about by the change in percentage of the *cis* to the *trans* states.

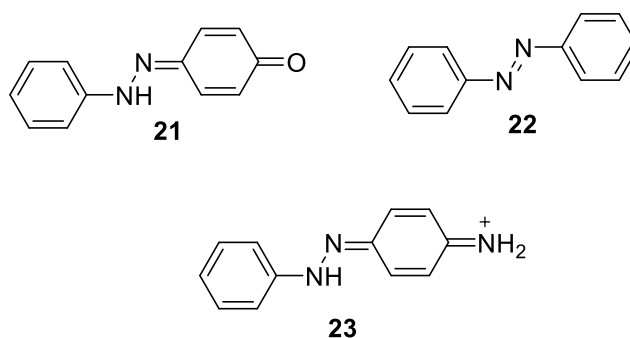
Spontaneous reversion from *cis* to *trans* is driven by the thermodynamic stability of the *trans* isomer. Conversion from the *cis* to the *trans* state is therefore exothermic and from this heat the thermodynamic energy difference for the isomers was calculated to be 12 kcal/mol. This energy difference is due to unfavourable steric interactions in the *cis* state and the planar *trans* state allowing favourable orbital overlap.



**Figure 17.** Bond lengths of azobenzene from crystal structures.

Not only are there steric and electronic differences between isomers, but bond lengths also change. Most notable is the change in the N-C bond length from 1.45 Å in the *cis* state to 1.41 Å in the *trans*. The average length of a C-N bond is 1.47 Å, shorter than usual bond in the *trans* isomer is due to increased orbital overlap.<sup>42</sup> The C-N-N bond angle of the *trans* isomer is 121°, very close to that of a trigonal planar molecule with the lone pairs in plane of the bond. The *cis* state cannot be planar as implied above (Figure 17), due to steric constraints one of the aromatic rings is rotated by 40° with respect to each other.<sup>41</sup>

Heartly *et al.* later discovered that substitution of the phenyl rings has a pronounced effect on isomerisation and could increase the speed of the thermal reversion from the *cis* to the *trans* isomer.<sup>43</sup> This effect was explained by proposing the formation of certain intermediates in the reversion pathway which favour additional single bond character between the nitrogen atoms, lowering the energy barrier to rotation (Figure 18).



**Figure 18.** Substitution of the *para* position effects the thermal relaxation of azobenzene.

*trans*-Azobenzene has an extinction coefficient of 22,000 M<sup>-1</sup>cm<sup>-1</sup> for the  $\pi$ - $\pi^*$  transition and of 440 M<sup>-1</sup>cm<sup>-1</sup> for the n- $\pi^*$  transition.<sup>44</sup> The absorbance maxima and extinction coefficients of these bands vary widely according to the substitution patterns of the azobenzene.

#### 1.4.1 Mechanism of *trans* to *cis* conversion

In order for the change in conformation of azobenzene to occur, the azo bond is required to behave more like an N-N single bond than an N=N double bond. One of the ways this can be achieved is by exciting an electron in the N=N double bond from the  $\pi$  bonding orbital to  $\pi^*$  antibonding orbital, breaking the N=N double bond and allowing free rotation or the direct inversion around the remaining N-N single bond. The same effect can also be achieved by excitation of an electron in the non-bonding orbital to the  $\pi^*$  orbital. The wavelength of light used to bring about these changes is dependent on the energy gaps between these two orbitals. This excited state can either then relax back into the *trans* state or rotate and relax into the *cis* state with an efficiency dependant on the lifetime of the excited state and the rate of rotation. Relaxation from the excited state *cis* to the ground state *trans* isomer is temperature dependent and requires a collision with a solvent molecule which receives the excess energy. The relaxation of azobenzene has been extensively modelled<sup>45,47,48,49</sup> and experimentally observed to be dependent both on the solvent and substitution of the azobenzene with the choice of solvent strongly influencing the mechanism of conversion.<sup>46</sup>

A range of factors affect the rate and extent of photo-conversion of azobenzene from the *trans* to *cis* excited state. It is not typically possible to achieve complete conversion to the *cis* isomer due to competing photo-conversion from *cis* to *trans* state. The

absorbance band of the *cis* state is usually not significantly altered by the change in conformation so that the  $\pi$ - $\pi^*$  or the  $n$ - $\pi^*$  band of the *cis* isomer can still be excited by the light used to form it from the *trans* isomer, causing the back conversion and a photostationary state with an equilibrium mixture of both isomers. Complete conversion is possible where absorption bands do not overlap; azobenzene can be forced to relax from the *cis* to *trans* isomer by irradiation of 450 nm light, which causes  $n$ - $\pi^*$  excitation of an electron from the *cis* N=N bond allowing free rotation around the N-N bond. It is also possible to obtain a fully *trans* state of azobenzene by heating an azobenzene in solution in the dark until it has fully relaxed.

One of the major problems with the use of photoswitches in general is susceptibility to photo bleaching, where chromophores degrade after a number of conversion cycles. Azobenzenes show little photo bleaching under typical conversion fluxes and can be cycled multiple times without significant degradation of the switch. This could be due to the fast (picosecond) isomerisation, as the excited electronic state intermediate is extremely short-lived, leaving little time for side reactions to occur.

#### 1.4.2 Solvent effects on azobenzene spectra

Ever since the discovery of azobenzene it has been known that solvents can affect the UV absorbance spectra of azobenzenes by changing the peak shape as well as their relative absorbance maxima.<sup>50</sup> Solvents can also alter the rate of *trans* to *cis* isomerisation, *cis* to *trans* reversion and isomer ratio of the photo-stationary state. For example, the rate of relaxation of *cis*-4-diethylamino-4'-nitroazobenzene is 10,000 times faster in dimethylsulfoxide than in cyclohexane.<sup>51</sup> The magnitude of solvent effects are dependent on the substitution of the azobenzene and the interaction of the substituents with the solvent. Solvents can also interact differently with *cis* and *trans* isomers due to the change in dipole moment brought about by the isomerisation.

#### 1.4.3 Tuning the absorbance spectrum of azobenzene

Many properties of azobenzenes can be tuned for particular applications; for example, an application may require particularly efficient conversion from the *trans* to *cis* state or *vice versa* or to avoid the use of high energy UV light sources that are damaging to cells. A pair of azobenzene photoswitches may also be modified to have sufficiently different absorbance maxima to allow their simultaneous use as a pair of different switches. Azobenzenes may also be tuned to avoid deleterious solvent effects or to

enhance environmental effects such as that of pH on their absorbance spectra. The most commonly used azobenzene photoswitches require irradiation by UV light to cause the conversion from *trans* to *cis* state. This can cause a number of issues when used for biological applications. Exposure of cell to UVA light (315–400 nm) can lead to damage from the formation of reactive oxygen and nitrogen species such that high exposure can be cytotoxic.<sup>52,53</sup> Although UVA is not directly absorbed by the DNA, it is absorbed by riboflavins, porphyrins and haem-containing proteins that can form singlet oxygen and superoxides which may damage cell membranes, other proteins or DNA.<sup>52</sup> However, UVA is not as damaging as the other wavelengths of UV light as UVC (220–280 nm) and UVB (280–315 nm) which can directly damage DNA by forming mono-addition and di-addition products of pyrimidine bases. Not only are these wavelengths of light harmful to cells and organisms, organisms have evolved mechanisms to protect themselves from harmful UV irradiation including production of melanin and other conjugated molecules that absorb UV light. Such absorbance of this range of light would lead to difficulty in using azobenzene *in vivo*.

## 1.5 Effects of substitution on azobenzene

### 1.5.1 Substitution of azobenzenes with heteroatoms at the *ortho* position

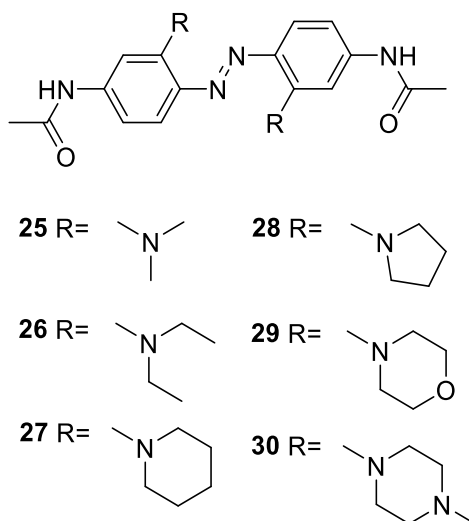


Figure 19. *Ortho* amino substitution.

Examples of substitution at the *ortho* position with respect to the nitrogen of the azobenzene illustrate the magnitude of effects on the half-life and *cis* to *trans* relaxation rates.

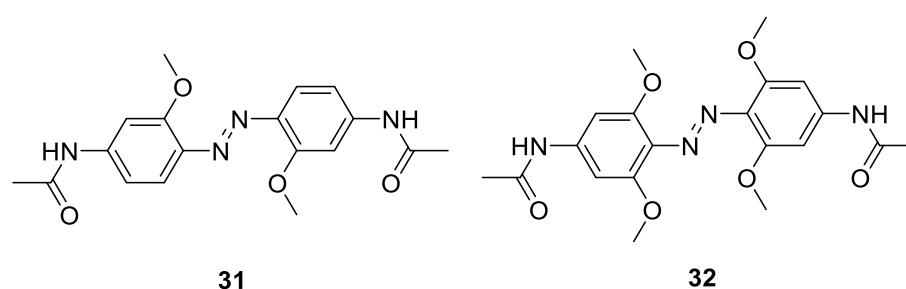
**Table 1.** Table of the effects of *ortho* amino substitution on azobenzene.<sup>54</sup>

Photoswitch	$\lambda_{\max}^{\pi-\pi^*}$ (nm)	$\epsilon^{[a]}$ ( $M^{-1} \text{ cm}^{-1}$ )	$t_{1/2}^{[a]}$ (s)	$t_{1/2}^{[b]}$ (s)
<b>25</b>	470	13,900	$3.3 \pm 0.3$	
<b>26</b>	488	13,000	$0.8 \pm 0.1$	
<b>27</b>	445	10,180	$6.0 \pm 0.2$	
<b>28</b>	513/537		$0.7 \pm 0.1$	
<b>29</b>	435	9,610	$302 \pm 4$	$8.1 \pm 0.2$
<b>30</b>	437	10,440	$210 \pm 10$	$27 \pm 1$

<sup>[a]</sup> in 70% acetonitrile/water, sodium phosphate buffer (1 mM, pH 7.0) 20 °C.

<sup>[b]</sup> in sodium phosphate buffer (10 mM, pH 7.0) 20 °C. <sup>54</sup>

Various *ortho*-amino substituents (Table 1) substantially red-shift the  $\pi-\pi^*$  absorbance band into visible wavelengths and can vary half-life of the *cis* azobenzene by a factor of  $\sim 100$ . However, these relaxation times remain relatively fast in comparison to cellular processes so these switches are poorly suited for *in vivo* control over longer lifetime cell functions. These fast rates are primarily caused by the presence of a nitrogen atom in the *para* position as explained in depth later (section 1.5.3). Faster switching azobenzenes that contain a *para* nitrogen atom have a variety of uses such as controlling nerve cell messages<sup>55</sup> and the control of potassium ion channels.<sup>56</sup>

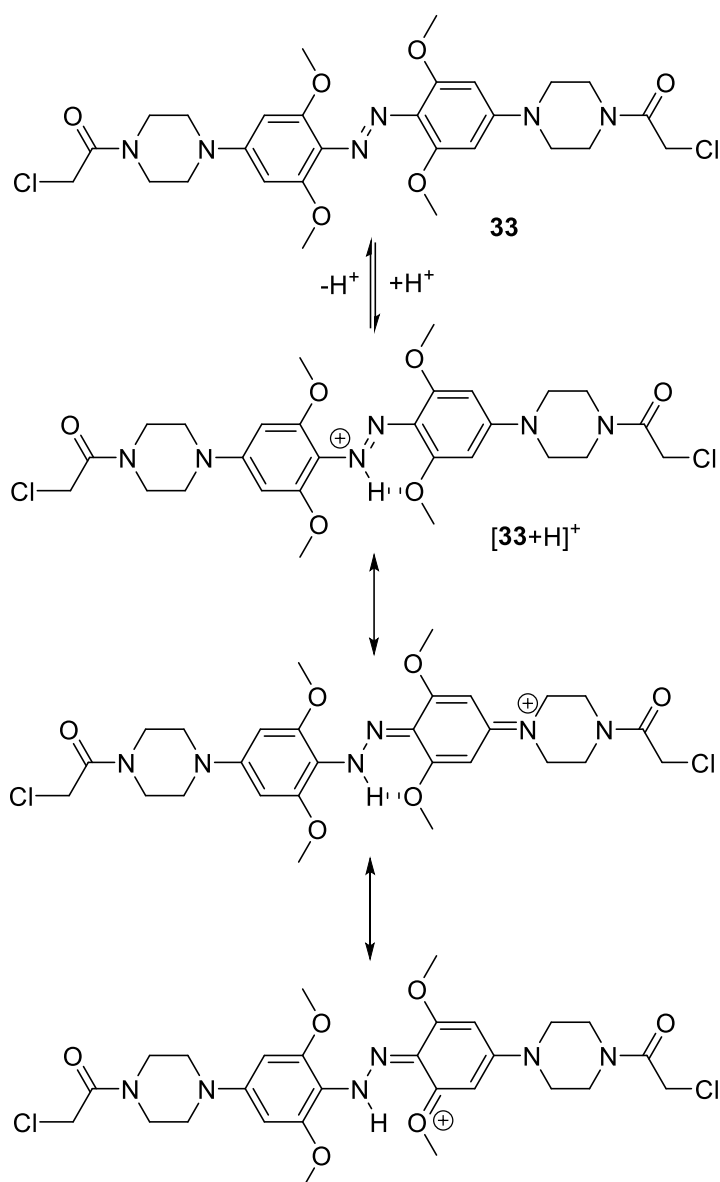


**Figure 20.** Compound **31** *N,N'*-(diazene-1,2-diylbis(3-methoxy-4,1-phenylene))diacetamide and compound **32** *N,N'*-(diazene-1,2-diylbis(3,5-dimethoxy-4,1-phenylene))diacetamide.<sup>57</sup>

In marked contrast to *ortho* amino substituents, the addition of *ortho* methoxy substituents (Figure 20.) resulted in **31**, whose *cis* isomer has a half-life of 2.4 days at 25 °C in sodium phosphate buffer (25 mM, pH 7.0).<sup>57</sup> The switches also show a shift in the

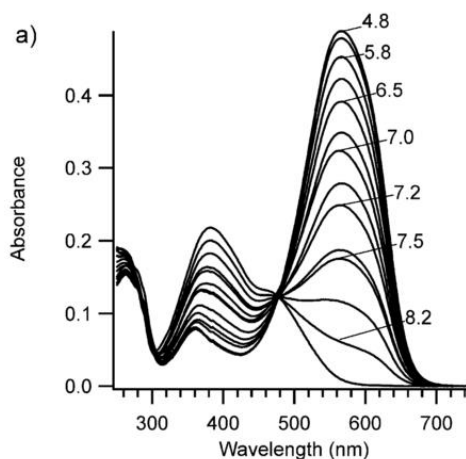


$n-\pi^*$  absorbance band shift in the *cis* isomer relative to the *trans* isomer with a separation between the  $n-\pi^*$  maxima of 35 nm for tetra-*ortho*-methoxyazobenzene **32**. The  $n-\pi^*$  band of *trans*-**32** is temperature sensitive, but at 25 °C the separation between the  $n-\pi^*$  bands of the isomers allows conversion of 100% *trans* to 70% *cis* by irradiation by 530-560 nm light. *cis*-**32** can then be switched back to the 80% *trans* by irradiating with UV light. Lowering the polarity of the solvent greatly increases the half-life of *cis*-**32**, probably due to it destabilising the azonium ion formed by protonation of the azo bond (Figure 21).



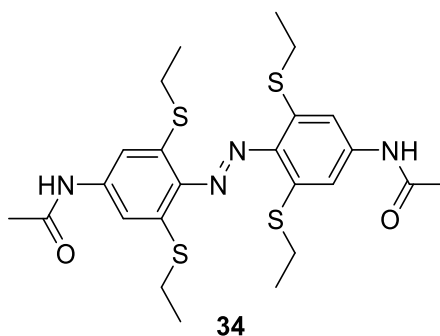
**Figure 21.** 1,1'-((Diazene-1,2-diylbis(3,5-dimethoxy-4,1-phenylene))bis(piperazine-4,1-diyl))bis(2-chloroethan-1-one) Azonium ion stabilization.

Protonation causes a large change in the UV/visible absorbance spectrum of **33**,<sup>58</sup> including a 5-fold change in extinction coefficient between pH 4.8 and 8.2 (Figure 22). Tetra-*ortho*-methoxyazobenzenes species are highly susceptible to glutathione reduction, an issue discussed in more depth later (Section 1.5.2).



**Figure 22.** pH dependent UV/visible absorbance spectra of compound **33**.<sup>58</sup>

To overcome this *pH* sensitivity and the vulnerability to glutathione reduction, the methoxy groups were replaced with thioethers (Figure 23).<sup>59</sup> The sulfur atom less readily forms hydrogen bonds to support the formation of azonium ions, and the resulting azobenzene **34** is not reduced by glutathione (10 mM, 37 °C, 12 hours) and, unlike **33**, **34** shows no sign of photo bleaching.<sup>59</sup>



**Figure 23.** *N,N'*-(diazene-1,2-diylbis(3,5-dithioether-4,1-phenylene))diacetamide.<sup>59</sup>

### 1.5.2 Tetra-*ortho*-haloazobenzenes

The observation that ethers and thioethers could produce such long-lived *cis* isomers and blue shifted absorbance maxima led to the substitution at the *ortho* position with halides to determine their effect on  $n-\pi^*$  band separation. Substitution with fluorine in all four *ortho* positions not only separated the  $n-\pi^*$  absorption bands of the *cis* and *trans* isomers, but also increased the quantum efficiency of conversion *via* absorption in the  $n-\pi^*$  band.<sup>60</sup> A range of tetra-*ortho*-fluoroazobenzene derivatives were subsequently synthesised and their properties investigated (Figure 24).<sup>61</sup>

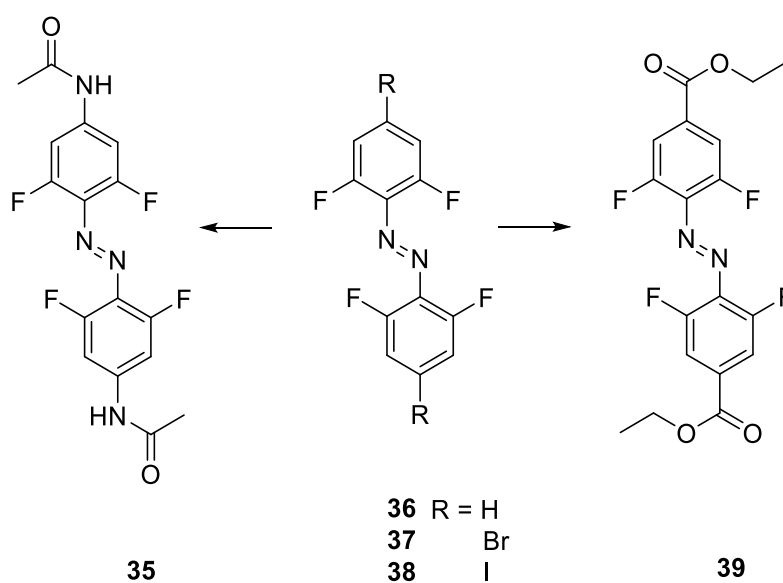
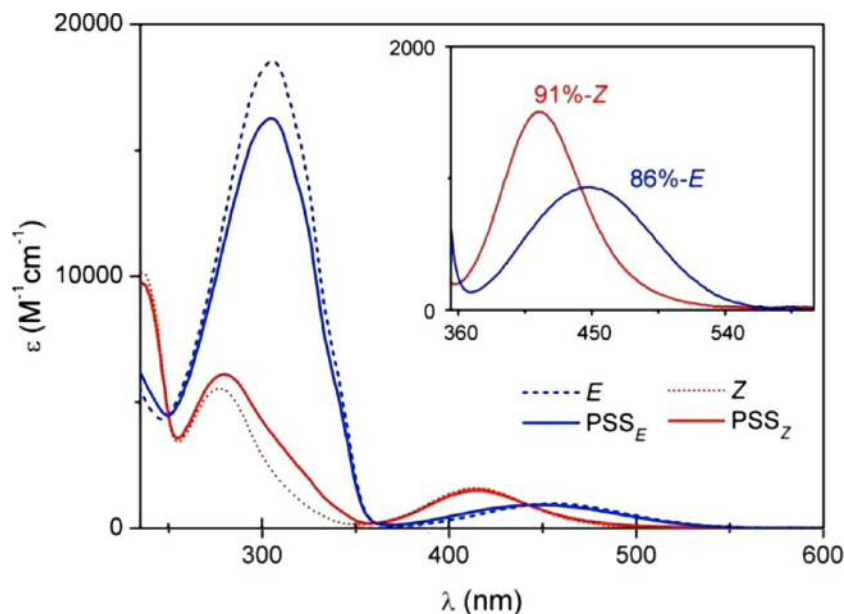


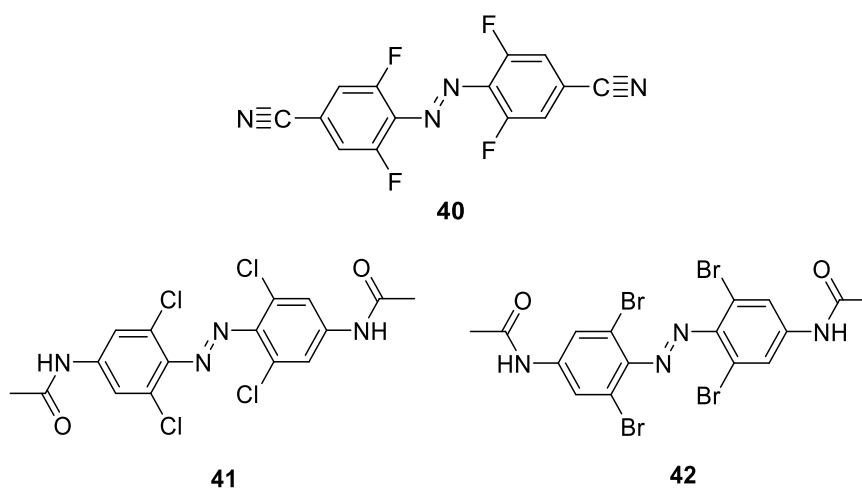
Figure 24. Tetra-*ortho*-fluoroazobenzene derivatives.<sup>61</sup>

Among them are the first examples of azobenzenes with a separation between the  $n-\pi^*$  bands of the *trans* and *cis* isomers greater than 50 nm (Figure 25).<sup>61</sup>



**Figure 25.** UV/visible spectra of *cis*-**36** (solid red) and *trans*-**36** (solid blue) photostationary states and calculated spectra for pure states (dashed red and blue, modified from reference 61).

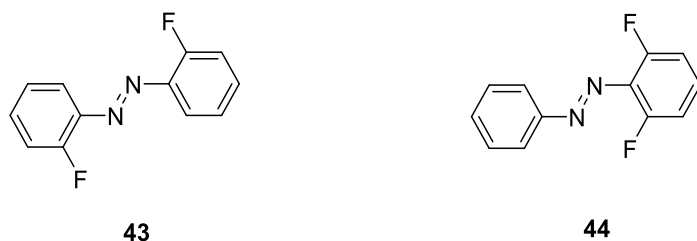
They also have extremely long lived *cis* isomers with *cis*-**39** in having a half-life of 700 days at 25 °C in dimethylsulfoxide.<sup>61</sup>



**Figure 26.** Tetra-*ortho*-halo substituted azobenzenes.

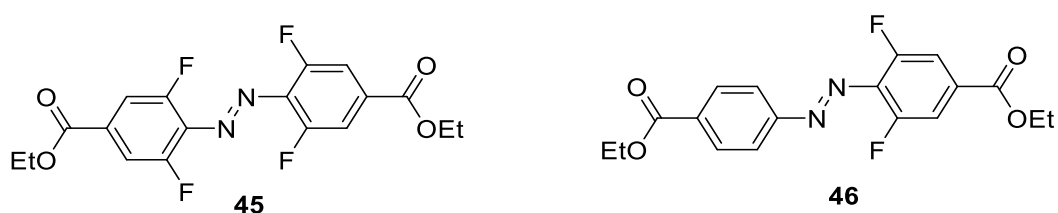
Other tetra-*ortho*-haloazobenzenes **41** and **42** show a similar band separation to that of the tetra-*ortho*-fluoroazobenzene, albeit with shorter *cis* isomer half-lives.<sup>62</sup> The addition of the *para*-chloroacetamide red-shifts their absorption spectra resulting in **41** and **42** photoswitching under irradiation with 600-630 nm light.<sup>62</sup> Calculations suggest

that the *cis* isomers of **41** and **42** are long lived due to the steric bulk of the halogens raising the barrier to interconversion which also makes the *trans* isomers non-planar, as previously observed with tetra-*ortho*-alkyl substituents.<sup>63</sup> However, these calculations cannot explain the extended half-life of *cis*-**40**.<sup>61</sup>



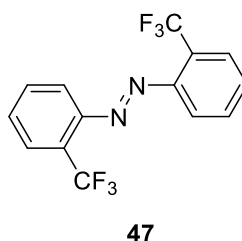
**Figure 27.** Non-symmetrical difluoroazobenzene **50** and symmetric difluoroazobenzene **44**.<sup>60</sup>

The difluoro symmetric and non-symmetric azobenzenes (Figure 27) show reduced band separations of 28 nm and 32 nm respectively. These results in a greater proportion of *cis* isomer at equilibrium for **43** than **44** due its larger band separation.<sup>60</sup>



**Figure 28.** 2,6-Difluoro and tetra-2,2',6,6'-fluoroazobenzenes.

This is also observed for the non-symmetrical di-*ortho*-fluoro **46** which has a band separation of 44 nm compared to 42 nm for tetra-*ortho*-fluoro substituted **45** (Figure 28), but these show widely varying proportions of their *cis* isomers at their photostationary states (90% and 61% respectively).<sup>60</sup>



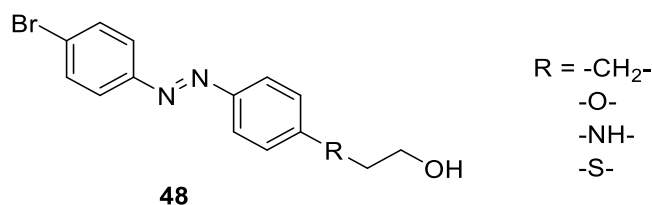
**Figure 29.** 1,2-Bis-(2-(trifluoromethyl)phenyl)diazene.

Exchanging the fluorine atom for trifluoromethyl groups which are solely electron withdrawing, with no back donation from the p orbital of the trifluoromethyl

(Hammett constants of  $\sigma_{\text{para}(\text{CF}_3)} = 0.54$  verses  $\sigma_{\text{para}(\text{F})} = 0.06$ ) resulted in **47**, with a reduced  $n-\pi^*$  band separation of 24 nm. However, these absorption bands are also very broad and did not allow selective isomerisation.<sup>60</sup> Effective band separation is therefore not solely dependent on the electron withdrawing nature of the *ortho* substituents.

### 1.5.3 Additional effects of *para* substituents

Electron withdrawing substituents *para* to the azo group increase the separation of the  $n-\pi^*$  bands of **48** to more than 40 nm by blue-shifting the *trans*  $n-\pi^*$  absorbance band and red-shifting the *cis*  $n-\pi^*$  band.<sup>62</sup> This effect lowers the energy of the LUMO of the *cis* isomer more than that of the *trans* form.<sup>60</sup> Electron donating groups have the opposite effect on the band separation; nitrogen substituents donate electrons into the azobenzene increasing the energy of the LUMO of the *cis* isomer less than that of the *trans* form.<sup>60</sup>

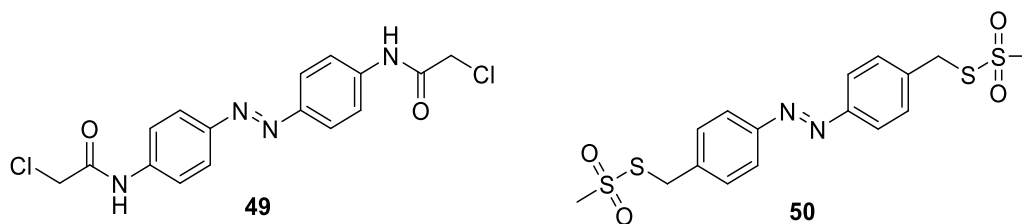


**Figure 30.** Azobenzene *para* substitution.<sup>64</sup>

<i>Para</i> -substituent	$\lambda_{\text{max}}^{\text{trans}}(\text{nm})$	<i>Trans-cis</i> conversion (%)	Half-life (s)
-CH <sub>2</sub> -	339	93	2.39 x 10 <sup>5</sup>
-O-	360	79	1.59 x 10 <sup>4</sup>
-NH-	429	34	3.30 x 10 <sup>2</sup>
-S-	377	77	1.42 x 10 <sup>4</sup>

**Table 2.** Effect of *para* heteroatom substituents on the stability of the *cis* isomer of azobenzenes.<sup>64</sup>

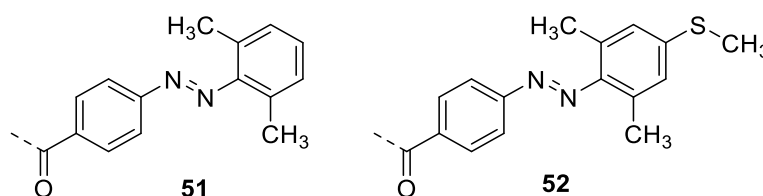
A *para* nitrogen substituent (Figure 30) causes the greatest red-shift (Table 2), this is at the expense of a greatly reduced *cis* isomer half-life and reduced conversion efficiency at its photostationary state. Sulfur and the oxygen do not shift the absorbance band as far as the nitrogen, but result in 100-fold longer *cis* isomer half-lives as well as roughly double the conversion efficiencies. The presence of a carbon at the *para* position blue shifts the  $n-\pi^*$  absorbance band, but also results in a 1000-fold increase in half-life compared to nitrogen.



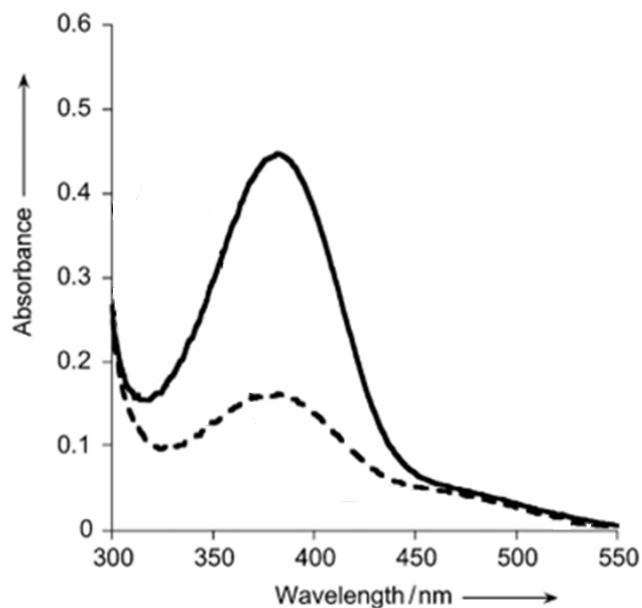
**Figure 31.** The nature of *para* substituents can have strong effects on the half-lives of *cis* isomers.

Another example of this was described by Wooley *et al.* (Figure 31);<sup>65</sup> changing the *para* substituent from the acetamide of **49** to methylene group of **50** elongated the half-life of the *cis* isomer from 8 minutes to 43 hours at 25 °C (5 mM phosphate, pH 7). This 300-fold difference transforms a photoswitch with limited *in vivo* applications into one that can remain active over the lifetime of a cell.

The introduction of a *para* thioether substituent causes a large red-shift in the absorbance maxima, but reduces the half-life of the *cis* isomer.<sup>66</sup> The shift in absorbance maxima is so large that photoswitches **51** and **52** could simultaneously be excited with different narrow bandwidth light sources.<sup>66</sup>



**Figure 32.** Installation of a thioether in the *para* position of **51** results in **52** exhibiting very different optical properties.<sup>66</sup>

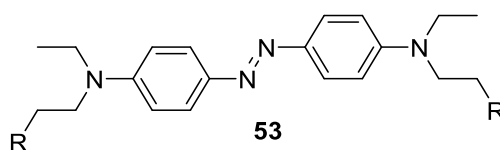


**Figure 33.** UV/visible spectra for **52** bound to DNA in buffer (10 mM phosphate pH 7.0, 100 mM sodium chloride) for the *trans* (solid) and irradiated *cis* rich states (dashed).<sup>66</sup>

**Table 3.** Half lives of *cis* isomers of **52** and **51** bound to DNA (5  $\mu$ m in 10 mM phosphate, pH 7, 100 mM sodium chloride).<sup>66</sup>

Azobenzene derivative	<i>Trans</i> at PSS	$t_{1/2}$ of <i>cis</i> [h] <sup>[c]</sup> 60 °C
<b>52</b>	48.3	6.4
<b>51</b>	50.9	25

The physical and optical properties of some azobenzenes are highly pH sensitive even without direct stabilisation of the azonium ions discussed in section 1.5.1. As for *ortho* amino substituents **25-30**, protonation of the basic nitrogen at the *para* position of **53** causes a change in the electronic structure altering the absorbance; a shift of up to 100 nm can be induced by varying the pH of the system. Not only does this alter the absorbance maxima of **53**, it also significantly changes the half-life of the *cis* isomer of this switch.



**Figure 34.** Azobenzene **53** has pH dependent UV/visible absorption spectra.<sup>67</sup>



### 1.5.4 Push-pull azobenzenes

Long-lived *cis*-isomers are ideal for controlling cell functions that take large proportions of the life cycle of a cell, such as transcription and some protein binding events, but azobenzenes with much shorter half-lives are desirable to control fast cell functions such as those of ion channels. Such azobenzenes are often non-symmetrical “push-pull” systems that have one electron deficient ring and one electron rich ring. This lowers the barrier to thermal revision giving short half-lives for the *cis* isomers.

### 1.5.5 Glutathione reduction

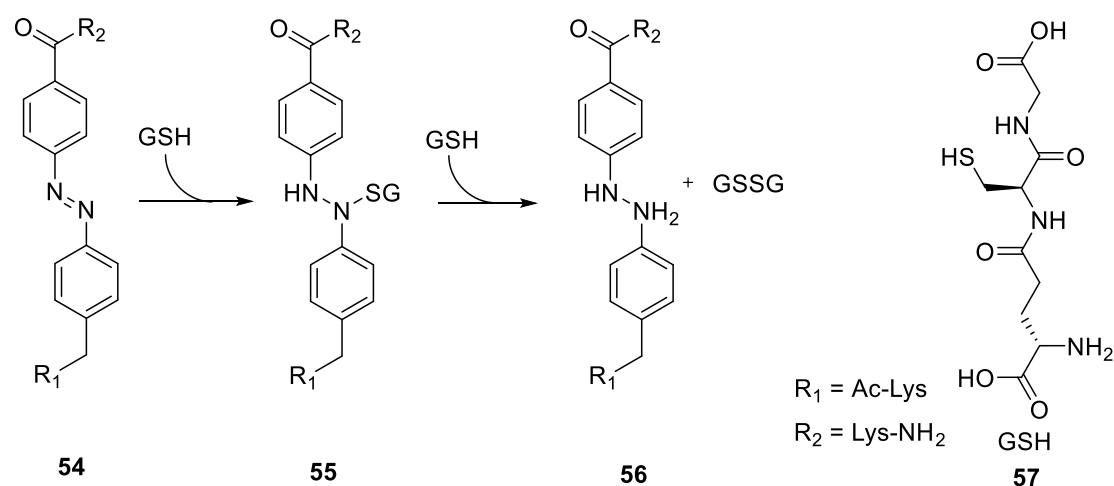


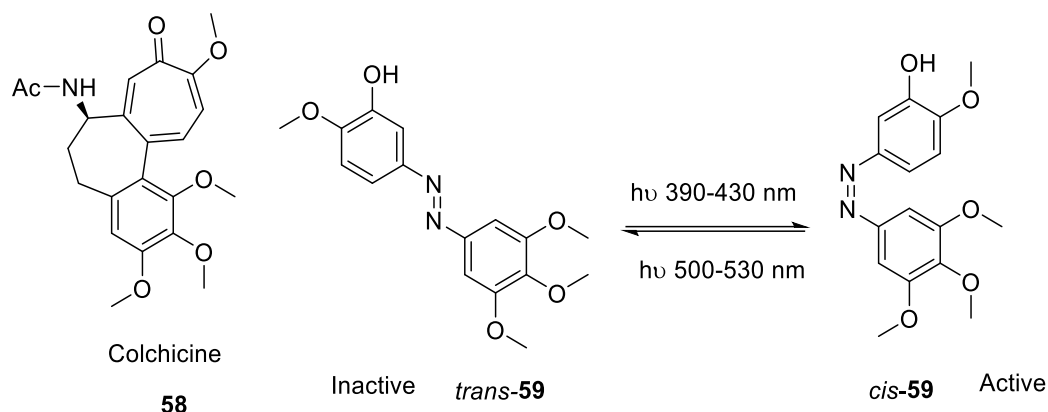
Figure 35. Glutathione(64), an intracellular redox buffer, reducing compound 54.<sup>68</sup>

Cells maintain a reducing potential by the balance of oxidised and reduced glutathione (Figure 35), whose reduced form can react with azo bonds.<sup>68</sup> The rate of azobenzene reduction is highly dependent on the nature of the substituents, especially at the *ortho* positions. For example, the tetra-*ortho*-methoxy substituted azobenzene 38 has a half-life of approximately one hour in the presence of reduced glutathione (10 mM)<sup>57</sup> due to the methoxy groups stabilising the azonium species as shown in section 1.5.1. Tetra-*ortho*-halo substituted azobenzenes show increased stability to glutathione reduction due to the steric bulk of the chloro and bromo substituents blocking the approach of glutathione to the azo bond.<sup>62</sup>

## 1.6 Azobenzene to control biological activity

### 1.6.1 Photopharmacology

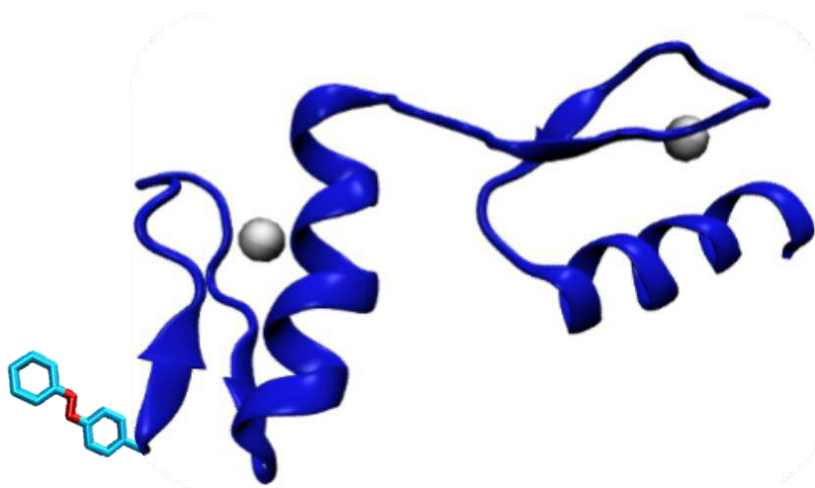
Azobenzenes can also be embedded into the structure of known drugs (Figure 36).<sup>69</sup>



**Figure 36.** Colchicine and photoswitchable colchicine mimic in the active and inactive state.<sup>69</sup>

The change in conformation of azobenzene in from can be used to control the binding of the colchicine-mimicking drug molecule **59**.<sup>69</sup> Dark state *trans*-**59** has an  $EC_{50}$  of 38  $\mu\text{M}$ , which falls to 0.5  $\mu\text{M}$  on formation of *cis*-**59** by exposure to 390 nm light. This property of **59** has been used to control mitosis and cell death,<sup>69</sup> and this concept of incorporating azobenzenes into existing drugs has also been applied to control G-protein receptors.<sup>70</sup>

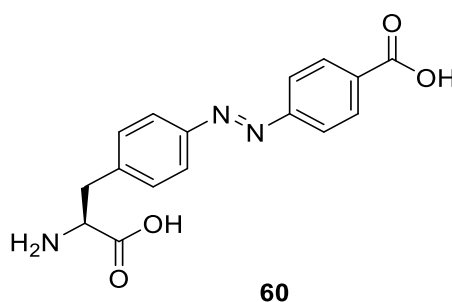
### 1.6.2 Azobenzene as a steric switch



**Figure 37.** Photo responsive tandem zinc fingers.<sup>71</sup>

To exert control over a peptide or protein conformation, azobenzene needs to be covalently attached to it in such a way as to act as a conformational switch. A single point of attachment was used to conjugate azobenzene to the tail of a protein

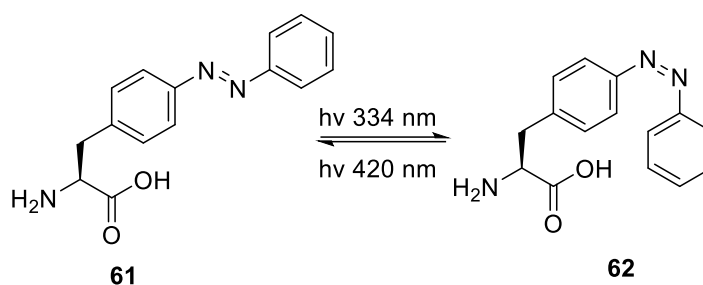
consisting of two zinc fingers.<sup>71</sup> Isomerisation of the azobenzene altered the binding of the zinc fingers to DNA by a factor of 2 in a gel mobility shift assay. Circular dichroism (CD) spectroscopy of the protein with the photoswitch as its *cis* and *trans* isomers revealed little difference in the conformation of the protein, implying a simple steric effect and not a structure change upon azobenzene switching. A pendant azobenzene installed at the end of a polypeptide has also been used to control  $\alpha$ -helical character of poly-L-glutamates.<sup>72</sup> Changes in the CD spectra signals at 208 and 222 nm indicative of  $\alpha$ -helical character were observed, although they were strongly influenced by the pH of the solution. The pK<sub>a</sub> of the glutamates were found to be dependent on the conformation of the azobenzene; when azobenzene was present as its *trans* isomer they had a pK<sub>a</sub> of 6.8 which fell to 6.3 with the *cis* isomer.<sup>72</sup> In addition, in the dark state the presence of the hydrophobic azobenzene caused the peptides to act as surfactant, but once irradiated the shape change lead to a reduction in these hydrophobic interactions and the formation of an  $\alpha$ -helix.



**Figure 38.** Azobenzene amino acid **60** used for incorporation into the homodimer interface of *Bam*HI and peptide backbones.

An amino acid with an azobenzene sidechain (Figure 38) was incorporated into the restriction enzyme *Bam*HI in a position where its sidechain formed part of the homodimer interface. The negative charge of the carboxylate group reduced endonuclease activity by causing unfavorable interactions with the nonpolar dimer binding site resulting in the breaking of the homodimer. These interactions were alleviated by photoswitching increasing activity of the enzyme.<sup>73</sup>

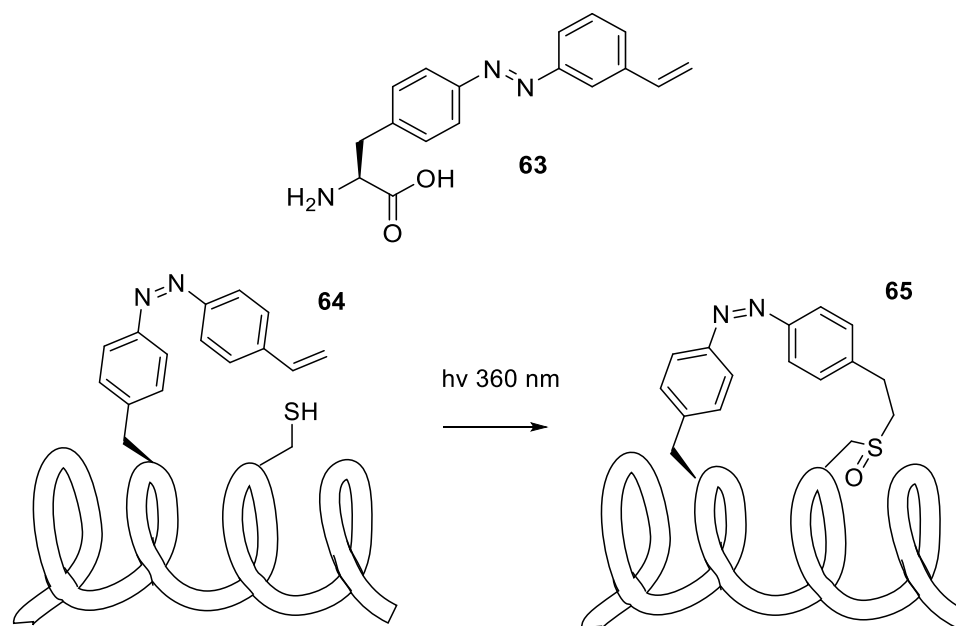
## 1.6.3 Genetically encoded incorporation of azobenzene



**Figure 39.** Photo switching of (S)-2-amino-3-(4-(phenyldiazenyl)phenyl)propanoic acid (**61**).<sup>75</sup>

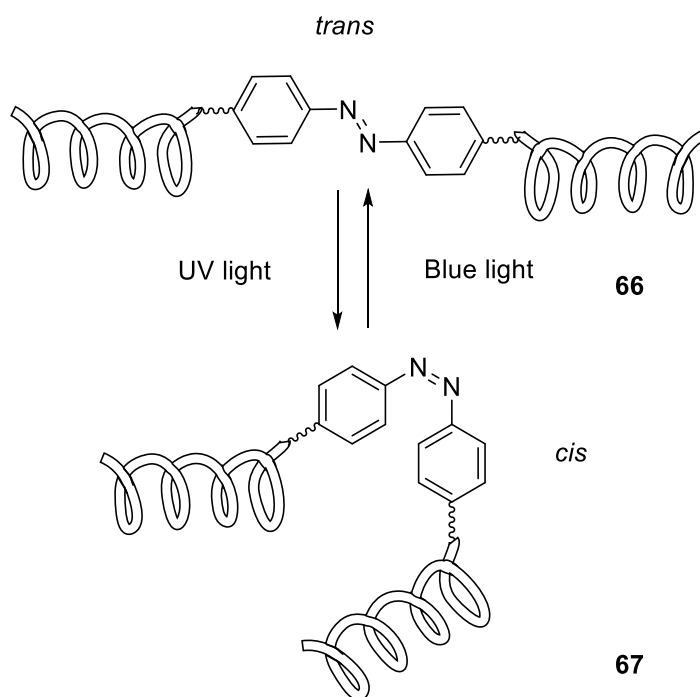
Azobenzene can be incorporated into proteins synthesised by *E. coli* using an orthogonal tRNA that recognises a less-used stop codon and a tRNA synthetase capable of priming this tRNA with an azobenzene-containing amino acid.<sup>75</sup> This allowed the direct incorporation of azobenzene into catabolite activator protein 9 (CAP 9, a transcription factor) which showed a four-fold difference between the binding affinities of proteins with azobenzenes in the *cis* and *trans* states ( $K_b = 4.0 \times 10^6$  M to  $K_b = 1.6 \times 10^7$  M) for DNA.

This technique of incorporating unnatural amino acids into biosynthesis was further developed by Hoppmann *et al.* who used an azobenzene with a vinyl group that could be incorporated into proteins by solid state peptide synthesis<sup>76</sup> or by incorporation by *E. coli*.<sup>77</sup> The vinyl group was then used to perform an intramolecular cyclisation with the sidechain of a cysteine residue with a *i,i+4* relationship to the azobenzene amino acid using a thiol-ene click reaction.



**Figure 40.** Azobenzene amino acid **63** contains a pendant vinyl group for forming a second point of attachment to a peptide *via* intramolecular thiol-ene click chemistry.<sup>76</sup>

This thiol-ene reaction unexpectedly caused the oxidation of the sulfur of the cysteine sidechain to a sulfoxide, but nevertheless upon irradiation with 360 nm light a conformational change was observed by CD due to *trans* to *cis* isomerisation.<sup>76</sup>

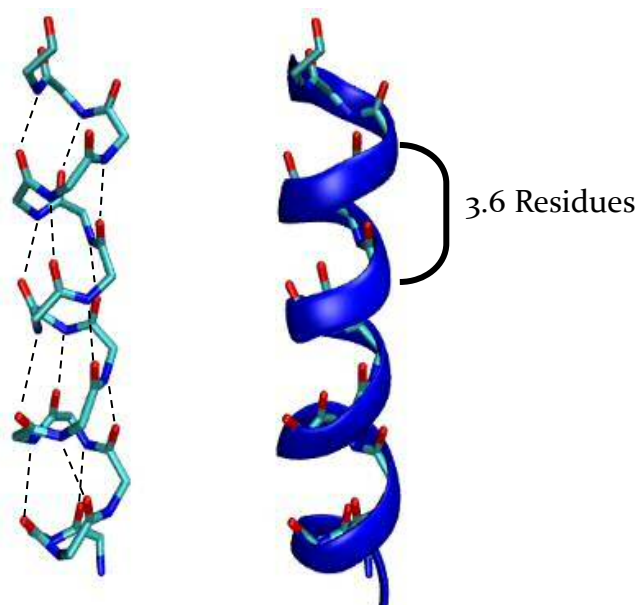


**Figure 41.** Conformational control of protein structure by azobenzene.<sup>74</sup>

Azobenzenes can also be placed directly between two proteins or peptides so that isomerisation causes changes in the secondary and tertiary structure (Figure 41). This approach was demonstrated by Trauner *et al*, with the *cis* isomer of a custom design peptide having a binding  $EC_{50}$  of 262 nM whilst the *trans* had an  $EC_{50}$  of 994 nM to a GLP-1R protein.<sup>74</sup>

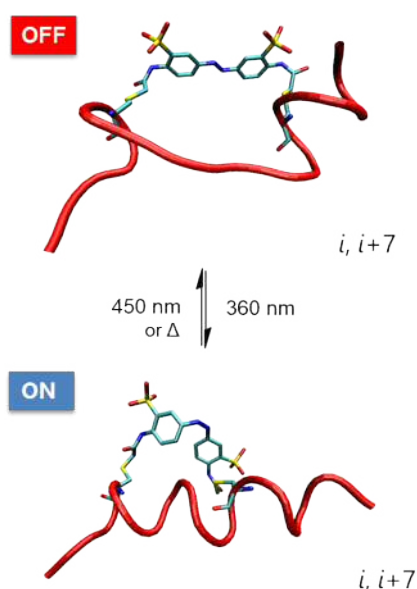
#### 1.6.4 Conjugating azobenzenes to pairs of cysteine sidechains

The shape of a protein can determine its binding affinity for other proteins, DNA, etc. so influencing the structure can control the binding affinity. In order to control this structure an azobenzene needs to be incorporated in an appropriate position to control key structural elements of the protein. An  $\alpha$ -helix is a coil of amino acids stabilised by the hydrogen bonding of carbonyl groups to amide groups four residues further along the polypeptide chain which fixes the helical register at 3.6 residues per turn. The hydrophobicity of the protein backbone is concentrated at the centre. The tendency of a peptide or protein to form  $\alpha$ -helices is dependent on its constituent amino acids. Amino acids such as alanine, lysine, methionine, glutamate and leucine favour helical structure as they support hydrogen bonding of the carbonyl to the amide nitrogen without disruptive steric bulk. Proline, for example, is highly disruptive as it lacks an amide proton to form hydrogen bonds.



**Figure 42.** Protein  $\alpha$ -helix showing hydrogen bonds (left) helical structure (right).

In order to achieve tighter control of the conformation of a  $\alpha$ -helical peptide, azobenzene can be bound to two points of the protein or peptide so that the  $\alpha$ -helix is encouraged to form or unfold with the change in shape of the azobenzene. This can be achieved in a number of ways, but most commonly two cysteine residues are incorporated into the protein or peptide of interest by site directed mutagenesis or solid phase peptide synthesis. The helical register causes the amino acid sidechains to point in the same direction every 3.6 amino acids, so the  $i, i+4$ ,  $i, i+7$  and  $i, i+11$  residue pairs sidechains project along the same face of the  $\alpha$ -helix. The spacing between cysteine residues used to attach the photo switch is chosen depending on the desired outcome of the switching. Two directions of control are possible; closer  $i, i+4$  or  $i, i+7$  cysteine spacings are often used, typically with chloroacetamidoazobenzenes, so that *trans* azobenzene disfavours formation of an  $\alpha$ -helical conformation. Conformational changes upon isomerisation to *cis* azobenzene favour the peptide to adopting a more  $\alpha$ -helical conformation (Figure 43).<sup>78</sup>



**Figure 43.** Azobenzene attached to  $i, i+7$  spaced residues.

A wider  $i, i+11$  spacing allows switching in the reverse sense, where the dark state azobenzene stabilises the  $\alpha$ -helix, but irradiation reduces this stabilisation (Figure 44).

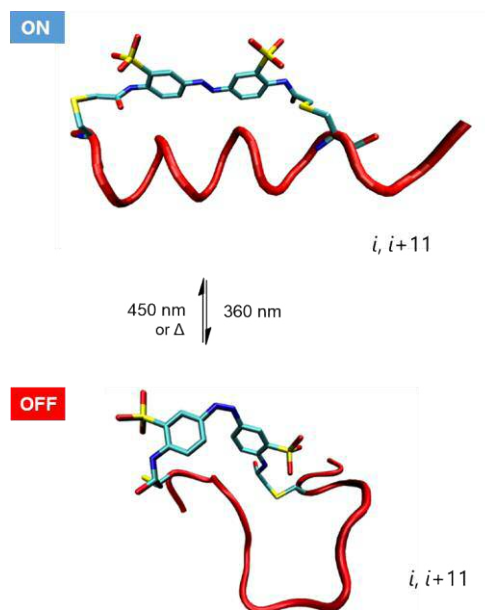


Figure 44. Azobenzene attached to  $i, i+11$  spaced residues.

The residues in between the anchor points have little or no effect on the stability of the *cis* isomer, even large sidechains such as those of valine and isoleucine show no appreciable disadvantage over alanine for the  $i, i+7$  spacing.<sup>79</sup> Residues lying between the cysteines usually point away from the cross linker, limiting their steric impact.<sup>79</sup>

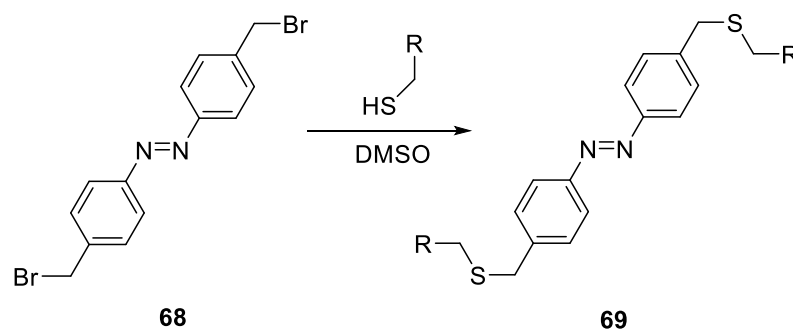
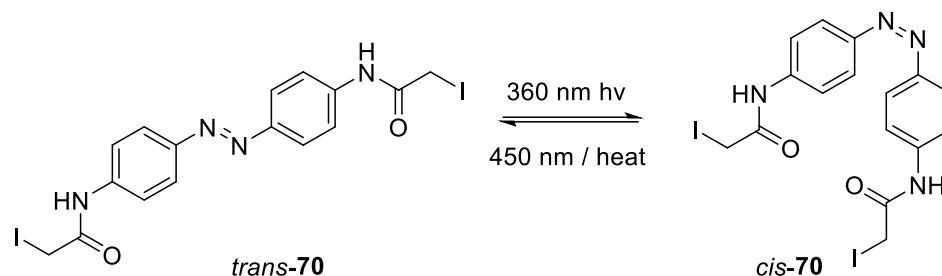


Figure 45. 4-(Bromomethyl)azobenzene reacting with thiols.<sup>80</sup>

(Bromomethyl)azobenzene **70** was used to control protein translocation by the SecYEG complex by  $S_N2$  reactions with the sidechains of cysteine residues.<sup>80</sup> Azobenzene **68** is poorly soluble in water, so the protein and azobenzene were dissolved in dimethyl sulfoxide whereupon a rapid  $S_N2$  reaction took place. Using an excess of **68** with a protein with a single available cysteine resulted in a pendant crosslinker upon which, after purification to remove excess unreacted **68**, could then be reacted with a second protein.<sup>80</sup> Haloacetamidoazobenzenes (Figure 46) have become the most popular



choice for the attachment of azobenzene to proteins due to their stability to buffered aqueous solutions and selective reactivity with cysteine sidechains once the pH of the solution is raised to 8.5.<sup>65,81,82</sup>

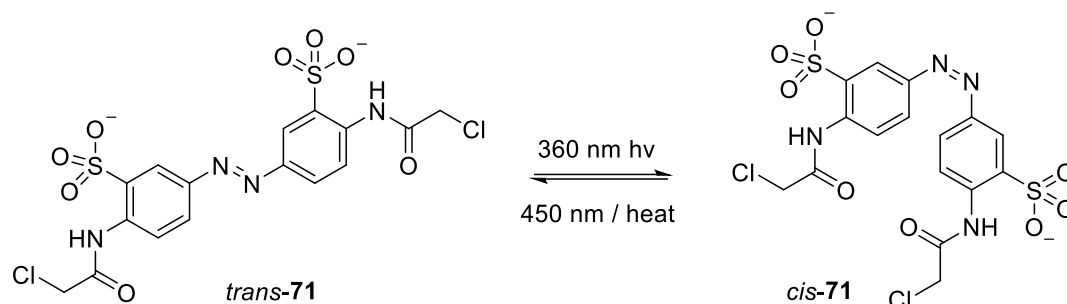


**Figure 46.** An azobenzene bearing an iodoacetamide anchoring group.

Reactions of **70** with peptides were performed in a solution of thirty percent dimethyl sulfoxide in water to overcome the limited solubility of the crosslinker. The half-life of *cis*-**70** is relatively short, 12 minutes at 25 °C,<sup>65</sup> and complete conversion to the *cis* state is not possible due to the overlap of UV spectra of the *cis* and *trans* isomer. However, once attached to a peptide containing *i, i+11* spaced cysteines, changes in the  $\alpha$ -helical character of the peptide upon irradiation were observed by CD spectroscopy with a change in degree of helicity from 65% in the *trans* state to only 35% helicity in the *cis* state.<sup>81</sup>

One of the major problems with planar cross linkers is their often poor solubility, which can limit the scope of the peptides that azobenzene can be attached and remain soluble in water. The water solubility of azobenzene can be improved by the addition of polar groups so that even moderately hydrophobic peptides remain soluble even after the cross linker has been attached. As a result **71**, whose sulfonate groups render it water soluble (Figure 47), has become the most widely used azobenzene switch, allowing

simple attachment of the cross linker to a peptide under mild conditions (50 mM phosphate buffer, pH 8.5).<sup>4,81</sup>



**Figure 47.** Water soluble, cysteine reactive 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido)azobenzene.<sup>4</sup>

The sulfonate groups have a positive effect on the half-life of the *cis* isomer of the switch compared to its unsubstituted equivalent when attached to a test peptide (Fk-11, Table 4).

**Table 4.** Half-lives of the *cis* isomers of Fk-11-71 and Fk-11-70.

	Half-life of <i>cis</i> isomers (minutes)	
	25 °C	37 °C
Fk-11-71	35 +/- 2	12 +/- 0.5
Fk-11-70	12.3 +/- 2	25 +/- 0.2

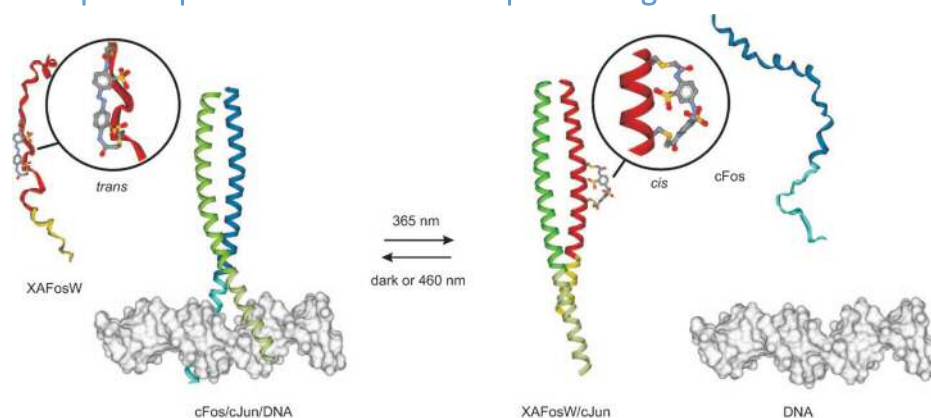
This improvement in half-life is thought to be a result of a larger barrier to rotation,<sup>83</sup> a hypothesis supported by the fact that the sulfonated molecule forms an 85% *cis* photostationary state after 5 minutes irradiation by 370 nm light (5 mW). The main two drawbacks with this switch are the UV light required to switch the azobenzene from the *trans* to the *cis* state and the relatively short half-life of the *cis* state at physiological temperatures. The later increases the frequency at which cells need to be irradiated to maintain significant proportions of the *cis* isomer.<sup>83</sup> Photocontrol of  $\alpha$ -helical peptides has also been used to target proteins that control cell apoptosis.<sup>12</sup> Small 16-mer peptides were modified with different spacing of  $i, i+7$  and  $i, i+11$ . As explained in section 1.62, these allow the control of the degree of  $\alpha$ -helicity once conjugated with a suitable azobenzene crosslinker, which is reflected in their binding affinities for Bcl-xL.<sup>12</sup>

**Table 5.** Binding affinities of Bak-7-74 and Bak-11-74 binding to Bcl-x<sub>L</sub>.

Peptide	Binding affinity (nM)	
	Bak- <i>i,i+11</i>	Bak- <i>i,i+7</i>
Non-crosslinked	328	134
Crosslinked (dark)	21	825
Crosslinked (irradiated)	48	42

The binding affinity of the Bak *i,i+7* peptide changes by a factor of 20 even with incomplete switching of the cross linker at the photostationary state. Whilst **71** has now been found to be useful in controlling the conformation of many peptides, the sulfonate groups were found to be undesirable when using this switch to control interactions with nucleic acids as the negative charges can distort these positively charged peptides away from the optimum geometries for binding.<sup>84</sup>

### 1.7 Examples of photocontrol of transcription using azobenzenes

**Figure 48.** bZIP Proteins binding to DNA.<sup>85</sup>

Leucine zipper proteins feature a pair of  $\alpha$ -helices that form a coiled coil. Such structures are often found at dimer interfaces, eg. associating a pair of DNA-binding regions so that both bind simultaneously in the major groove (Figure 48). Since the combined avidity of the two proteins is greater than their individual affinities, DNA binding can be controlled by controlling the dimer formation. Woolley *et al.*<sup>85,86</sup> have shown that azobenzene isomerisation can control protein secondary structure and consequently DNA binding.

Controlling the binding of transcription factors is not the only way to use azobenzene to control gene expression; DNA itself can be modified with azobenzenes to control duplex formation.<sup>87,88</sup> For example, it has been shown that DNA modified with d-theroninol linkers (Figure 49<sup>66</sup>) transcribe normally with *trans* azobenzene but *cis* azobenzene can block the T7 RNA transcription site.<sup>89</sup> When two of these azobenzene are placed prior to the T7 promoter region in DNA in the *cis, cis* state only 10 % of the normal level of transcription was observed compared to greater than 70% in the *trans, trans* state.<sup>89</sup>

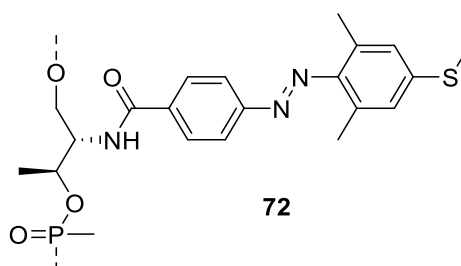


Figure 49. D-Theroninol azobenzene derivative.<sup>66</sup>

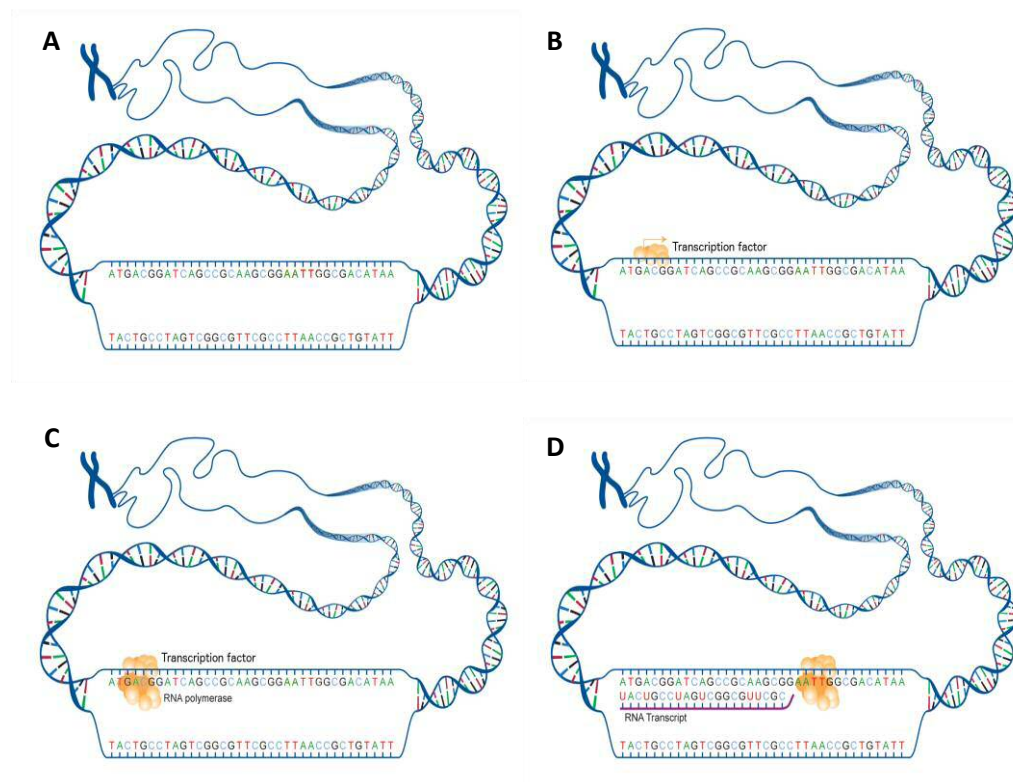
# **Chapter 2**

# **Zinc Finger Peptides**

## 2 Zinc finger peptides

### 2.1 DNA transcription

Cell function is largely controlled by protein concentrations and enzyme activities. Protein concentrations can be controlled by their rate of production or degradation. Controlling the rate of production is less metabolically wasteful and can be regulated by external or internal stimuli. Gene expression can be controlled by a large variety of proteins including the *lac* repressor for prokaryotes and CAP proteins (catabolite activator protein) for eukaryotic cells.



**Figure 50.** Graphical representation of DNA transcription (modified from Ref. 90) A) DNA strands separate locally B) A transcription factor binds its cognate DNA sequence C) RNA polymerase binds to the transcription factor D) RNA polymerase begins to assemble an RNA strand based on the DNA template.

Transcription is the process by which DNA is transcribed to pre-mRNA. This pre-mRNA is processed to mRNA which can stimulate further transcription or be translated to the required protein. In eukaryotic cells transcription is initiated by the binding of a protein called a transcription factor (Figure 50) to a specific DNA sequence as

controlling expression of the well-defined subsets of a eukaryotic genome that define cell identity and function by prokaryote-style repressors would be very wasteful. However, there are some DNA sequences that are transcribed without the need for transcription factors as they contain a promoter site which directly induces RNA polymerase binding and transcription.<sup>91</sup> These sequences typically encode for proteins that need to be constantly expressed for cell function, but they are often present at different concentrations at different stages of the cells lifetime with their expression level tightly controlled by repressors.

Prokaryotes contain a single DNA-dependent RNA polymerase, a 390 kDa complex consisting of 5 core subunits.<sup>92</sup> Together with the transcription factor, these six subunits make up the RNA polymerase holoenzyme. Unlike its eukaryotic equivalents, which pause when a mismatched base is added and can then reverse and remove errors, this complex does not contain a separate proof-reading subunit. The polymerase is therefore relatively error-prone, inserting the wrong base pair roughly every  $10^5$  to  $10^6$  bases. This limited fidelity is less critical for RNA than for DNA synthesis as mRNA is eventually broken down and the nucleotides recycled. The genome in prokaryotic cells is much smaller than that of the eukaryotes and as they lack a nucleus, transcription of DNA and translation of mRNA both occur in the cytoplasm. Genes encoding metabolic pathways in prokaryotes are typically found grouped together in the DNA. This allows expression of the entire group in a single transcript and leads to a single point of control over a range of proteins.

Eukaryotes contain three RNA polymerases; RNA polymerase I binds directly to promoters in the gene sequence to synthesise the RNA component of ribosomes and polymerase III makes tRNA and some small specialized RNAs. RNA polymerase II is the general purpose polymerase; it is much more complex than the *E. coli* polymerase but they share some highly conserved structural features. This polymerase is not active in the absence of a transcription factor.

### 2.1.1 Prokaryotic transcription control

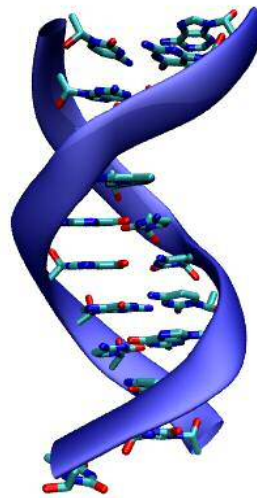
The extent of protein expression is controlled by the availability and activity of mRNA with transcription factors controlling the rate at which the RNA polymerase is recruited to particular promoter sequences to produce mRNA. Transcription is principally controlled by moderating the binding of the transcription factor to the promoter region

of the DNA or to the RNA polymerase. For example, the sigma 70 subunit (a prokaryotic transcription factor) of *E. coli* RNA polymerase binds a conserved DNA sequence at several positions before the start site of transcription. The spacing of these binding positions define the *E. coli* transcription start position; the first is at -10 with respect to the start of transcription at a 5`-TATAAT-3` sequence and the second is found at position -35 at a 5`-TTGACA-3` sequence. These sites are highly conserved but do show some variation between different strains of *E. coli*. Variation from the consensus sequence among *E. coli* genes can also impose transcription control, as a single site mutation can affect the binding of the sigma 70 subunit and vary gene expression by several orders of magnitude.<sup>92</sup> Other proteins can either activate or repress transcription by binding to these consensus sequences. For example, when glucose levels are low cAMP is formed that binds to the catabolite activator protein. This protein can bind to promoters within the lac operon to increase the transcription of enzymes necessary for the digestion of lactose, but this site is blocked by the lac repressor protein which binds to the *lac* promoter site between the RNA polymerase binding sites preventing RNA polymerase binding. If lactose is present, it binds to the lac repressor protein which then dissociates from the DNA, allowing transcription to occur. This combination of interlocked regulatory proteins ensures that enzymes necessary to metabolise lactose are only produced when glucose is absent and lactose is present. Transcription can also be repressed by proteins binding further downstream blocking the RNA polymerase as it processes along the DNA

### 2.1.2 Eukaryotic transcription control

In eukaryotic cells expression is mainly controlled by the affinity of transcription factors for their DNA target and their affinity for, and ability to, organise complexes with the RNA polymerase holoenzyme. Transcription is also strongly influenced by the acetylation or deacetylation of the histone proteins that DNA is wound around. Acetylation reduces the affinity of DNA to histone proteins increasing the availability of sequences at these sites to RNA polymerases. Conversely, deacetylation increases binding of DNA to histone proteins reducing the availability of sequences to RNA polymerases. Transcription factors can also recruit cofactor or corepressor proteins to the DNA to increase or decrease mRNA production.



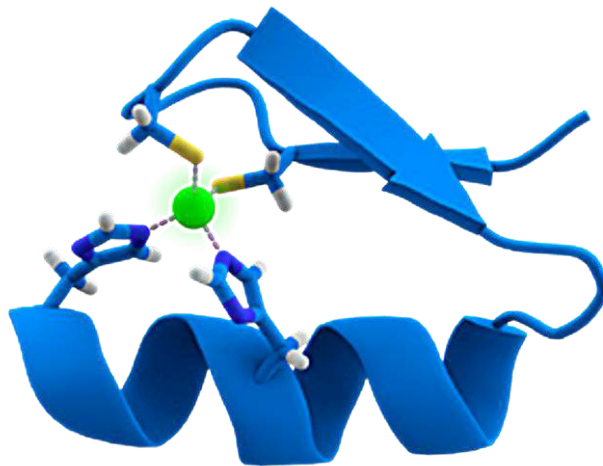


**Figure 51.** DNA  $\alpha$ -helix.

In order for transcription to occur, the RNA polymerase holoenzyme needs to bind to the DNA and form an active complex. DNA has three major helical forms, the most common of which is the “B DNA helix” which features 2 grooves (Figure 51). The minor groove exposes the glycosidic bonds of the sugar phosphate backbone and the major groove exposes the edges of the bases.<sup>92</sup> In B DNA the minor groove is 5 Å wide and 8 Å deep, not large enough to accommodate an  $\alpha$ -helix or other protein secondary structure elements, and mainly exposes functionality conserved by all DNA bases. In contrast, the major groove is 12 Å wide and 8 Å deep, large enough cavity to fit an  $\alpha$ -helix, and contains base specific hydrogen bond acceptor and donor patterns that allow bases to be recognised without the need for significant a conformational change of the DNA helix.

## 2.2 Zinc finger transcription factors

The class of transcription factors examined during this project are called zinc finger transcription factors. Zinc finger transcription factors (ZF-TF) are a wide-ranging class of proteins that variously regulate individual genes or clusters of related genes. For example, the Snail family of ZF-TFs are thought to play a critical role in cell myogenesis.<sup>93</sup> These transcription factors contain a number of individual zinc fingers (Figure 52), usually between four and six.



**Figure 52.** A single Cys<sub>2</sub>His<sub>2</sub> zinc finger is shown in blue with the sidechains of two cysteine and two histidine residues coordinating a zinc ion (green).

Shown in Figure 52 is a single Cys<sub>2</sub>His<sub>2</sub> type zinc finger made up of two short peptide strands and an  $\alpha$ -helix in a fold known as a  $\beta$ - $\beta$ -helix held together by the coordination of the zinc ion from which the fingers derive their name. The zinc ion is coordinated to the sidechains of two histidine residues of the  $\alpha$ -helix and two cysteine residues, one from each of the short peptide strands. In the absence of a zinc ion the peptide does not form an  $\alpha$ -helix and shows little if any binding affinity for DNA. Each individual zinc finger unit coordinates to three bases in an E-box transcription factor binding site. These three base pairs in the DNA are not sufficiently selective on their own but when between four and six zinc fingers are bound together in a chain, enough bases are contacted to select unique sites within the genome. Although there is some overlap between the DNA bases contacted by adjacent fingers, this can be seen as a modular design and has allowed the construction of functioning artificial zinc finger proteins with as many as eighteen separate zinc fingers. However, since overlap effects are not easily predicted large libraries of these modules have had to be created in order to get sequence specific recognition.<sup>94</sup>

### 2.2.1 Designed zinc finger proteins

Zinc fingers have other uses besides transcription factors. When coupled to an otherwise non-selective restriction endonuclease such as the catalytic domain of *FokI* the resulting construct was used to reduce the expression of mutant Huntingtin protein, one of the major factors in Huntington's disease, in mice.<sup>95</sup>

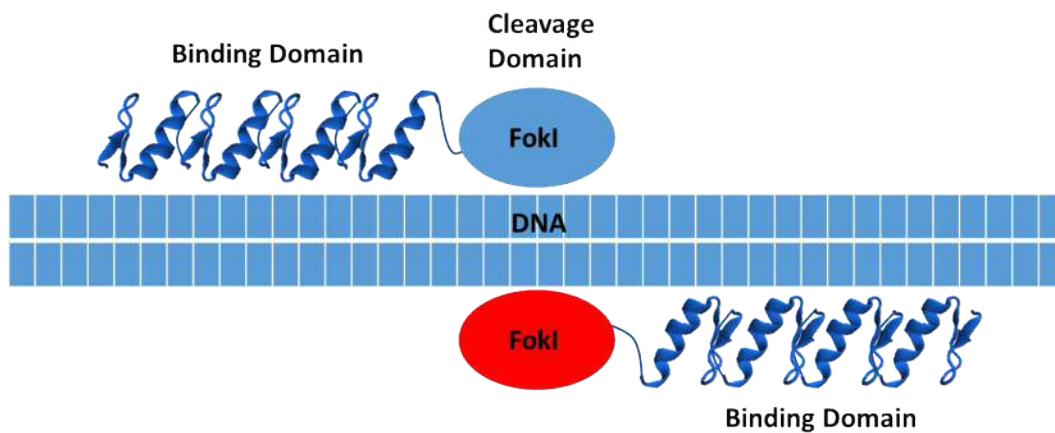


Figure 53. Zinc finger based *FokI* dimer binding.<sup>96</sup>

*FokI* restriction nuclease monomers are inactive and weakly associating and must form a dimer to cut DNA strands. The zinc finger binding to the DNA allows *FokI* to form an active homodimer at sites defined by the selectivity of the fused zinc finger and create double strand breaks (DSB, Figure 53). These fusions of zinc finger domain and a non-selective endonuclease are called zinc finger nucleases (ZFN).<sup>21,97</sup> These breaks can then be repaired by the error-prone non-homologous end joining (NHEJ).<sup>98</sup> Repair of DSB caused by ZFN can be used to effectively introduce mutations at the site of the double strand breaks.<sup>99</sup> This error-prone repair can result in frame shifts in the coding DNA causing destruction of a start codon, stop codon or simply generating an incorrect protein product. The outcome of repaired DSB can be influenced by exogenously introduced DNA matching the overhang at the cleavage site allowing the introduction of designed DNA oligonucleotides into genes.<sup>99</sup> Using tailored zinc fingers it is theoretically possible to selectively cleave at unique sites on the genome, but practically this technique is limited to palindromic target sequences by off target effects. However, it is possible to use two individual monomers with oligate heterodimer *FokI* catalytic domains, eliminating the need for a palindromic sequence and diminishing off-target effects, although it requires the identification of at least eight synergistic zinc

fingers.<sup>96,98</sup> Zinc fingers like these have been shown to be intrinsically cell penetrating, avoiding a common pitfall of peptide therapeutics.<sup>98</sup>

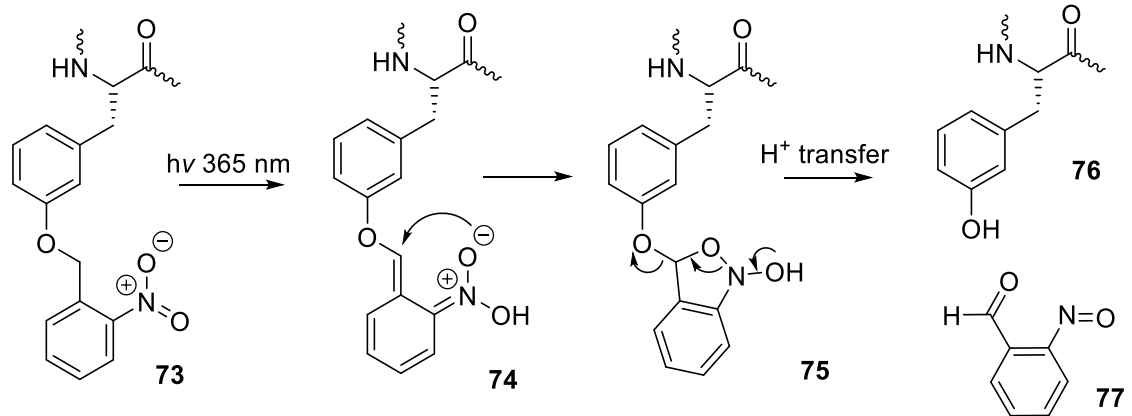
## 2.3 Photo activated zinc finger transcription factors.

### 2.3.1 Slug and snail

Initial work to determine the viability of creating photoswitchable zinc fingers was carried out on peptides derived from a zinc finger transcription factor called Slug (formally Snail2). This transcription factor was attractive because it plays a role in a number of functions in cell development,<sup>93</sup> most notably in the formation of neural crest cells,<sup>100</sup> but it also controls the expression levels of E-cadherin in adult cells. E-cadherin is a calcium dependent transmembrane protein that is responsible for adhering cells together. The level of expression of E-cadherin is one of the main factors that controls the transformation a benign cancer cell to a malignant cancer cell.<sup>101,102</sup> If a photo-responsive transcription factor targeting Slug promoter sequences could be constructed then it might be possible to trigger epithelial to mesenchymal transitions (EMT) and contribute to the debate over the role of EMT in metastasis.

## 2.4 Existing photo activated zinc finger peptides

Several photo-activated zinc finger nucleases have been reported, the most common being photo caged zinc finger nucleases<sup>103</sup> using, for example, an *ortho*-nitrobenzyl group as a bulky protecting group to prevent enzyme activity.<sup>21</sup> When subsequently removed with UV irradiation, the completely inert protein was converted into one with 70% of the activity of the wild type species.<sup>4</sup> Unfortunately, irradiation of *ortho*-nitrobenzyl groups releases nitrate by-products which can be toxic to some cells.<sup>10</sup> Combined with the harmful UV light these by-products significantly reduce cell viability.



**Figure 54.** Photo-decaging of a tyrosine residue.

This process only allows for the selective activation, not deactivation of the nuclease; a drawback common to all decaging procedures. As a single  $\alpha$ -helix is primarily responsible for ZF-TF binding to DNA it was chosen for modification to allow bidirectional control using an azobenzene photoswitch using an  $i,i+7$  cysteine spacing for low background activity.

#### 2.4.1 Design of the Slug finger peptides

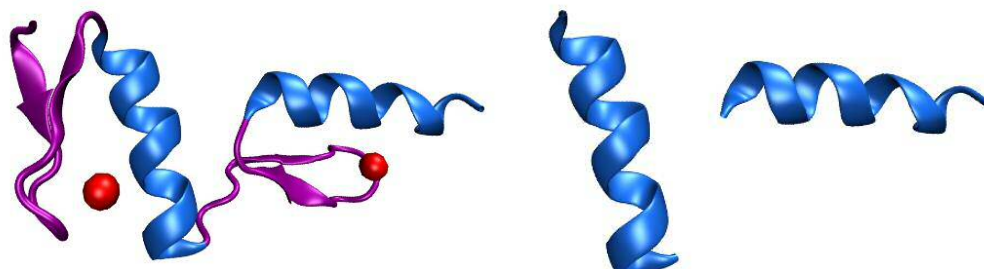
The initial aim of the project was to take the zinc finger transcription factor and excise only the residues from the  $\alpha$ -helix that bind in the major groove of DNA. The wild type Slug zinc finger protein contains four canonical Cys<sub>2</sub>His<sub>2</sub> zinc fingers plus a fifth atypical zinc finger (Figure 55).

```

MPRSFLVKKH FNASKKPNS ELDHTVIIS PYLCESYPMP VIPKPEILTS
GAYSPITVWT SAVPFHSPLP SGLSPLTGYS SSLGRVSPLP SSDTSSKDHS
GSESPISDEE ERLQPKLSDP HAIEAEKFQC NLCNKTYSTF SGLAKHKQLH
CDAQARKSFS CKYCDKEYVS LGALKMHIRT HTLPCVCKIC GKAFSRPWLL
QGHIRTHTGE KPFSCPHCNR AFADRSNLRA HLQTHSDVKK YQCKNCSKTF
SRMSLLHKHE ESGCCVAH
  
```

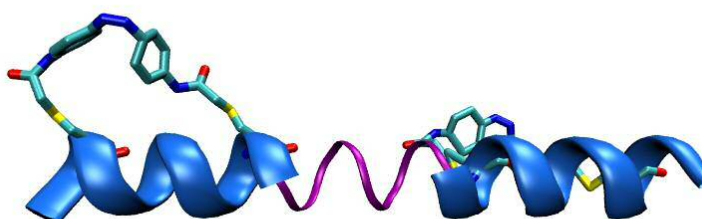
**Figure 55.** Sequence of the product of the *SNAI2* gene in *Homo sapiens* (UniProtKB: O43623). The chosen DNA-binding helices are highlighted in red and blue.

The Slug sequence was homology modelled onto the archetypal zinc finger ZIF268 (PDB 1AAY) using SwissModel online tools and the fingers that aligned most closely to the ZIF 268 sequence were selected for initial testing (Figure 56).



**Figure 56.** Cartoon of SNAl2 zinc finger residues after homology modelling their sequence onto the structure of ZIF268 (PDB 1AAY). Full schematic (left)  $\alpha$ -helical regions only (right).

The sheets and loops were removed as the photoswitch will fulfill the role of zinc ion in adding rigidity to stabilise the helix. This leaves only the short  $\alpha$ -helix regions that make the primary contacts with DNA bases during binding. Lysine and histidine residues in each helix were replaced by cysteines to create an  $i, i+7$  spaced pair for the incorporation of a photoswitch (Table 6). These residues were chosen because they face away from the DNA binding faces and should not significantly affect the binding of the peptide to DNA. An additional peptide was designed with these two helices joined by a short linker (Figure 57).



**Figure 57** Double slug with photoswitch.

	Sequence
Slug Finger 1 WT	Ac-AW <b>L</b> QGHIR <b>T</b> HTG-NH <sub>2</sub>
Slug Finger 1	Ac-AW <b>C</b> QGHIR <b>C</b> TG-NH <sub>2</sub>
Slug Finger 2 WT	Ac-RSN <b>L</b> RAHLQ <b>T</b> HSD-NH <sub>2</sub>
Slug Finger 2	Ac-RSN <b>C</b> RAHLQ <b>C</b> SD-NH <sub>2</sub>

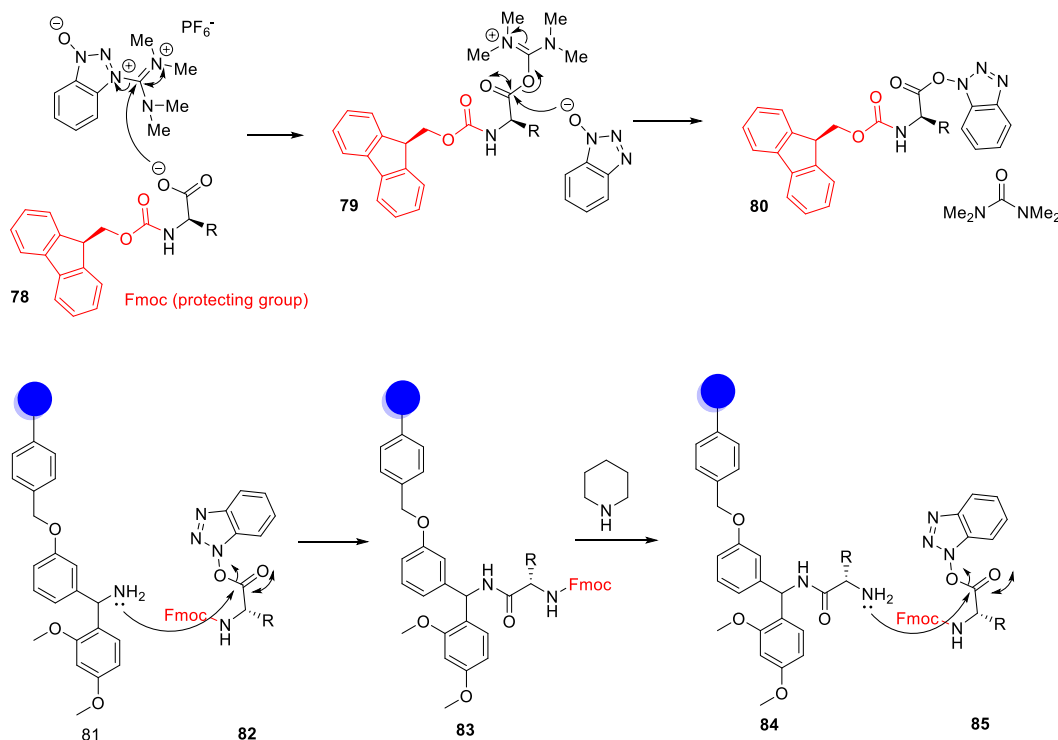
## Double Slug Finger



**Table 6.** Sequences of Slug zinc fingers (Linker in red, original residues in blue, modified in green).

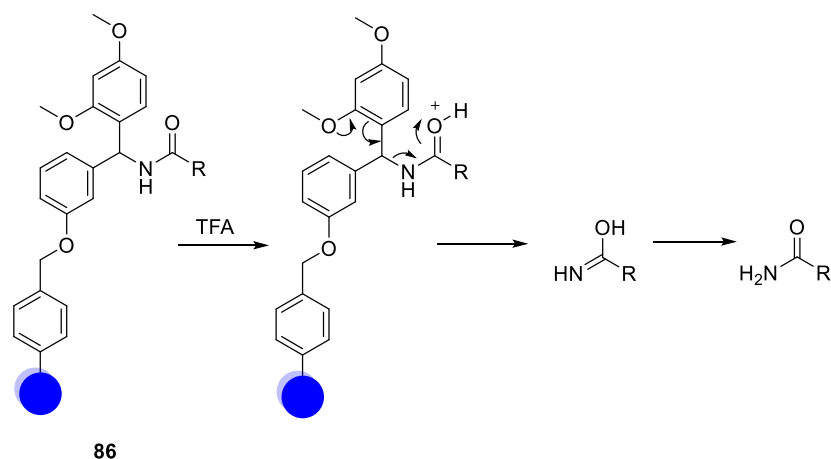
The short glycine-based linker was chosen because of its flexibility to allow the  $\alpha$ -helices to align with the major groove of the DNA. They were called Slug Finger 1, Slug Finger 2 and Slug Double Finger (Table 6).

## 2.4.2 Solid state peptide synthesis



**Figure 58.** Elementary steps of solid phase peptide synthesis.

Peptides were assembled iteratively by solid phase peptide synthesis using standard 9-fluorenylmethoxycarbonyl (Fmoc) protected amino acid building blocks with orthogonal acid-labile sidechain protecting groups. This method of forming a peptide bond is approximately 95-99 % efficient, with diminishing returns limiting its practical application to between 20 and 60 amino acids.



**Figure 59.** Cleaving of peptide from a solid support with the use of TFA (where R is the peptide).

Once chain extension was complete, the N-terminus was deprotected with piperidine, then the resin washed and amine protected as an acetamide using acetic anhydride. The peptide was then cleaved from the resin and the amino acid sidechains simultaneously deprotected with trifluoroacetic acid containing tri-*iso*-propylsilane (5%), water (5%) and 3,6-dioxa-1,8-octanedithiol (5%) scavengers (Figure 59). After half an hour, the reaction mixture was filtered to remove the spent resin and the TFA was removed under a stream of nitrogen and the residue dispersed in diethyl ether and chilled in a freezer overnight. Filtration provided a crude peptide mixture that was purified by HPLC as described in the methods section. Purified peptides were then checked by matrix-assisted laser desorption ionisation/time of flight (MALDI-TOF) mass spectroscopy to ensure they gave masses corresponding to the intended products (Table 7).

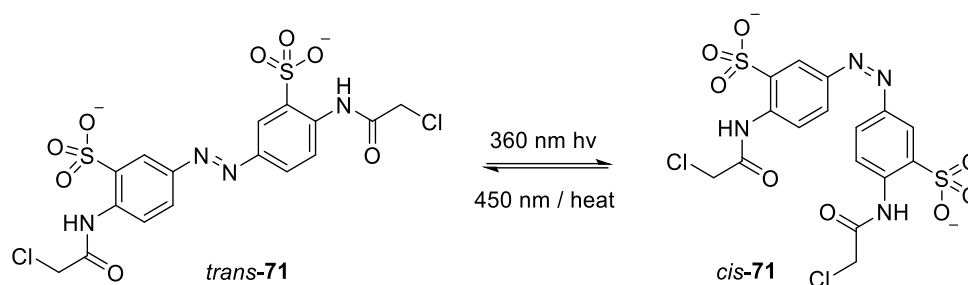
**Table 7.** Expected and observed masses for Slug peptides.

	Calculated mass	Observed mass
Slug Finger 1	1373.60	1373.44 <sup>a</sup>
Slug Finger 2	1531.72	1533.89 <sup>a</sup>
Double Slug Finger	3204.56	3185.85 [M -H <sub>2</sub> O]

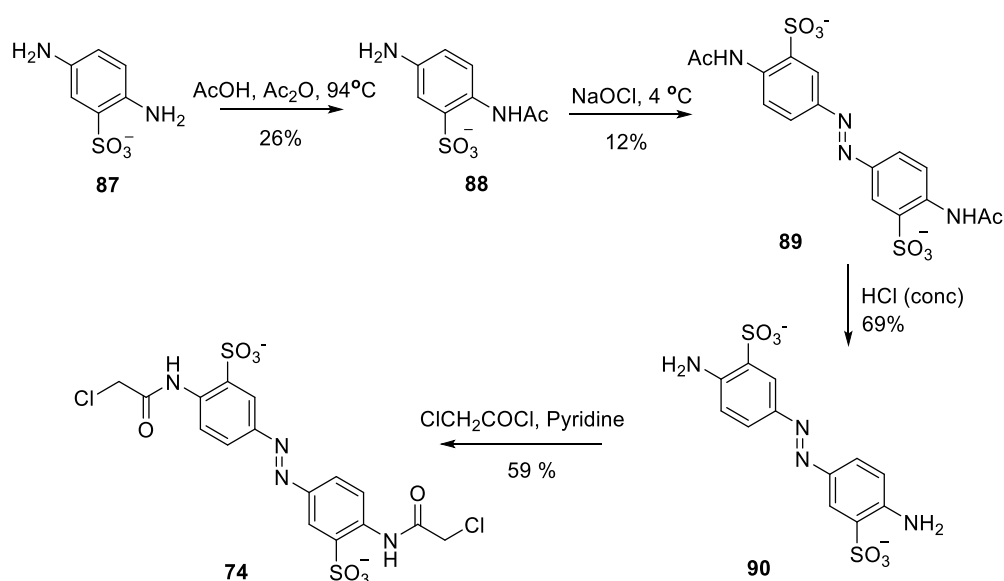
a) The masses vary due to the picking of the peak from the MALDI spectra



## 2.4.3 Synthesis of sulfonated azobenzene crosslinker

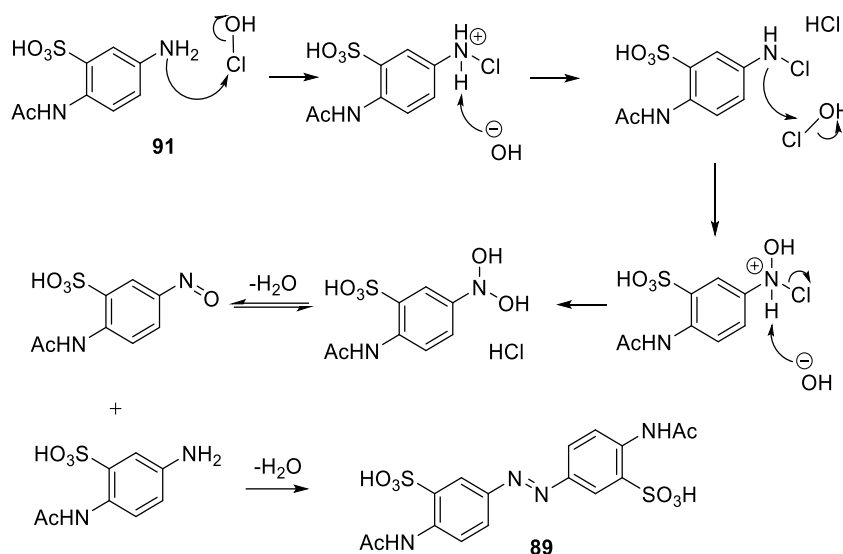


**Figure 60.** 3'-Bis(sulfonato)-4,4'-bis(chloroacetamido)azobenzene (**71**) was chosen for its water solubility.



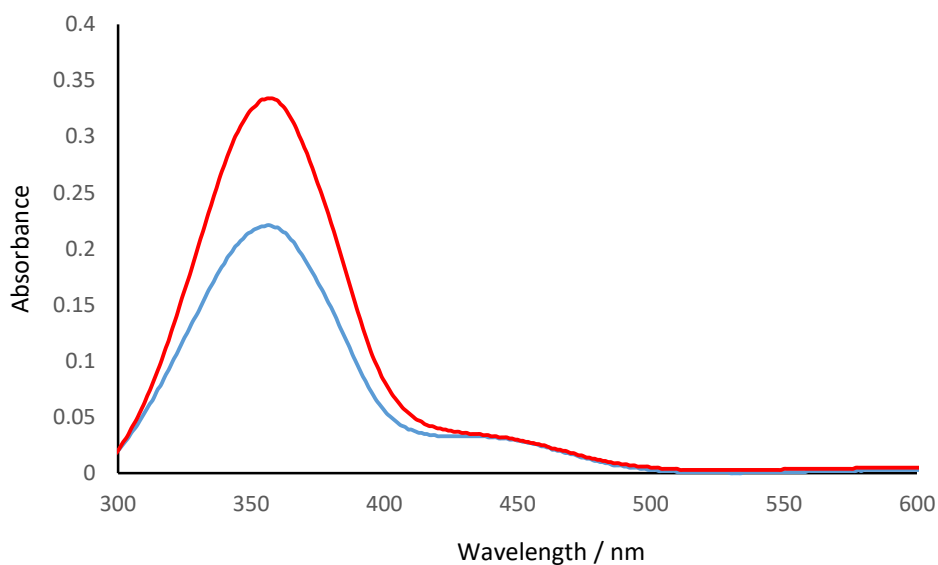
**Figure 61.** Scheme showing the synthesis of BSBCA (**71**).<sup>104</sup>

The first step is a selective acetate protection with acetic anhydride under acidic conditions. Selectivity is achieved because the sulfonate group changes the  $pK_a$  of the amine in the *ortho* position, making it harder to protonate. In initial experiments performing the reaction below 90 °C resulted in a 50:50 mix of both the singly and doubly acetylated protected species, but rigorously maintaining the temperature at 94 C greatly improved the ratio for the desired monoacetate. Oxidative coupling then converted two molecules of an aniline containing molecule (**91**) to azobenzene (Figure 62).



**Figure 62.** Mechanism for oxidative coupling of anilines to form azobenzene.

The reaction proceeds through oxidation of the amine by hypochlorite to a hydroxylamine, then to a nitroso group followed by condensation with another molecule of amine. The acetate protecting groups were then removed using hydrochloric acid and the resulting solution was neutralised and lyophilised. The free amine was then converted into the chloroacetamide by reaction with chloroacetic anhydride to furnish **71** in an overall yield of 1.3 %. The UV/visible absorbance spectra of **71** displayed distinct extinction coefficients for the  $\pi$ - $\pi^*$  transitions of the *trans* and *cis* isomers (Figure 63).

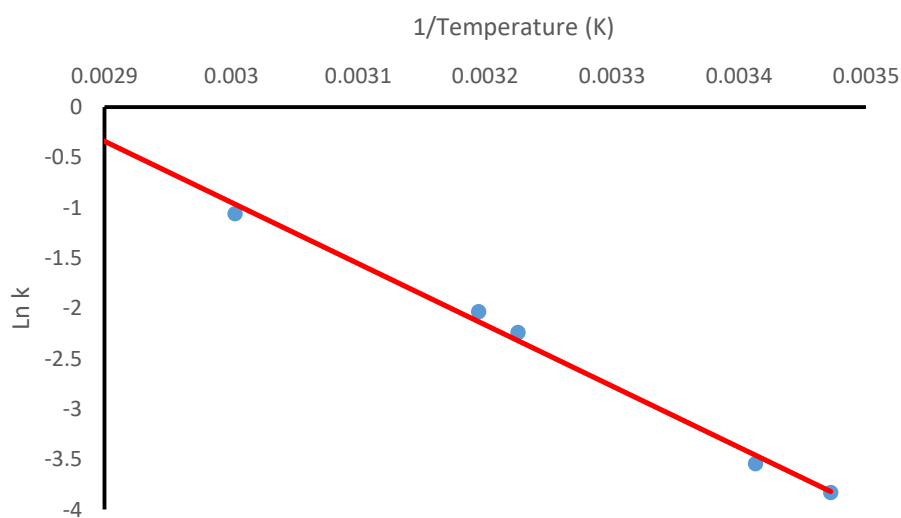


**Figure 63.** UV absorption of **71** in water in its photostationary light state (blue) and dark state (red).

The  $\pi$ - $\pi^*$  band is evident in the UV/visible absorbance spectrum at 360 nm and the n- $\pi^*$  at 450 nm.

**Table 8.** Half-lives of *cis*-**71** in water.

Temperature	Half-life (min) $\pm$ S.D.
15 °C	31.9 $\pm$ 4.7
20 °C	24.0 $\pm$ 1.1
37 °C	6.5 $\pm$ 1.3
40 °C	5.3 $\pm$ 0.9
60 °C	2.0 $\pm$ 0.4



**Figure 64.** Arrhenius plot for **71** in water.

The half-life of *cis*-**71** varies with temperature were calculated (Table 8), with an Arrhenius plot indicating an activation energy of  $\sim 50$  kJ mol<sup>-1</sup>. Based on the equation below.

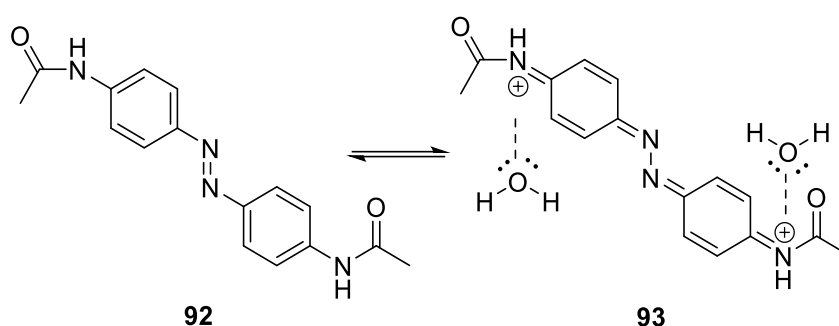


These results also hint that this switch may be of limited use *in vivo* because the half-life, even when cooled to 15 °C, is only ~30 minutes; close to the time it takes for *E. coli* to replicate. However, the extinction coefficients of samples do not show any degradation on switching or photo bleaching after 10 cycles, suggesting pulsed or constant irradiation can be used to overcome the short half-life. The half-life of the *cis*-**71** photo switch was also measured in dimethylsulfoxide (DMSO) to compare it to other switches that are not water soluble.

**Table 9.** *cis*-**71** relaxation times in DMSO.

Temperature	Half-life $\pm$ S.D. (min)
20 °C	435 $\pm$ 30
40 °C	56 $\pm$ 2
60 °C	7.7 $\pm$ 0.1

The relaxation times of *cis*-**71** are much longer in DMSO than in water. This is due to the water stabilizing of the resonance structures that cause a greater contribution of single bond character in the azo bond (Figure 65)



**Figure 65.** Iminium ion resonance structure of azobenzene stabilised by interaction with water.

The positively-charged iminium ion resonance structure can be better stabilized by hydrogen bonding from the water than DMSO. Hydrogen bond interactions can also have a great effect on the absorbance spectra of these switches as described in section 1.5.3. Compound **49** was synthesised from commercially available **94** to determine the

effect of the sulfonate groups on the half-life and the UV/visible absorbance of **71** (Figure 66).

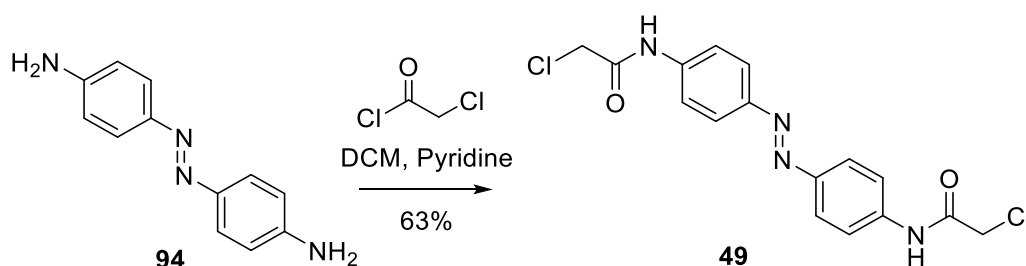


Figure 66. Acetamidoazobenzene crosslinker **49**.

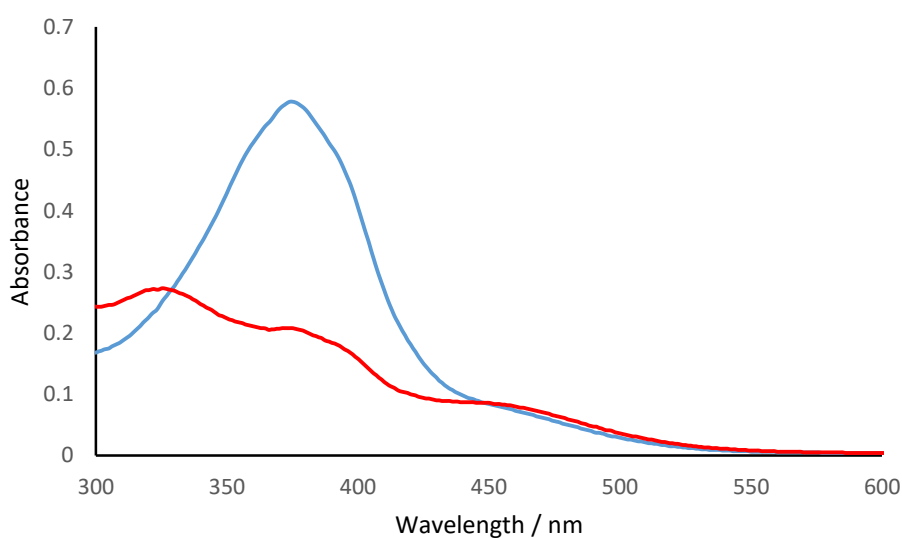


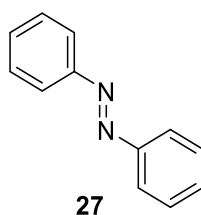
Figure 67. UV absorption of **49** in DMSO in its dark state (blue) and photostationary light state (red).

Table 10. Azobenzene relaxation times in DMSO at varying temperature.

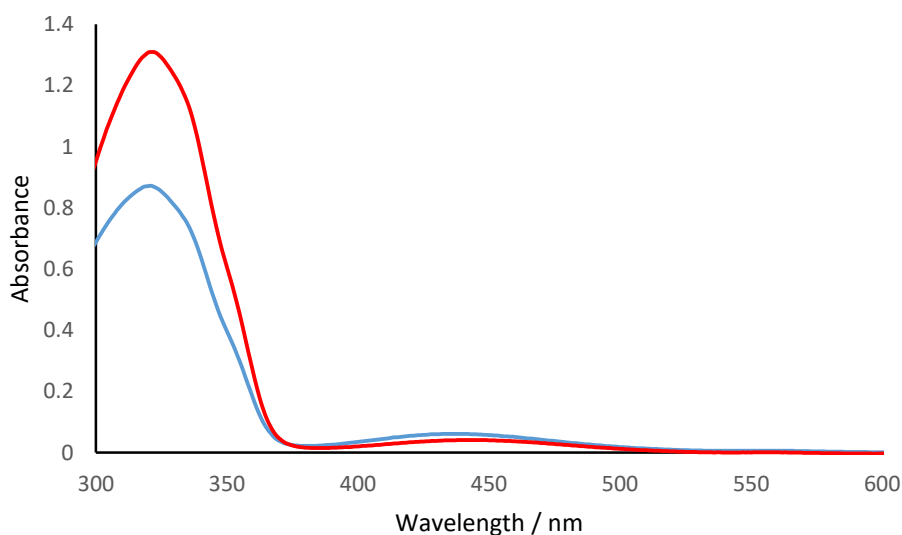
Temperature	Half-life $\pm$ S.D. (min)
20 °C	429 $\pm$ 22
40 °C	84.9 $\pm$ 3.3
60 °C	11.9 $\pm$ 0.2

Comparing the half-lives of *cis*-**71** and *cis*-**49** it is clear that the addition of the sulfonate group has only a small effect on the half-life. Therefore, the decrease in the half-life of *cis*-**71** as compared to **27** (azobenzene shown below) must be as a result of the *para*-chloroacetamide. The sulfonate shows very little shift in the UV/visible absorbance

spectrum or in the photostationary state brought about by the irradiation of 360 nm light. This shows that while the addition of the chloroacetamide group provides a convenient anchor point to the peptide, it greatly reduces the half-life and consequently the usefulness of the switch for *in vivo* control of transcription. This is emphasised by the properties of the parent azobenzene.



**Figure 68.** Azobenzene.

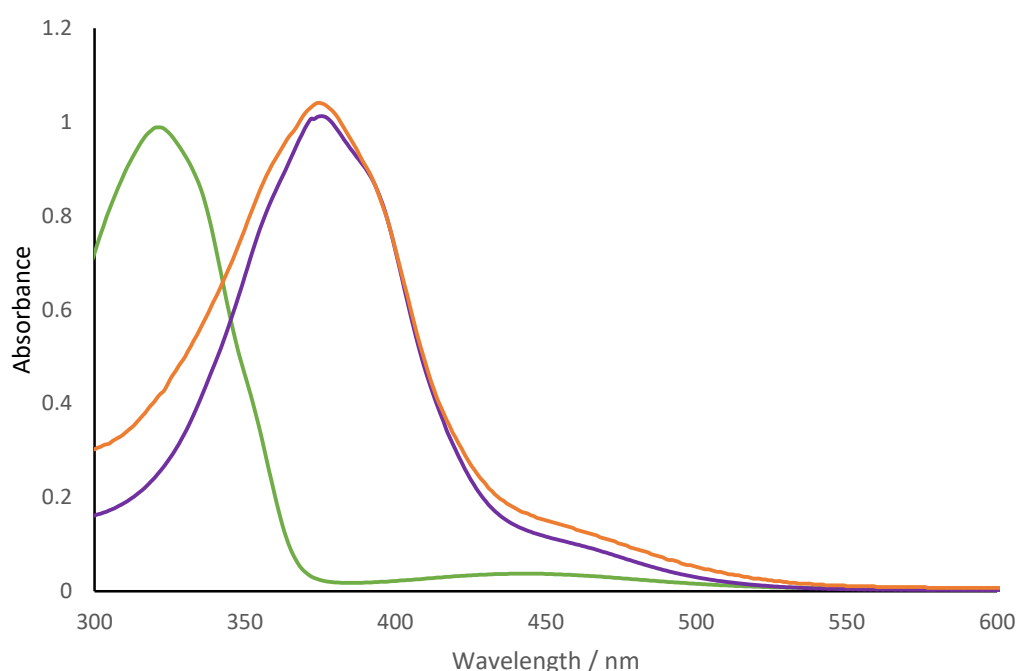


**Figure 69.** UV absorption of azobenzene in DMSO in its dark state (blue) and photostationary light state (red).

**Table 11.** Relaxation times of *cis*-27 in DMSO at different temperatures.

Temperature	Half-life $\pm$ S.D. (min)
40 °C	485.5 $\pm$ 22.7
60 °C	219.8 $\pm$ 4.6

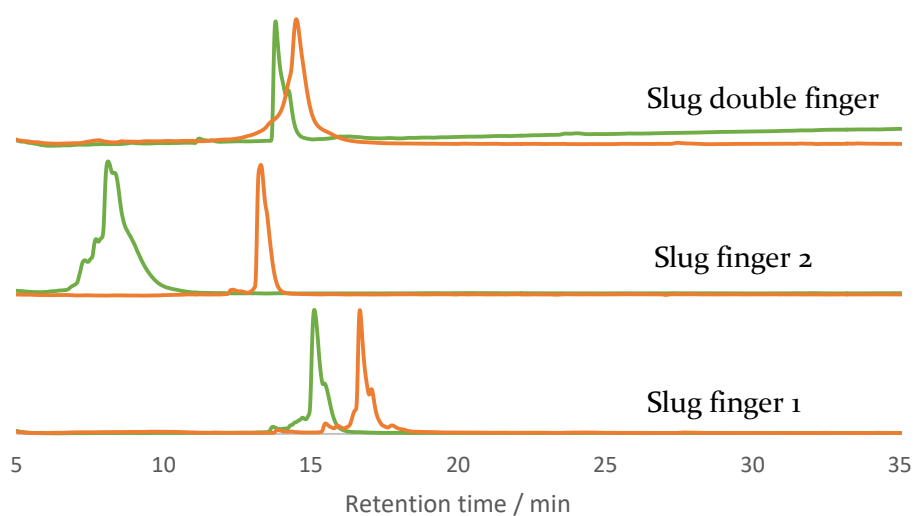
Azobenzene compound **27** showed no appreciable relaxation at 20 °C, even after 24 hours, so no relaxation curve could be plotted. At 40 °C or 60 °C there was a change but it was so small that it led to large standard deviations in readings taken. Comparing these data to *cis*-**49**, whose half-life at 60 °C is only ~12 minutes, the acetamido group diminished the *cis* isomer half-life by 20-fold. UV/visible spectra of **27** indicate the addition of the chloroacetamide in **49** causes a 40 nm redshift in the  $\pi$ - $\pi^*$  as shown below in Figure 70.



**Figure 70.** Absorbance spectra of **27** (green), **49** (purple) and **71** (orange) in DMSO.

#### 2.4.4 Purification of slug peptides

The Slug finger peptides were cross linked with BSBCA (**71**) in pH 8.5 buffer as described in the methods section, purified by HPLC and lyophilised. Solid samples were dissolved in phosphate buffer and their purity was checked by analytical HPLC showing the formation of a product with decreased polarity compared to the native peptide (Figure 71). Following separation of crosslinked peptide from the starting peptide, MALDI-TOF mass spectra were used to confirm the identity of the isolated species (Table 12).

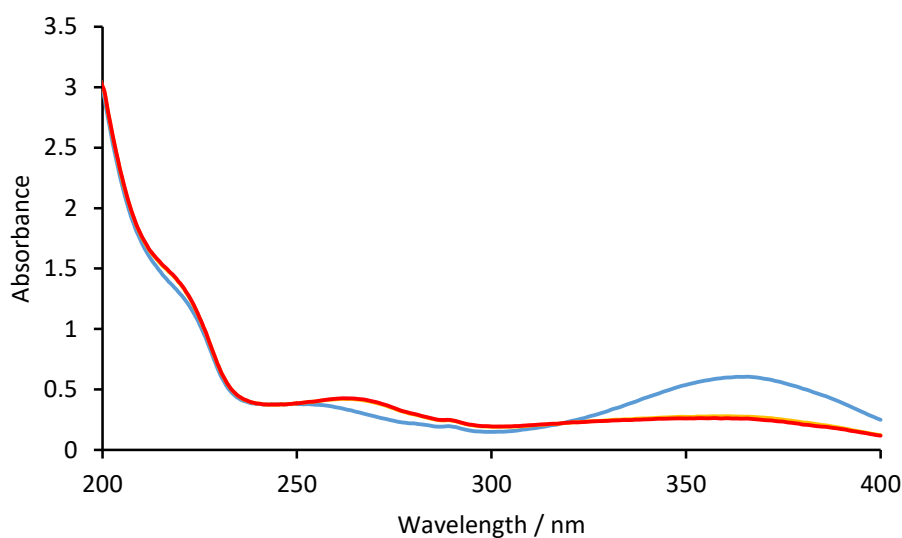


**Figure 71** HPLC chromatograms showing absorbance at 210 nm for native peptides (green) and crosslinked peptides (orange).

**Table 12.** Table of crosslinked Slug finger peptide masses.

Crosslinked peptide	Expected m/z	Obtained m/z
Slug Finger 1-71	1826.80	1825 1869 [M+2Na-H] <sup>+</sup> , 1893 M+3Na.
Slug Finger 2-71	1984.92	1984 [M-H] <sup>+</sup> 2006 [M+Na] <sup>+</sup> 2023 [M+K] <sup>+</sup>





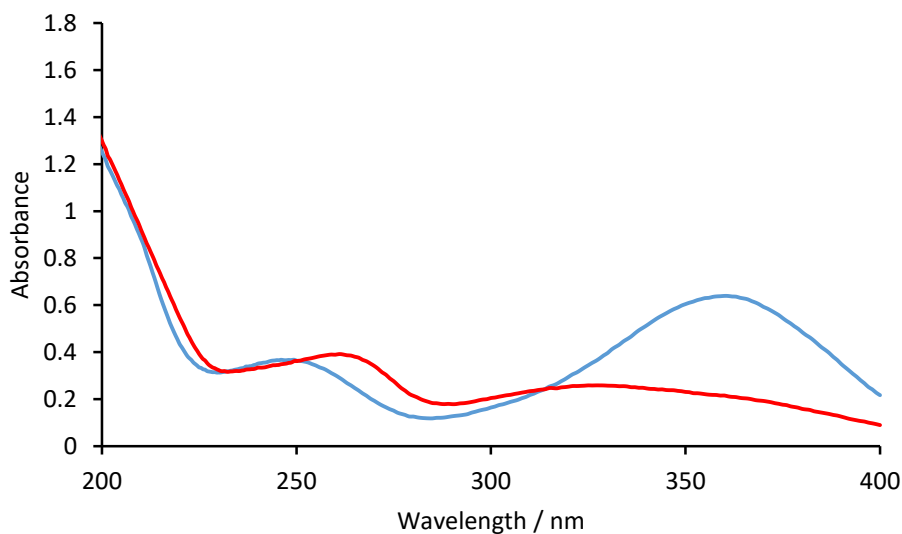
**Figure 72.** UV/visible absorption spectrum of Slug Finger 1 cross-linked with **71** the dark state (blue) and the 360 nm photostationary (red) state.

The absorbance maximum of the  $\pi$ - $\pi^*$  band of **71** is not shifted to any great extent upon binding to the peptide in this case although there is a shift in the absorbance in another band at 270-280 nm.

**Table 13.** Relaxation times of Slug Finger 1-**71**.

Temperature	Half-life (min) $\pm$ S.D.
15 °C	178 $\pm$ 17
20 °C	90 $\pm$ 15
40 °C	20.7 $\pm$ 0.4

Slug Finger 1-**71** displays a large increase in the half-life of the light state of the switch compared to *cis*-**71** alone; between 4 and 5 fold. This suggests that the  $\alpha$ -helical conformation of the peptide adds stability to the *cis* state, increasing the half-life of the switch.

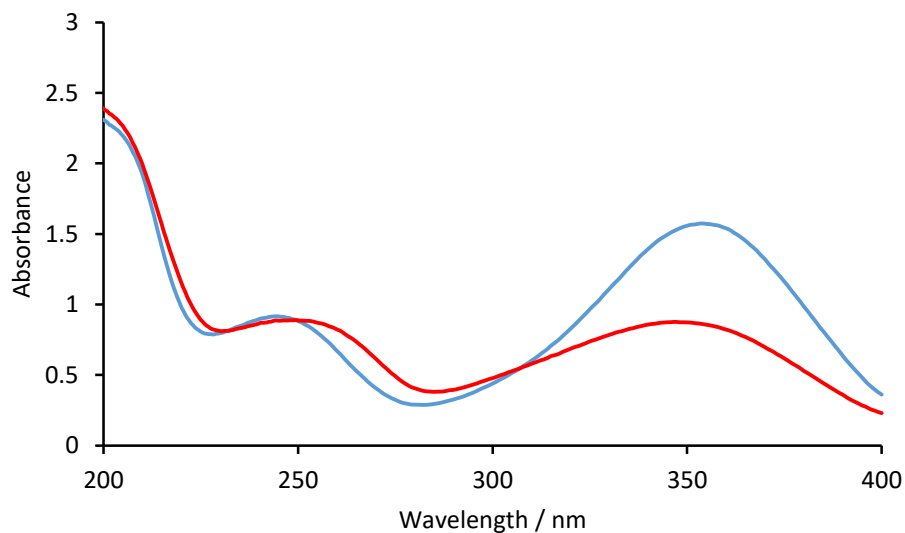


**Figure 73.** UV/visible absorption of Slug Finger 2-71 in dark (blue) and 360 nm photo stationary state (red).

**Table 14.** Half-lives of Slug Finger-2-71.

Temperature	Half-life (min) $\pm$ S.D.
15 °C	51 $\pm$ 4
20 °C	32.5 $\pm$ 0.3
40 °C	9 $\pm$ 1

Similar to Slug Finger 1-71, Slug Finger 2-71 demonstrates no significant absorbance band shifts, but in this instance the half-life of the light state of Slug Finger 2-71 shows no appreciable change to that of the *cis*-71 in water suggesting no helical stabilisation is derived from the more  $\alpha$ -helical conformation of the peptide.



**Figure 74.** UV/visible absorption spectra of the Slug Double Finger-71-71 in dark (blue) and 360 nm photo stationary state (red).

The extent of switching of Slug Double Finger-71-71, as judged by the reduction in 360 nm absorbance, is not as pronounced as that of the other cross-linked peptides. This can be partially explained by taking into account that the efficiency of 71 switching is limited; if each switch had an independent switching efficiency of 80% at best the total the percentage of the double switched would only be 64% .

**Table 15.** Relaxation times of Slug Double Finger-71,71.

Temperature	Half-life (min) $\pm$ S.D.
15 °C	39 $\pm$ 2
20 °C	227 $\pm$ 1
40 °C	6.4 $\pm$ 0.2

The relaxation of Slug Double Finger-71,71 is faster than that of both of the switches on the individual peptide sections. This indicates that the relaxation cannot be an independent event where each section relaxes with its own characteristic half-life. In fact, the relaxation curve is closer to fitting first order kinetics (Figure 75).

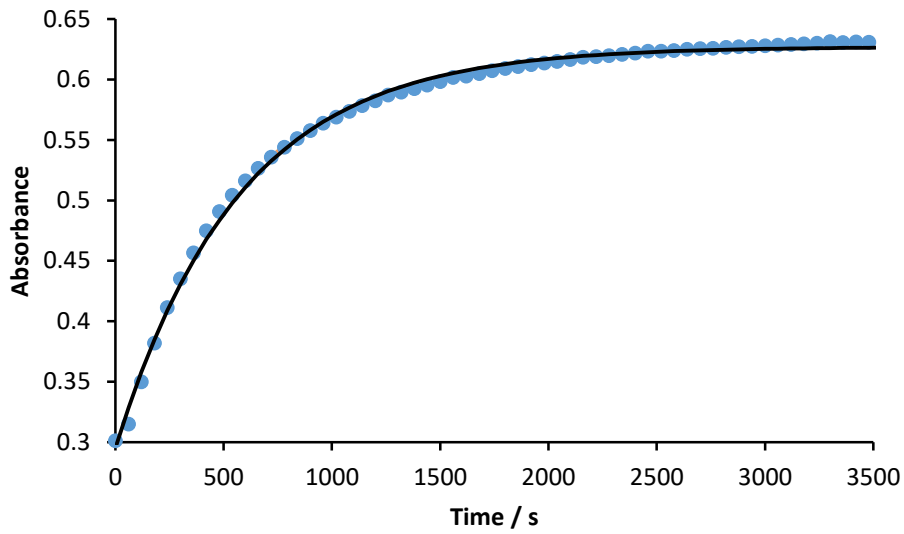


Figure 75. Relaxation curve Slug Double Finger-71,71 at 40 °C overlaid with a first order kinetics fit.

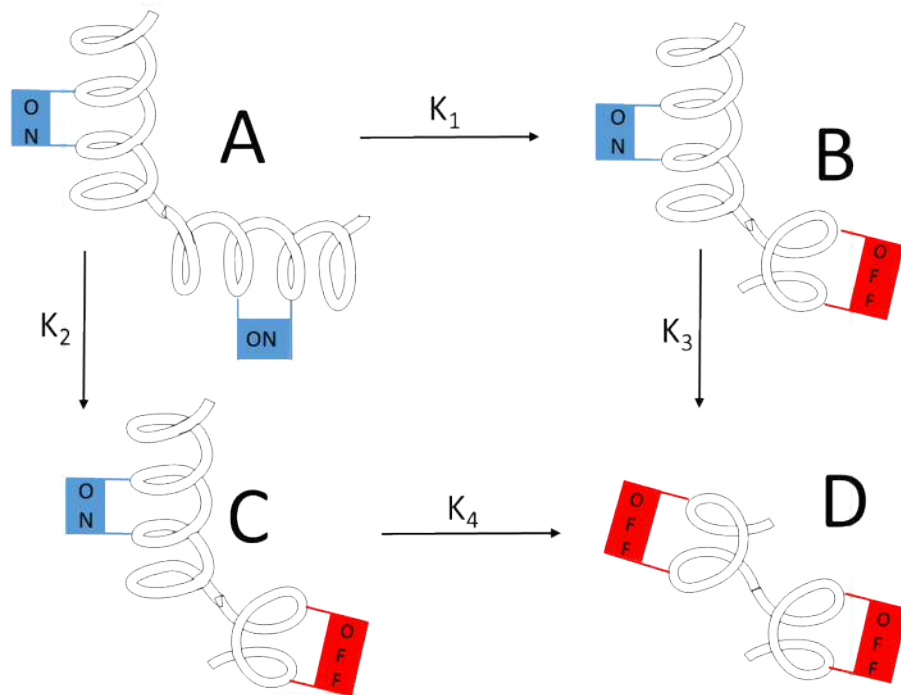


Figure 76. Schematic representation of independent Slug Double Finger-71,71 relaxation.



The closeness of the fit to a first order relaxation curve is not concordant with independently relaxing crosslinkers with 3-fold different half-lives which would result in a curve with two stages. Therefore, the relaxation of the one switch must directly affect the relaxation of the other.

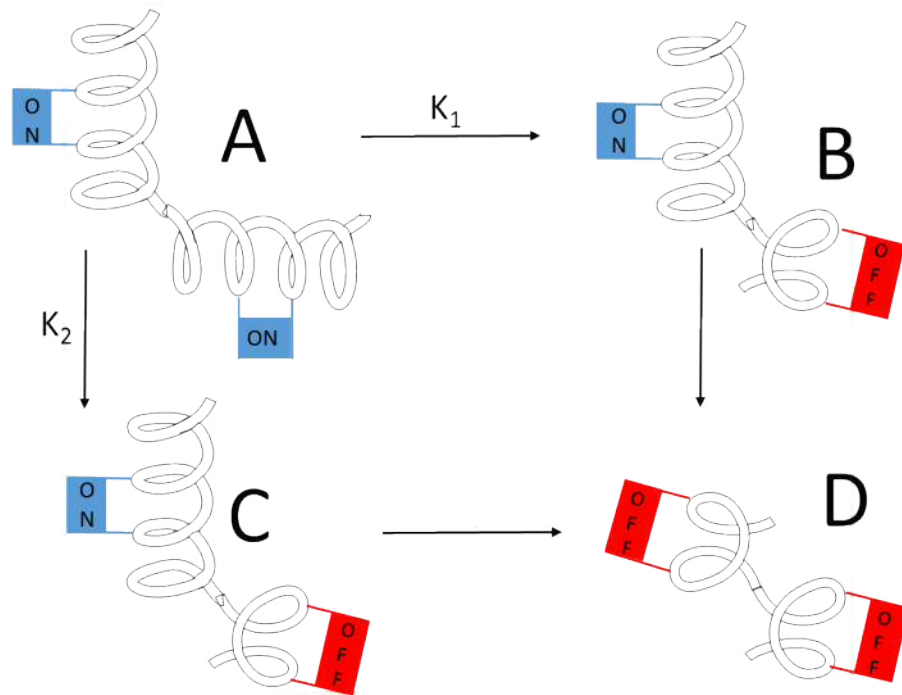
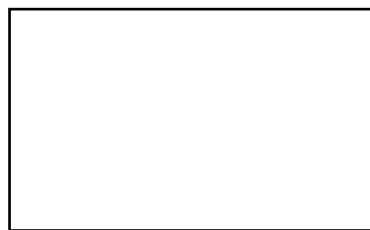


Figure 77. Schematic representation of non-independent Slug Double finger-71,71 relaxation.

If treated as a two body problem where both switches are in the *cis* in state A and states B and C are *cis*, *trans* and *trans*, *cis* states a combination of relaxations resolved to the equation:

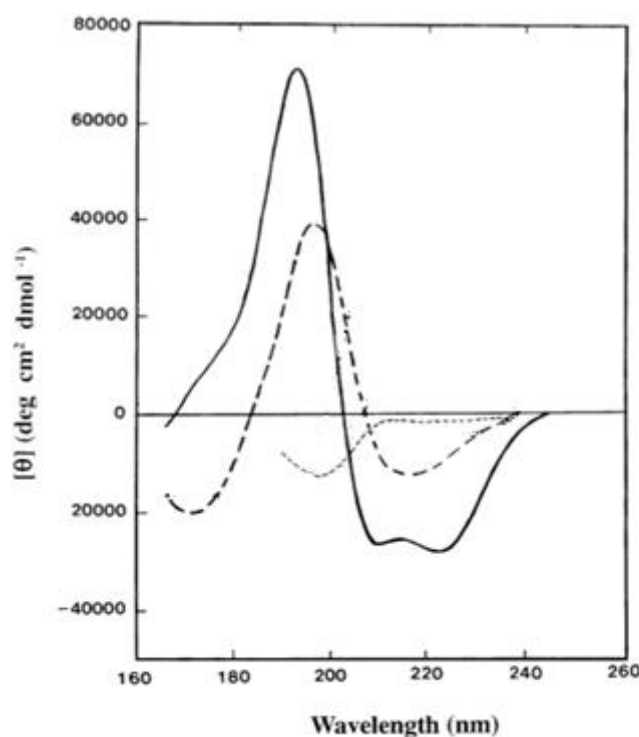


However, as there are no discernible features on the relaxation curve in order to measure any of these variables independently and no way in which to switch A without

switching B or *vice versa* it is very difficult if not impossible to determine the half-life for each individual switch.

## 2.5 Circular dichroism

Circular dichroism was used in order to investigate the effect of the photoswitch on the conformation of the cross linked peptide. Circular dichroism measures the difference in extinction coefficients of clockwise and anticlockwise anti clockwise circularly polarised light. Because the peptide is inherently chiral due to the chirality of its constituent amino acids as well as the secondary structure of the peptide being chiral, right and left-handed circularly polarised light interact differently with the molecule. In such cases samples are said to have ellipticity polarisation.<sup>105</sup>



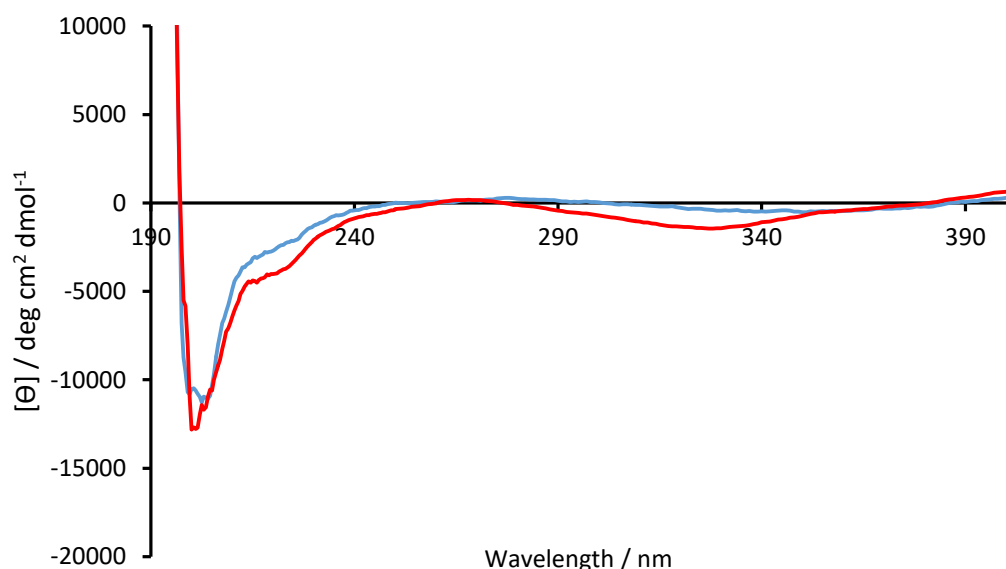
**Figure 78.** CD spectra of different protein secondary structure motifs:  $\alpha$ -helix (solid line), anti-parallel  $\beta$ -sheet (long dashed line), unstructured peptide (dotted line). Modified from reference <sup>105</sup>.

As seen in Figure 78, random coil can be distinguished from  $\alpha$ -helix by the shape of the spectrum and is usually estimated by measuring ellipticity at 208 and 222 nm.<sup>105</sup> Circular dichroism is concentration, temperature and solvent dependent. Low salt phosphate buffer was used for these experiments because phosphate does not show strong absorbance in the region of interest whilst sodium chloride absorbs strongly in the 180 nm region which can lead to saturation of the detector. Circular dichroism values are quoted in mean residue ellipticity based on the following formula:

- $\theta_{\text{MRE}}$  = mean residue ellipticity
- $n$  = number of residues
- $c$  = concentration
- $l$  = path length
- $\theta$  = ellipticity

The concentrations of azobenzene crosslinked peptides were calculated from the absorbance of the azobenzene at 370 nm using an extinction coefficient of  $29,000 \text{ cm}^{-1} \text{ mol}^{-1}$ .<sup>50</sup> The concentrations of unconstrained Slug Double Finger and Slug Finger 1 were calculated from the absorbance of tryptophans and tyrosine sidechains at 280 nm using extinction coefficients of  $5500 \text{ cm}^{-1} \text{ mol}^{-1}$  and  $1490 \text{ cm}^{-1} \text{ mol}^{-1}$  respectively. This was not possible for Slug Finger 2 due to the lack of absorbing residues so the concentration was estimated by comparing absorbance at 210 with Slug Finger 1 and scaling by the number of residues present in the respective peptides.

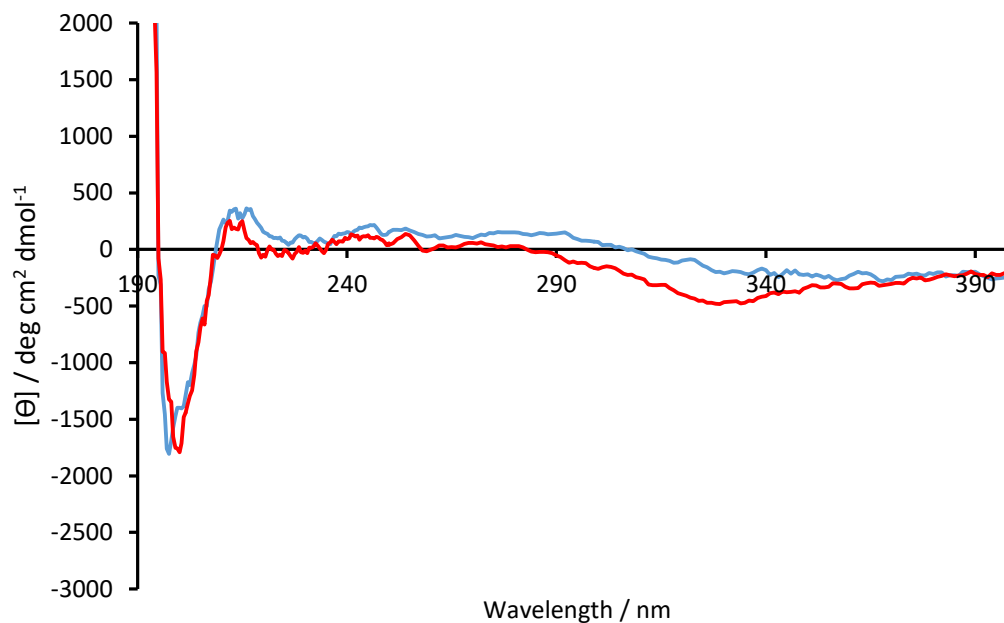
These concentrations were then used to calculate the mean residue ellipticity for each of the samples.



**Figure 79.** CD spectra of Slug Finger 1-71 the dark state (blue) and the 360 nm photo stationary state (red).

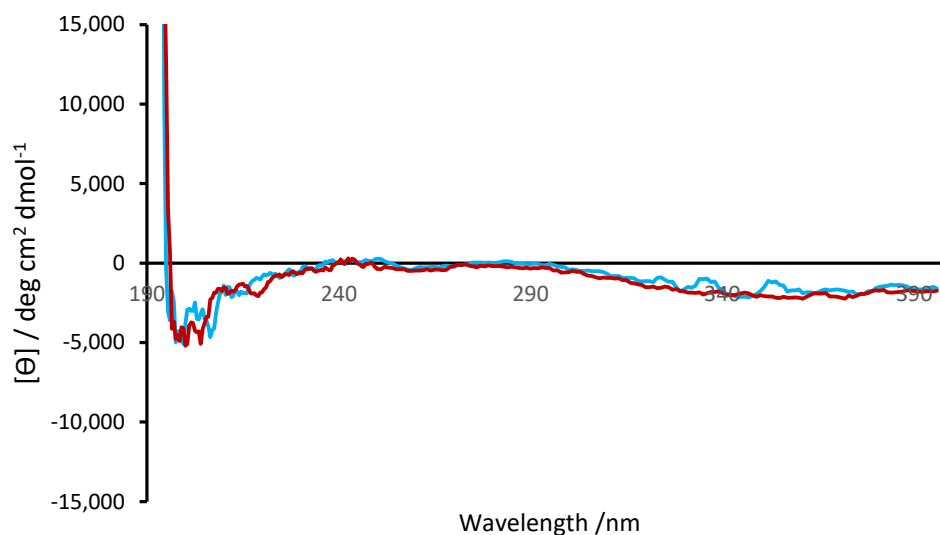
CD spectra of Slug Finger 1-71 show some change in  $\alpha$ -helical character and also some change in the 330 nm region due to the azobenzene being in a chiral environment.





**Figure 80.** Circular dichroism spectra of Slug Finger 2-71 in the dark state (blue) and the 360 nm photo stationary state (red).

The atypical finger Slug Finger 2-71 shows a slight change in structure but it is not as clear as that of Slug Finger 1-71.



**Figure 81.** Circular dichroism Slug Double Finger-71,71 in the dark state (blue) and the 360 nm photo stationary (red)

Slug Double Finger-71,71 shows no observable change between light and dark states. This could be due to the reduced half-life of the switches and the lower effective switching of the peptide due to the two azobenzenes present.

## 2.6 Fluorescence anisotropy

Fluorescence anisotropy was used to determine the binding affinities for cross-linked Slug Finger peptides to target DNA in both light and dark states. Fluorescence anisotropy can be used to detect the difference in rotational rates of a chromophore when it is free in solution or bound to another molecule. Fluorescein was used as the fluorophore for these experiments.

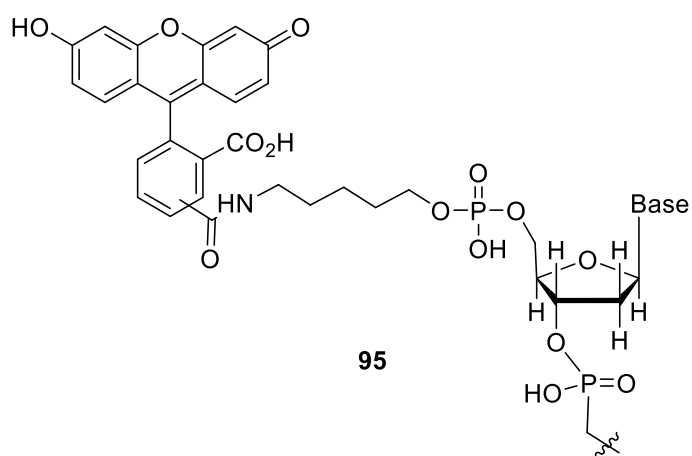
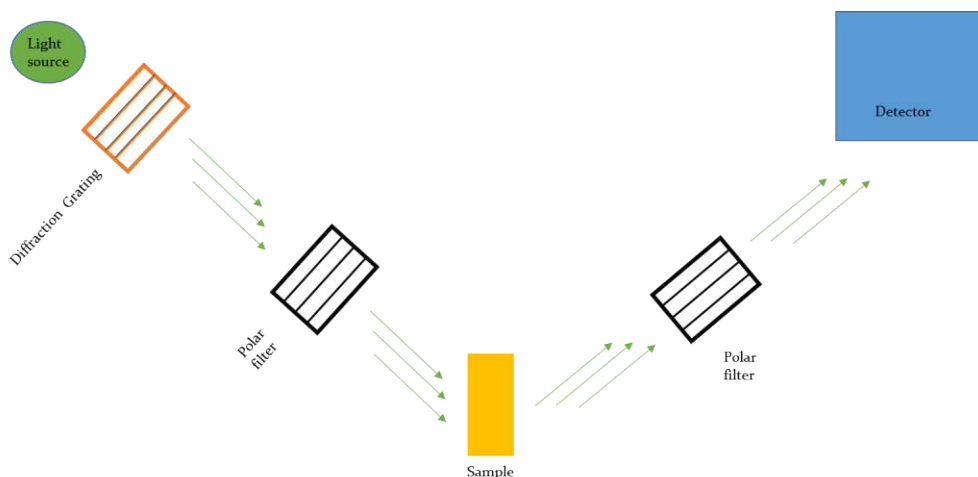


Figure 82. Fluorescein attached to DNA.

Fluorescein was excited with plane polarised 495 nm light causing excitation of an electron, after a short delay this electron returns to the ground state releasing a slightly less energetic photon. If the decay of the excited state is slower than the rate of molecular tumbling in solution then the polarisation will be scrambled, with the extent of scrambling being proportional to the tumbling rate. The rate of rotation in the solution can be used to measure relative size of the molecule and therefore the binding of one molecule to another to create a larger, slower tumbling complex can be measured.



**Figure 83** Diagram of fluorescent anisotropy apparatus optical set up.

One strand of an annealed DNA duplex containing the sequence of the Slug binding site was labelled with fluorescein (3' FAM-ATCCTGTCAGGTAGTTCTGC 5') and irradiated at 495 nm through a polarising filter. Emission at 520 nm was recorded using matched and crossed polarising filters to calculate the fluorescence anisotropy:

Anisotropy was calculated using the following formula.

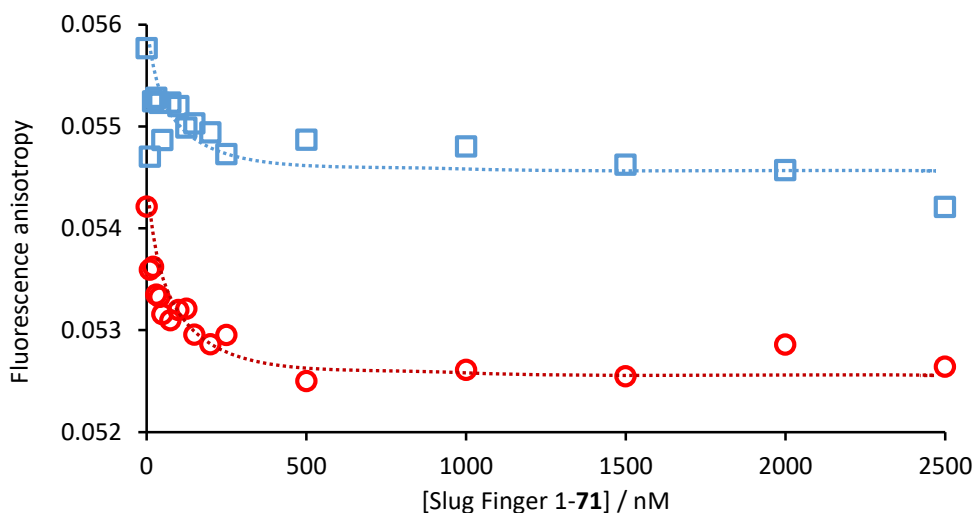
Where

- GF is the grating factor. —
- $I_{vh}$  is the intensity with the gratings vertical and horizontal.
- $I_{vv}$  is the intensity with both gratings vertical.

The labelled DNA tumbles at a characteristic rate which is altered when Slug Finger peptides bind to the DNA so that titrations to saturation can be used to calculate association constants. In order to calculate the  $K_d$  the concentration of the DNA was determined by the absorbance of fluorescein:

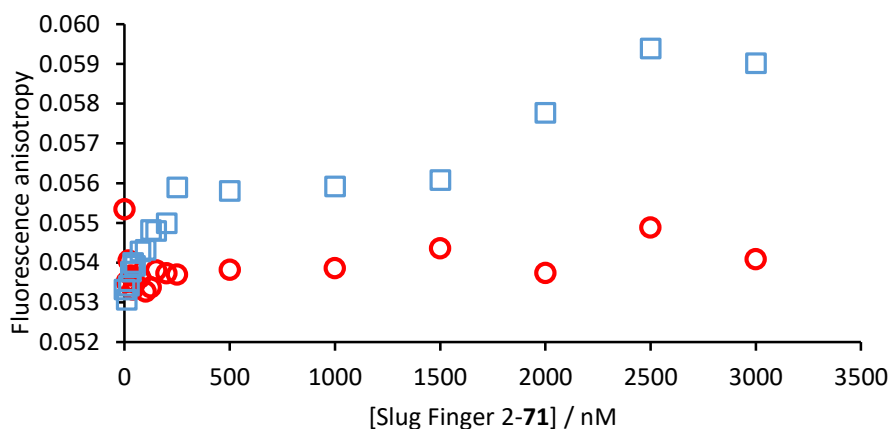
- $A$  = Absorbance = 0.0774405 @ 495 nm
- $E$  = extinction coefficient for fluorescein = 83,000  $M^{-1} cm^{-1}$
- $l$  = path length of cell = 0.5 cm
- $c$  = concentration

The graphs below show the fluorescence anisotropy of binding of the peptide to labelled tagged DNA (5 nm) at varying concentration of Slug Finger 1-71.

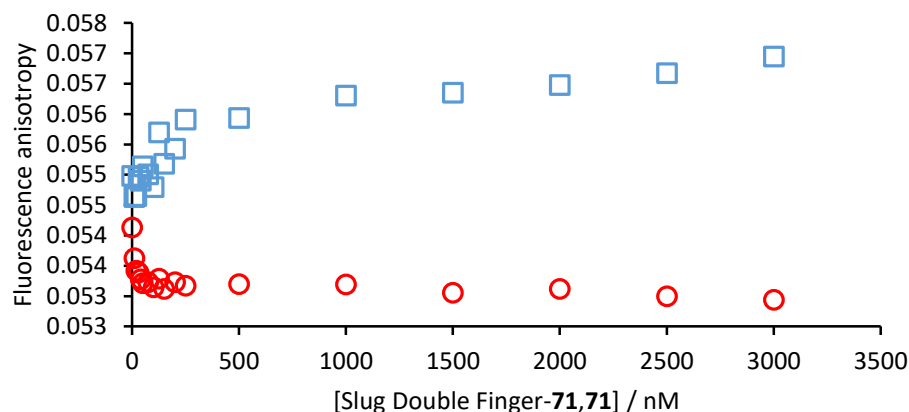


**Figure 84.** Fluorescence anisotropy binding data for Slug Finger 1-71 in both light (blue squares) and dark adapted (red squares) states. Curves represent the best fit for the data see section 6.17.

Irradiated Slug Finger 1-71 bound to fluorescently labelled DNA with a  $K_d$  of approximately 100 nM, whereas in the dark adapted state it less clear interaction. Data for Slug Finger 2-71 and Slug Double Finger-71,71 were also not convincing, with the dark adapted state of Slug Finger 2-71 showing very similar behaviour to that of the irradiated peptide. An apparent inversion in polarity was observed for Slug Double Finger-71,71 (Figure 86), possibly due to distortion of the B-form DNA structure.



**Figure 85.** Fluorescence anisotropy binding data for Slug Finger 2-71 in both light (blue squares) and dark adapted (red squares) states.



**Figure 86.** Fluorescence anisotropy binding data for Slug Double Finger-71,71 in light (blue squares) and dark adapted (red squares) states.

## 2.7 Conclusions

The binding of the Slug Finger 1-71 crosslinked peptide appears to be affected by the change of conformation brought about by isomerisation of the photoswitch although the binding data is not definitive proof due to the error in the data. Slug Finger 2-71 does not show the same extent of change, but this could be due to the Slug Finger 2 helix not being responsible for strong binding to DNA, but merely providing additional recognition for the DNA sequence. Neither Slug Finger 2-71, nor Double Slug Finger-71,71 show the same clear change in CD spectrum upon isomerisation as Slug Finger 1-71.

To further investigate the binding of Slug Finger crosslinked peptides to DNA several other techniques were investigated to attempt to more clearly discern if isomerisation had an effect, as the changes in the anisotropy brought about by the binding of the slug peptides to the DNA were nearing the limitation of the anisotropy values obtainable by the machine. Several electronic mobility shift assays (EMSA) were carried out but these yielded very similar results. Isothermal titration calorimetry (ITC) was also attempted, but required a concentration of peptide greater than was achievable in the titration buffer for the size of the ITC cell.

This project might not have worked due to our minimalist design omitting additional DNA-binding residues found in the loop region of the ZF-TF that have been shown to make important water mediated interactions with DNA<sup>94</sup>. Furthermore, *cis*-71 has too short a half-life to be ideal for this application and the wavelength the switch requires

to change conformation is damaging to cells. In order to create a truly biologically relevant tool switches must have greater half-lives and require lower energy photons for switching between states.

# Chapter 3:

## Visible Light Photoswitches

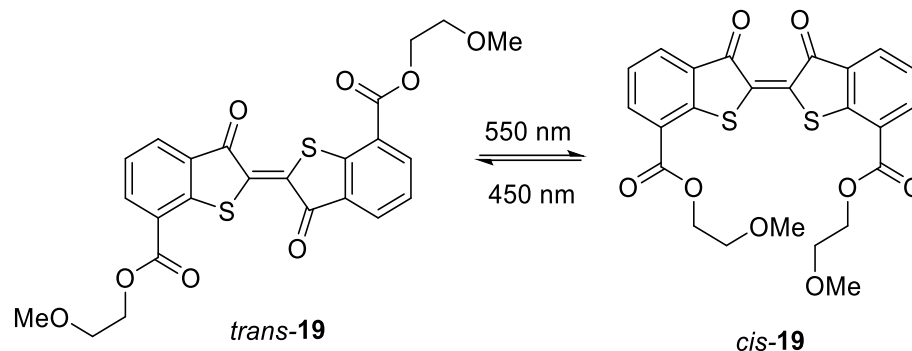
### 3 Visible light photoswitches

#### 3.1 Introduction

An ideal photoswitch would offer bidirectional switching so that both ‘on’ and ‘off’ states are readily optically accessible. The actuation wavelengths should fall into the therapeutic window where light passes harmlessly through the surrounding tissue and not be absorbed by compounds such as flavins, melanin or haem. For most applications the excited state should possess a half-life in the region of hours to days, or ideally be stable until the application of a stimulus to reverse. The ideal photoswitch would also be soluble in water. Current photoswitches are variously limited by their half-lives, and their rate and extent of switching with very few possessing combinations of properties in the range that is viable for *in vivo* experiments. The most pressing problem, however, is the wavelength required for switching, with many photoswitches requiring harmful UV irradiation.

#### 3.2 Thioindigo

One example of a photoswitch with improved optical and physical properties compared to **71** are thioindigos; structurally related to the indigo family as described in the introduction (section 1.3.4). Thioindigo switches from the *trans* state to the *cis* state upon irradiation with 550 nm light and can be converted back to the *trans* state with 450 nm light or by heating.



**Figure 87.** Thioindigo modified to act as a isomerisation dependant ligand for small cations.<sup>106</sup>

These properties were demonstrated by the adaptation of thioindigo for photo regulated ion binding (Figure 87). In the *trans* state the two methoxy ether chains on the 7 and 7' positions have lower affinity for alkaline metals than in the *cis* state. This property was used to transport metals from aqueous solutions then selectively release them upon irradiation with 450 nm light.<sup>106</sup>



The main barrier to the use of thioindigo as a biological photoswitch is that thioindigo is an entirely planar molecule, resulting in extremely poor solubility in solvents that cannot disrupt  $\pi$ -stacking between thioindigo molecules.<sup>33</sup> Thioindigo is useful for vat dyeing clothing precisely because of its poor solubility, leaving the dye impregnated in the fabric retaining the colour. The main advantages of thioindigo compared to azobenzenes and spiropyrans are that the *cis* state is more kinetically stable, so that once switched thermal reversion to the *trans* state is very slow but can be stimulated using shorter wavelength light. Thioindigo dyes are typically made by the same means as indogoid dyes (Figure 90), by oxidation of a thioindoxyl precursor that oxidises to thioindigo in air.

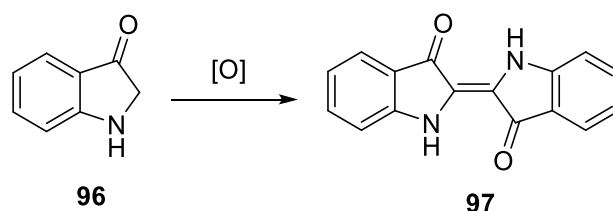


Figure 88. Oxidation of indoxyl to indigo.

Electron withdrawing substituents in the 5- or 7-positions are known to red-shift the absorption spectra,<sup>107</sup> once reduced to amines then acylated to amide groups they could be used as a handle to attach these molecules to proteins.

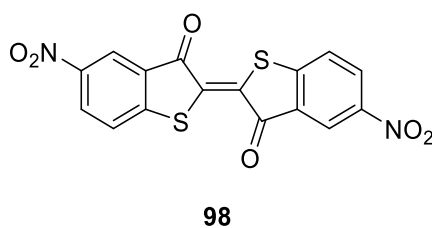
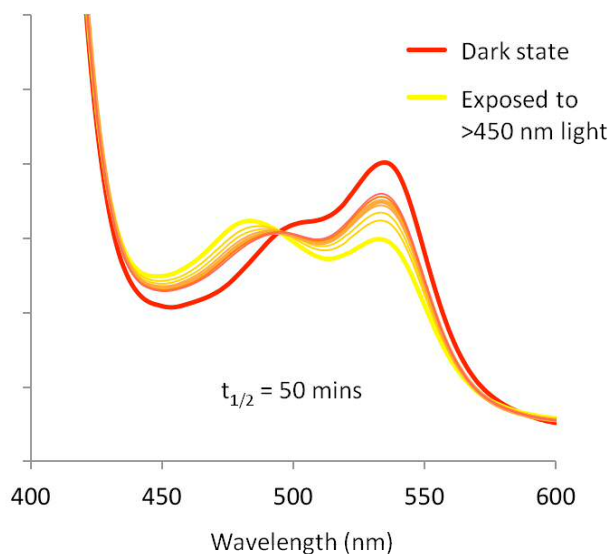
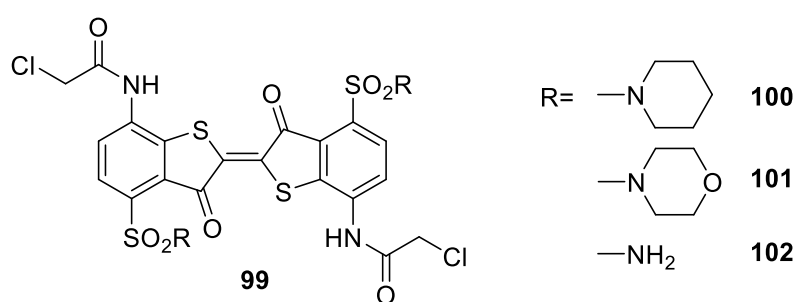


Figure 89. 5,5'-Dinitro thioindigo (98).



**Figure 90.** 5,5'-Dinitrothioindigo (**98**) UV/visible spectra.

5,5'-Dinitrothioindigo (**98**) shows rapid switching at blue/green wavelengths and an extended lifetime of its light state compared to BSBCA (**71**). Unfortunately, preliminary work also found **98** to be extremely insoluble, and when a 7-(chloroacetamido)-thioindoxyl was attached to two cysteines incorporated in a Bak peptide and a thioindigo formed on the peptide the product was insoluble.<sup>36</sup> Therefore, following the example of BSBCA, a synthesis was proposed to add sulfonyl groups to thioindigo to improve solubility (Figure 91).<sup>4</sup>



**Figure 91.** Idealised thioindigo photoswitch.

However, as sulfonate groups may interfere with the subsequent thioxazole formation, this route incorporated sulfonamides formed either with highly polar or non-planar amines (morpholine and piperidine) to help overcome the self  $\pi$ -stacking of the thioindigo molecules.

## 3.2.1 Retrosynthesis of thioindigo derivatives

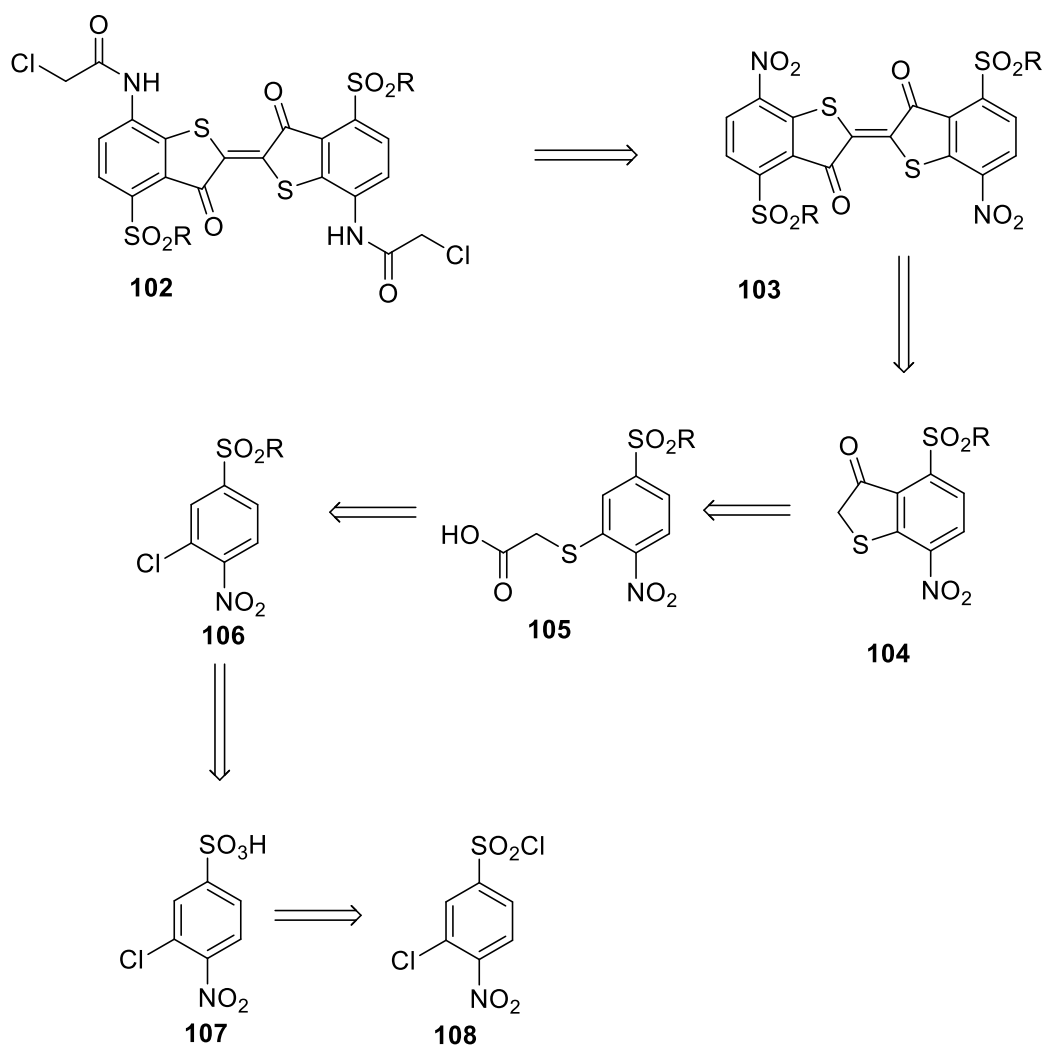


Figure 92. Retrosynthesis of sulfonamidothioindigo.

The first retrosynthetic step is to break the chloroacetamide bond followed by oxidation of the amine to a nitro group to protecting it from undergoing reactions in proceeding steps. Separating the two symmetrical halves of the thioindigo to form thioindoxyl compound **104**. Removal of the thioglycolate is then followed by the scission of the sulfonamide groups to give commercially available 3-chloro-4-nitrobenzenesulfonic acid.

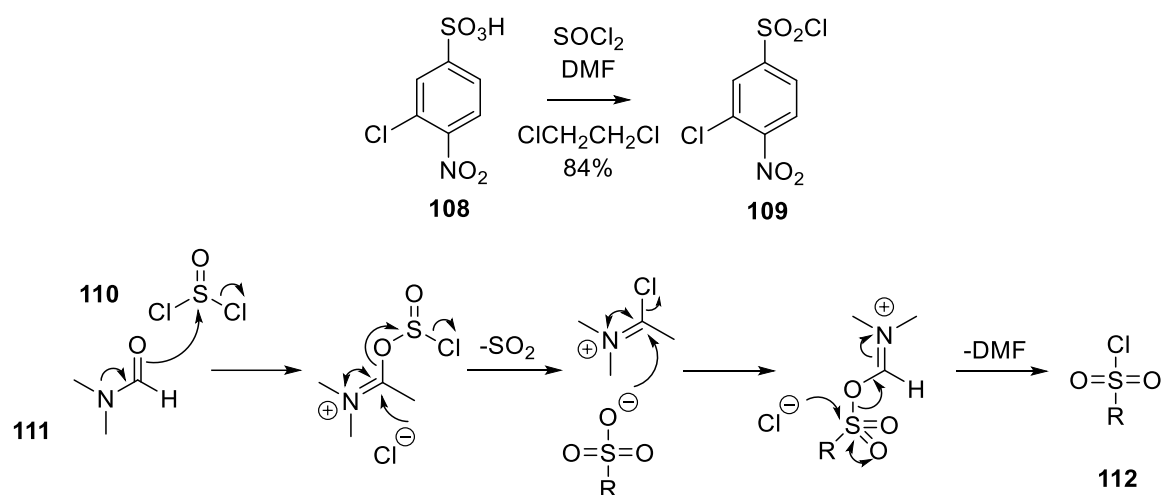


Figure 93. Chlorination of sulfate group.

The first step was chlorination of 3-chloro-4-nitrobenzenesulfonic acid (**108**) to 3-chloro-4-nitrobenzenesulfonyl chloride (**109**). This was achieved using a variation of the Vilsmeier–Haack reaction described in a patent EP1367 058 A1. Dimethylsulfoxide (DMF, **111**) is chlorinated to form an activated species, which reacts with oxygen on the sulfate group which is then displaced by a chloride ion yielding the corresponding chlorosulfonate. This can then be converted to the corresponding sulphonamide with a non-nucleophilic base and any primary or secondary amine.

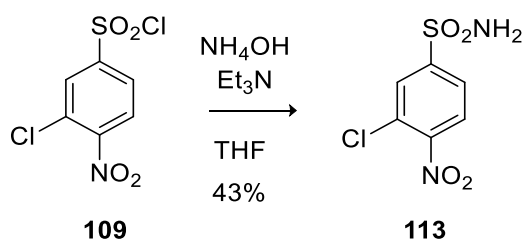
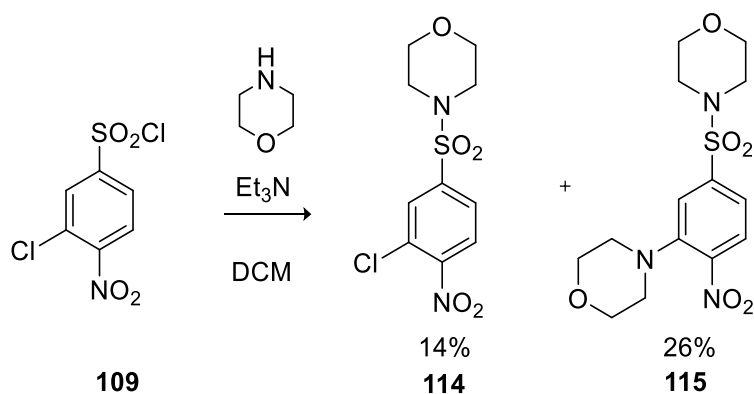


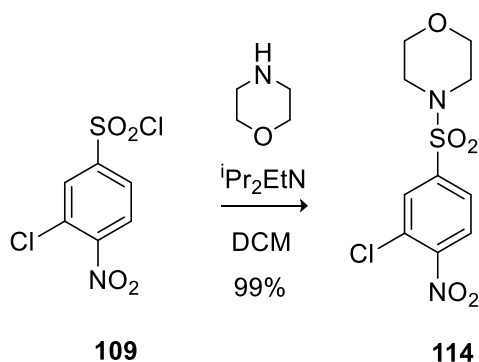
Figure 94. Sulphonamide formation.

Unexpectedly, using morpholine resulted in nucleophilic aromatic substitution and gave a mixture of compounds **114** and **115**. Using substoichiometric morpholine (0.9 equivalents) failed to improve the selectivity.



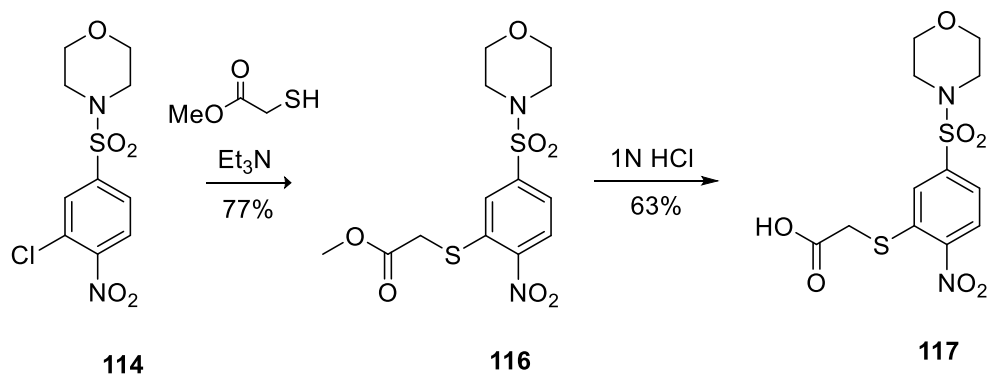
**Figure 95.** Formation of sulphonamide and an unexpected byproduct from morpholine.

However, using diisopropylethylamine as the base rather than triethylamine or bulk morpholine resulted in selective sulphonamide formation in a 99% yield. This was unexpected, because the base was expected simply to react with the hydrochloric acid formed during the reaction rather than play an active part in either reaction.



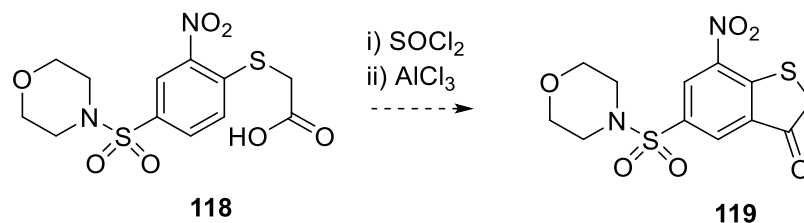
**Figure 96.** Selective sulphonamide formation with morpholine.

This compound exhibits solvatochromism; upon concentration the solution changes colour from an off yellow solution to a dark green oil once the solvent is removed due to the  $\pi$ -stacking of the system. The next step was an aromatic substitution with methylthioglycolate where a thiolate conducts a nucleophilic attack on the aromatic ring. The hydrochloric acid by-product was initially removed by the addition of potassium hydroxide, but this caused hydrolysis of the sulphonamide to the sulfate. Using a milder reagent such as triethylamine retained the sulphonamide group.



**Figure 97.** Nucleophilic aromatic substitution with methylthioglycolate.

Following deprotection of the methyl ester with dilute hydrochloric acid (1 M) to yield **117**, closure of the five-membered ring was attempted. Unfortunately, this reaction was unsuccessful due to the electron deficient ring preventing efficient cyclisation.



**Figure 98.** Thioindoxyl ring formation using a Friedel-Crafts reaction.

**Table 16.** Conditions used to attempt to close the thiooxazol ring.

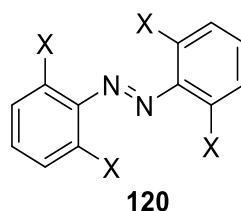
Reagent	Conditions	Reference
Chlorosulfonic acid, bromine (0.3 eqv)	2 hours, 20 °C	37
Chlorosulfonic acid, bromine (0.3 eqv)	overnight, 20 °C	37
Thionyl chloride	overnight reflux	
Thionyl chloride, then aluminium chloride, 1,2-dichloroethanol	overnight, 20 °C	38
Thionyl chloride, then aluminium chloride, 1,2-dichloro benzene	overnight reflux	38

The inability to detect product formation under any of the attempted conditions (Table 16) strongly suggests that the presence of a sulfonamide group at the 5 position of the thioindigo simply makes the ring too electron deficient for the subsequent ring closure

by radical reaction or nucleophilic attack on the acyl chloride from the ring. It is not appropriate to reduce the nitro group to the amine at this point because it would act as a competing nucleophile. Convenient amino protecting groups to mask an amine at this stage are mainly acid labile and would be unlikely to survive the chlorination/cyclisation reactions intact. Therefore, synthesis was halted at this point and attention focussed on more tractable structures.

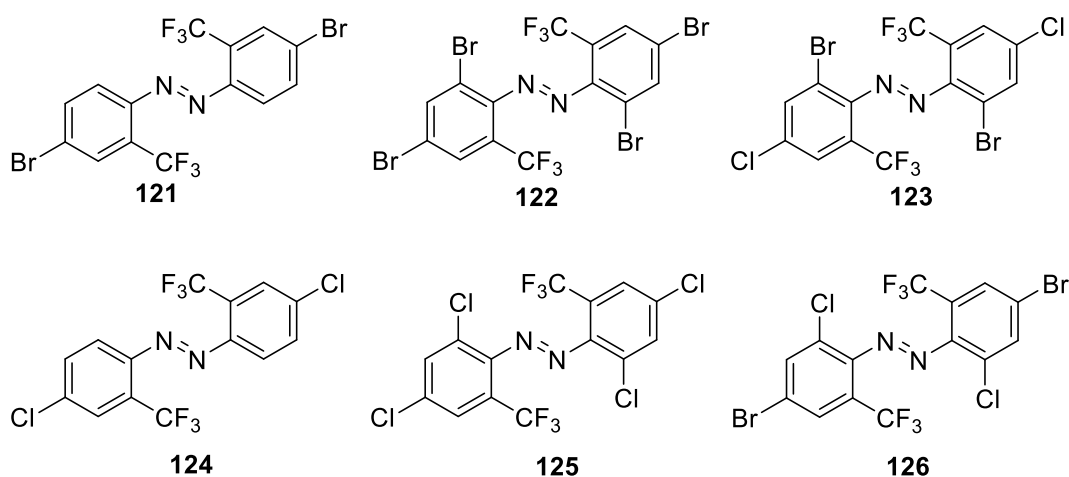
### 3.3 Trifluoromethyl derived photo switches

Contemporary reports by Hecht<sup>60</sup> and Woolley<sup>62</sup> highlighted the use of *ortho* halogen substituents to increase the half-lives of azobenzene derived photo switches.



**Figure 99.** Tetra-*ortho*-substituted azobenzene.

Substitution of the *ortho* positions of the azobenzene not only resulted in increased half-lives of *cis* isomers but also effected the UV absorbance of the spectra, causing separation of the  $n-\pi^*$  bands of the *cis* and *trans* states. This band separation allows the use of lower energy light to switch the states, albeit at lower efficiency as the extinction coefficients of the  $n-\pi^*$  bands are much lower than those of the  $\pi-\pi^*$  band. A range of trifluoromethyl azobenzenes were synthesised to investigate if these effects were due to steric factors or the electronegativity of the halogen substituents, (Figure 100).

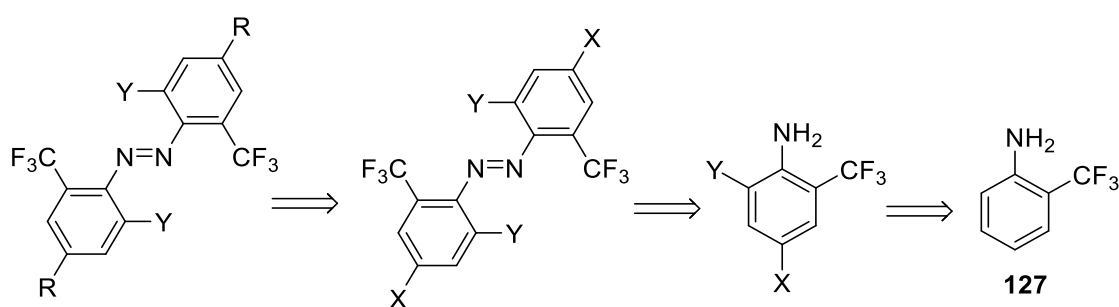


**Figure 100.** Trifluoromethyl-containing azobenzene photoswitches.

These candidates were chosen due to the ready synthetic accessibility of sufficient molecules to compare electrostatic and steric effects on half-life of the photoswitch and the band separation between absorbance spectra in the *cis* and *trans* states. Additionally, the *para* halogen substituents should be useful for introducing linkers to attach these improved azobenzenes to peptides for photocontrol.

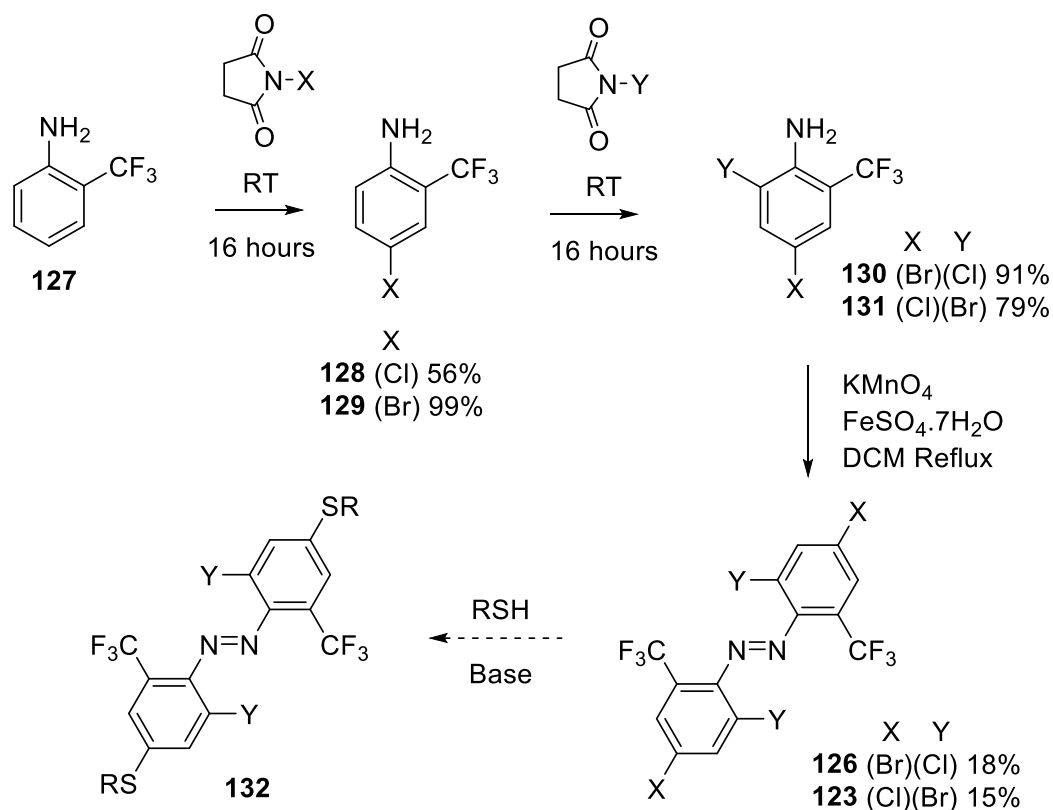


## 3.3.1 Retrosynthetic analysis and synthesis



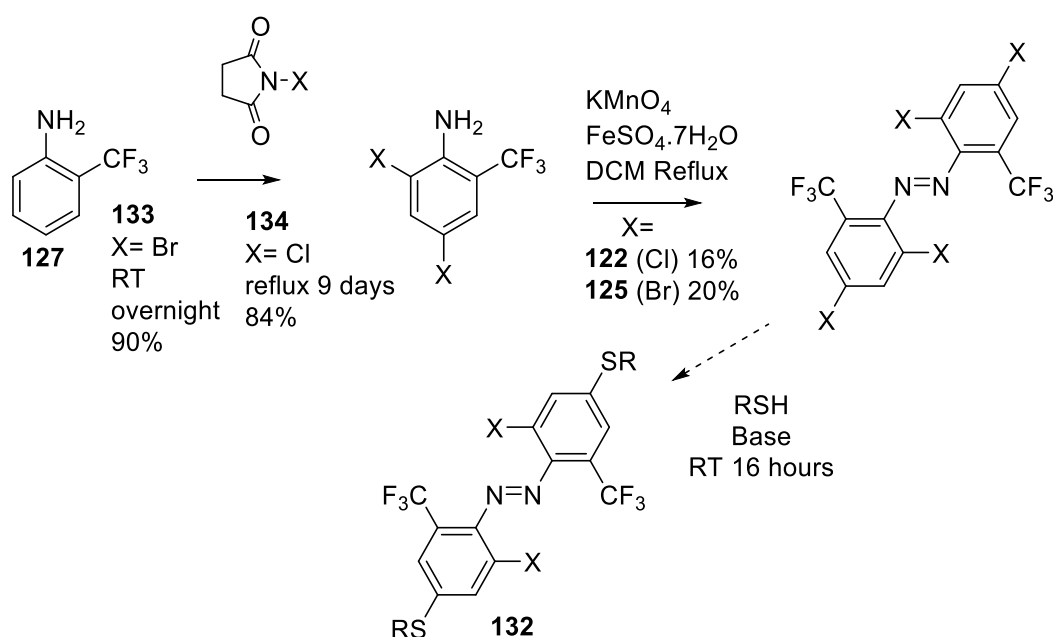
**Figure 101.** Retrosynthetic analysis of *ortho*-trifluoromethyl azobenzenes.

Removal of these proposed connecting groups (denoted by R in Figure 101) leaves a multiply halogenated azobenzene, which is then broken at the symmetrical azo bond. Appropriately substituted rings are prepared by the selective halogenation of the *meta* and *para* positions. This can be achieved by first halogenating the *para* position with stoichiometric quantities of either chloro- or bromosuccinimide. Additional halosuccinimide is used to introduce the *ortho* substituents (Figure 102).



**Figure 102.** Synthetic route to *ortho*-trifluoromethyl azobenzenes.

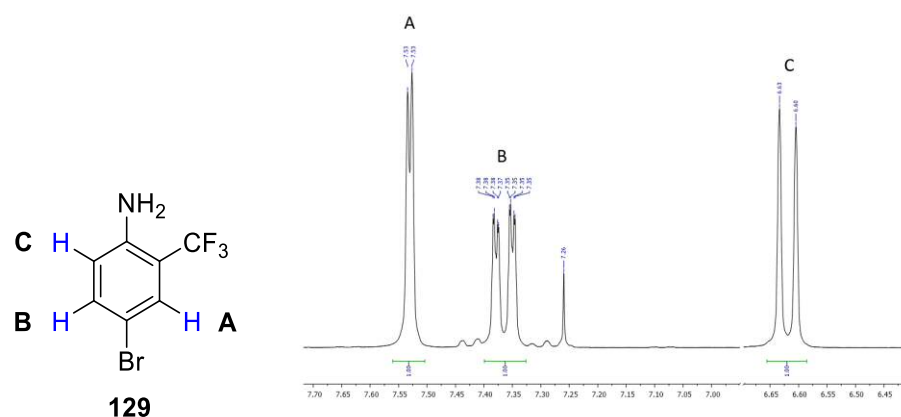
Selective halogenation of the *para* position was achieved using 1 equivalent of *N*-bromosuccinimide or 0.9 equivalents *N*-chlorosuccinimide.<sup>61</sup> The chloride reagent was less selective than the bromide reagent, leading to a reduced yield due to unwanted *ortho*-substitution even when using substoichiometric reagent (0.9 equivalents). Reaction with *N*-iodosuccinimide was attempted, but was not successful. Introducing a second halogen to the *ortho* position on the already electron deficient ring system required heating and extended reaction times compared to the introduction of the *para* halogen.



**Figure 103.** Synthetic route to *ortho*-trifluoromethyl azobenzenes.

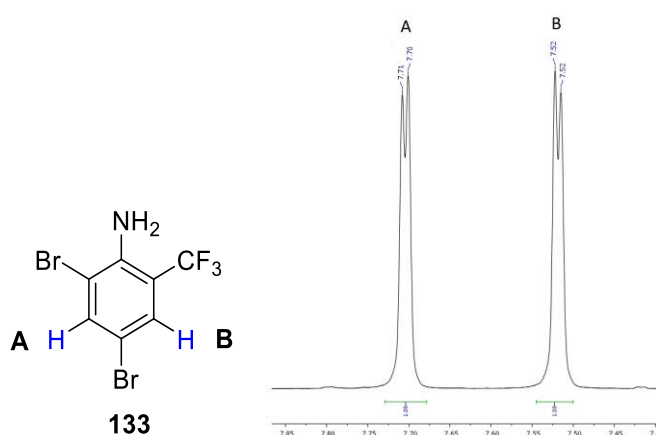
Whereas the reaction to form 2,4-dibromo intermediate **133** was complete after stirring overnight at room temperature, 2,4-dichloro intermediate **134** required 9 days refluxing in chloroform with 2.2 equivalents of *N*-chlorosuccinimide. This is a result of the reduced reactivity of the mono-chlorinated aromatic ring; chlorine is the most electron withdrawing halogen on an aromatic ring due to its electro negativity and reduced electron back donation from its p-orbitals in comparison to the more electronegative fluorine.

The correct substitution patterns were confirmed by the J-coupling values in NMR spectra. For example, selective bromination of the *para*-position gives a characteristic double doublet for position B on compound **129** (Figure 104).



**Figure 104.**  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ) spectrum of 4-bromo-2-trifluoromethylaniline (**129**).

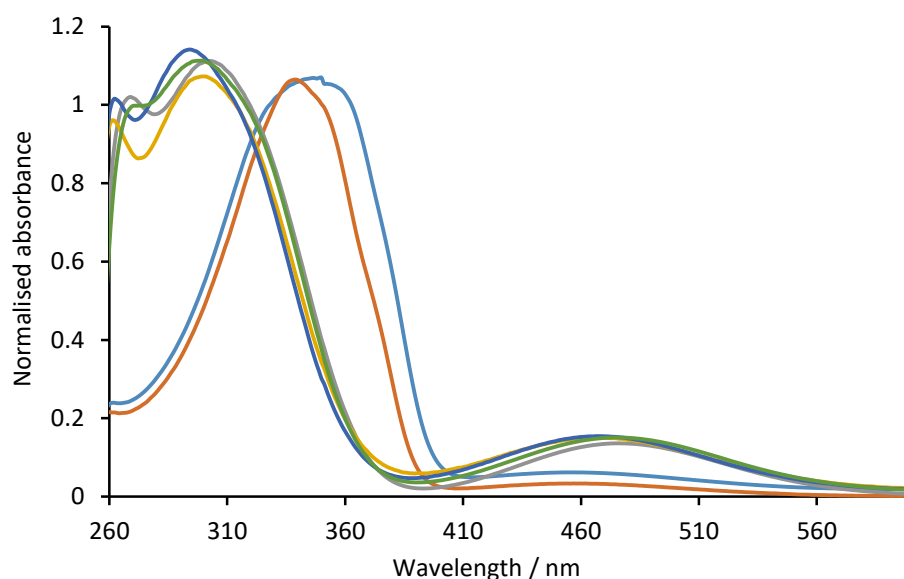
Proton B is split by proton C with  $J_{\text{B-C}} = 8.5$  Hz, a magnitude appropriate for an *ortho* coupling, and proton A with  $J_{\text{A-B}} = 2.0$  Hz as expected for a *meta* relationship. Intermediate **129** can subsequently be brominated again to yield **133** whose NMR shows the expected two peaks between 7 and 8 ppm with a coupling constant of 2.3 Hertz, consistent with *meta* substitution (Figure 105).



**Figure 105.**  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ) spectrum for 1,3-dibromo-6-trifluoromethylaniline from 7.4 -7.8 ppm (**133**).

The final step for the synthesis of **132** was unsuccessful but did lead to some interesting properties.

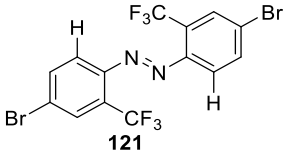
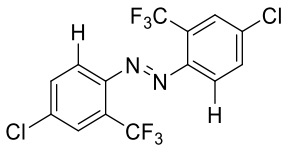
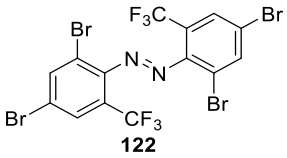
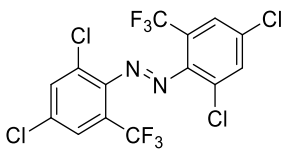
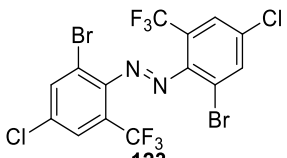
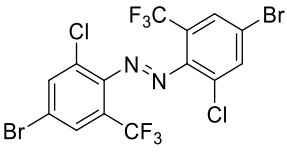
## 3.3.2 Spectra of halotrifluoromethyl azobenzene photoswitches



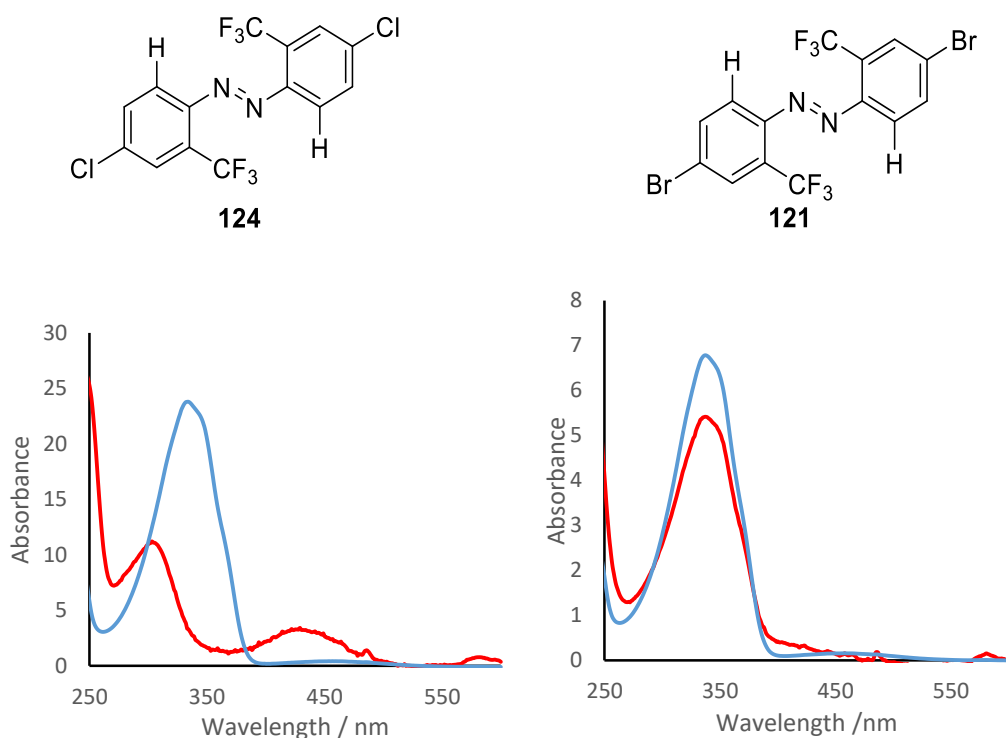
**Figure 106.** UV/visible spectra of halogenated trifluoromethyl azobenzenes. Compound **121** (blue), Compound **124** (red), Compound **122** (orange), Compound **125** (green), Compound **126** (green) and Compound **123** (light blue),

The most apparent change in the UV absorbance spectra of halogenated trifluoromethyl azobenzenes (**121-124**) is a general blue-shift in the absorbance of the  $\pi$ - $\pi^*$  orbital from 360 nm for the di-*ortho*-trifluoromethyl substituted azobenzenes to ~300 nm for the tetra-*ortho* substituted azobenzenes. Photoswitching properties of these compounds were examined by recording spectra before and after illumination. Where an appropriate filter set was available to selectively irradiate the  $n$ - $\pi^*$  transition, it was used; otherwise the unfiltered output of a 250 W mercury vapour discharge bulb was used. The conversions for the di-*ortho* substituted azobenzenes were calculated based on the 360 nm absorbance (assuming that 360 nm is zero for *cis*). With the tetra-*ortho* substituted azobenzene this was not possible due to both the *cis* and the *trans* states absorbing 360 nm light.

**Table 17.** Absorbance maxima, switching extents and thermal reversion half-lives for **121-126**.

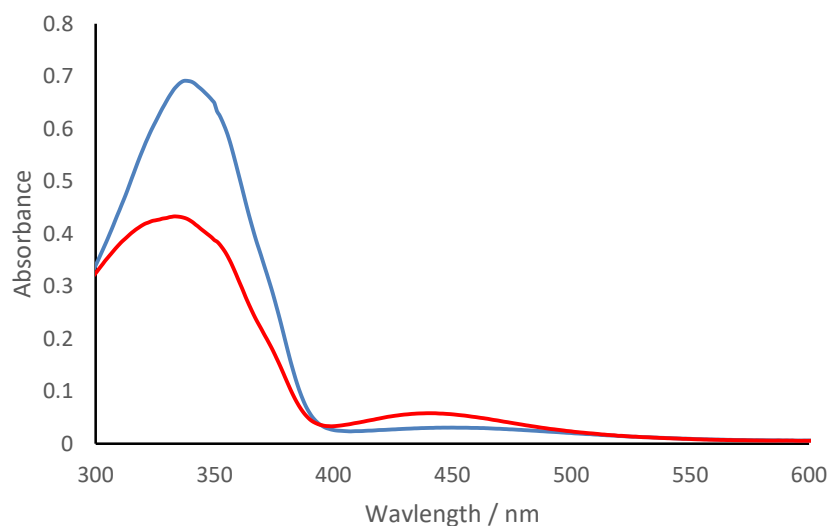
Switch	$\pi-\pi^*$	$n-\pi^*$	% <i>cis</i> at PSS	Filter	Half-life (min)	
	(nm)	(nm)			DMSO	20 °C
 <b>121</b>	350	450	10%	> 450 nm	180 ± 3	
 <b>124</b>	340	440	50%	> 450 nm	150 ± 14	
 <b>122</b>	300	480		none	~17 <sup>a</sup>	
 <b>125</b>	300	470		none	~69 <sup>a</sup>	
 <b>123</b>	300	460		none	~66 <sup>a</sup>	
 <b>126</b>	290	460		none	~140 <sup>a</sup>	

<sup>a</sup> Complete conversion of the *trans* dark state to that of the *cis* light state was not possible due to both the *cis* and *trans* absorbing the wavelength of irradiation. Low conversions gave potentially spurious *cis* isomer half-lives.



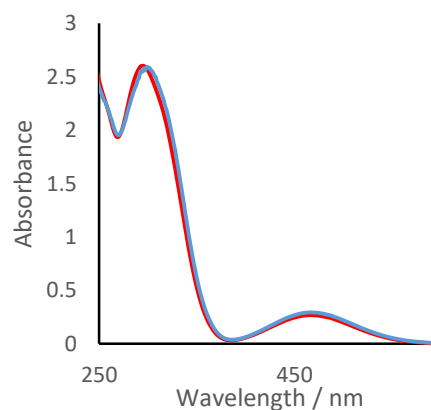
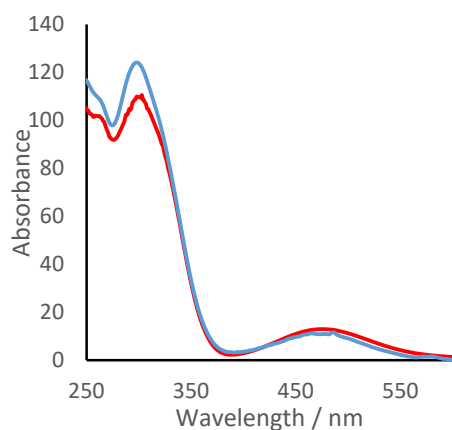
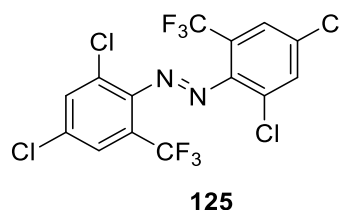
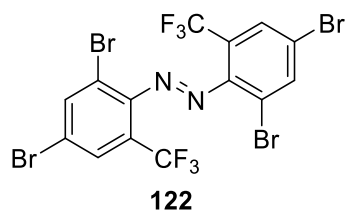
**Figure 107.** UV/visible absorbance spectra of **124** (left) and **121** (right) in their *trans* (blue) and *cis* (red) states.

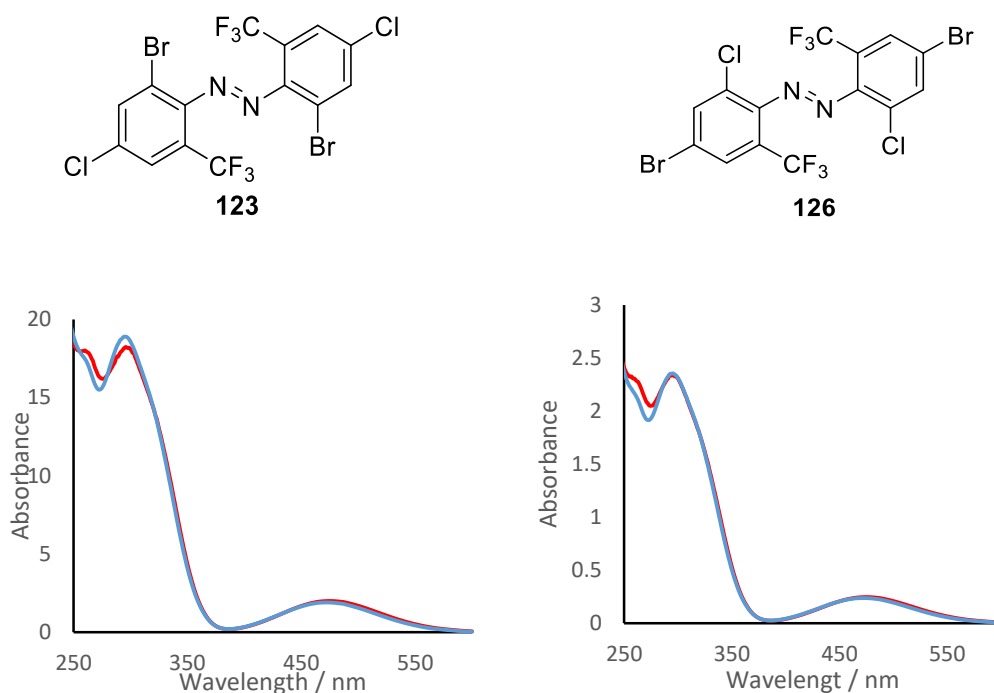
Compounds **121** and **124** reveal the remarkable influence of the *para* substituent with *para*-chloro **124** showing clear differences between the spectra of the *cis* and *trans* states whereas *para*-bromo **121** shows very little. However, even **124** shows no separation between its  $\pi$ - $\pi^*$  or  $n$ - $\pi^*$  absorbance bands in the *trans* and *cis* states. Spectra of **121** recorded at its dark and photostationary states show significant switching between the *trans* and *cis* states indicated by changes in the strength of the  $\pi$ - $\pi^*$  absorption bands with an estimated conversion of ~30% based on 360 nm absorbance as shown in Figure 108.



**Figure 10.8.** UV/visible absorbance spectra of **121** in the dark *trans* state (blue) and light *cis* state (red).

The *tetra-ortho* switches proved hard to switch due to the very similar spectra of *their trans* and *cis* isomers and a lack of appropriate filters to irradiate at 300 nm, the absorbance maximum of the  $\pi$ - $\pi^*$  transition. A slight shift in the  $n$ - $\pi^*$  between the *trans* and *cis* states is evident for **122**, but it is not large enough to generate a significant proportion of *cis-122* at its photostationary state.





**Figure 109.** UV/visible absorbance spectra of **122**, **123**, **125** and **126** in the dark *trans* state (blue) and light *cis* state (red).

The position of the  $\pi$ - $\pi^*$  band is significantly changed by the presence of a second halogen substituent in the *para* position, due to the shift in electronic structure of both the *cis* and *trans* state of the trifluoromethyl azobenzenes. There is some band separation between the  $n$ - $\pi^*$  orbitals in the spectra of some of these species, notably **123** and **126**, but the difference is not large enough to allow significant conversion from the *trans* to *cis* state by irradiation using 530 nm light. The extent of switching at the photostationary state is increased by the reduction of steric bulk in the *ortho* position, but the rate of reversion from *cis* to *trans* isomers is only partly determined by steric bulk as hindered *cis*-**126** has a shorter half-life than *trans*-**121**. Nevertheless, the presence of a second *ortho*-halo substituent in addition to the *ortho*-trifluoromethyl still leads to greatly increased photoswitch half-lives compared to that of the standard photoswitch **74**, due to its steric effects increasing the energy barrier to the rotation around the azo bond required for relaxation.<sup>60</sup> Further development was clearly still needed to create biologically useful switches.

### 3.4 Functionalising azobenzene to attach to a peptide

Because *ortho*-trifluoromethyl azobenzene photoswitches could not match the superior band separation and half-life properties of the tetra-*ortho*-halo substituted azobenzenes reported by Woolley<sup>62</sup> and Hecht<sup>60</sup>, it was decided to use *ortho*-halo



substituents and vary the linking structure at the *para* position of the aromatic ring to create novel structures. Azobenzene **137** was synthesised by literature procedures (Figure 110) to confirm its properties and as a starting point for further modification.

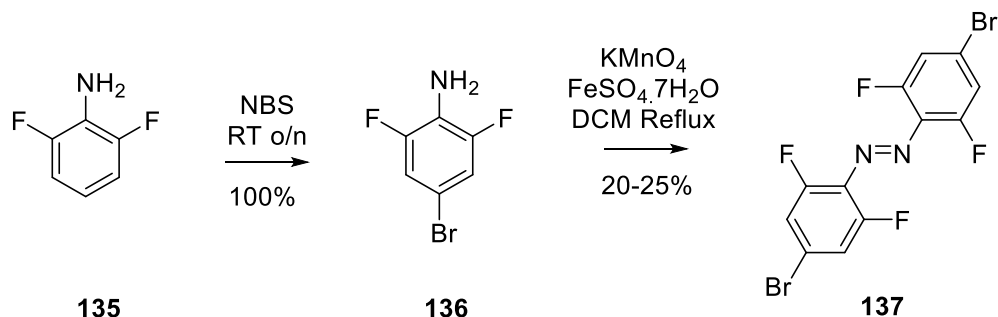


Figure 110. Synthesis of tetra-*ortho*-fluoroazobenzene.

The half-life of *cis*-**137** is 420 minutes at 40 °C in dimethyl sulfoxide corresponding to roughly 20 *E. coli* cell cycles.

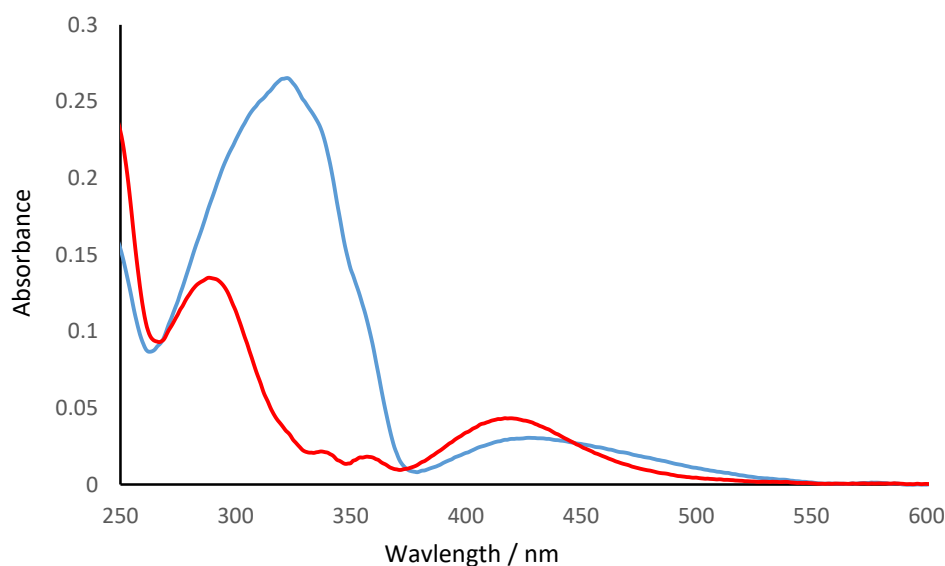
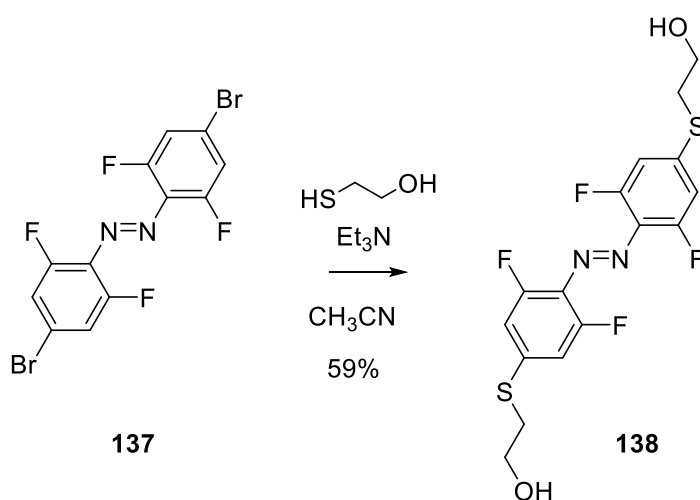


Figure 111. UV/visible absorbance spectra of *trans*-**137** (blue) and *cis*-**137** (red).

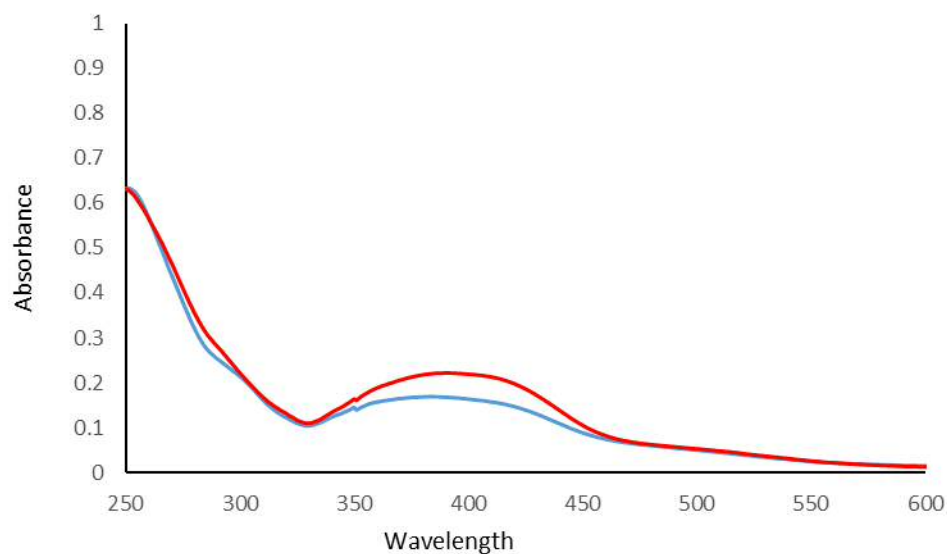
Unlike the *ortho*-trifluoromethyl compounds **121-126**, **137** does not show as great a  $\pi$ - $\pi^*$  band shift from the typical azobenzene value of 360 nm with an absorbance maximum at 340 nm. The other benefit of these switches is that they exhibit large shifts between the  $n$ - $\pi^*$  absorbance bands of the *cis* and *trans* isomers allowing the molecule to be switched with light above 500 nm as described in Section 1.5.2. This lower energy is less harmful for biological systems and penetrates more deeply through tissue. Starting from **137** allows for ready functionalisation of the *para* position by nucleophilic

aromatic substitution reactions, but the use of a nitrogen was avoided due to its half-life reducing properties as described earlier (Section 1.5.3). The groups installed must balance their electronic effect on the half-life of the light state of the azobenzene and allow functionalisation of the aromatic ring with a range of connecting groups to attach the switch to peptides and proteins. These connecting groups will be used to tune solubility, cell selectivity and the bioavailability of the molecule by providing attachment points for polyarginines, ethyleneglycol polymers or targeting proteins or sugars to encourage selective uptake by specific cell lines *via* endocytosis.<sup>108</sup>

A first choice of nucleophile was influenced by the work of Sawada, et al,<sup>64</sup> who showed that replacing the *para* amino group with a sulfur atom lead to a longer lived *cis* state and a greater extent of switching. This is due to the  $3sp^3$  orbital of the sulfur being larger in size and overlapping poorly with the smaller  $2sp^3$  of the carbon. This reduces the influence of its electron donating resonance forms, giving more prominence to its electron withdrawing character which reduces the rate of *cis*-azobenzene relaxation by destabilizing the azonium ion formation. 2-Mercaptoethanol was chosen as a prototype sulfur nucleophile to react with **137** to yield **138**.



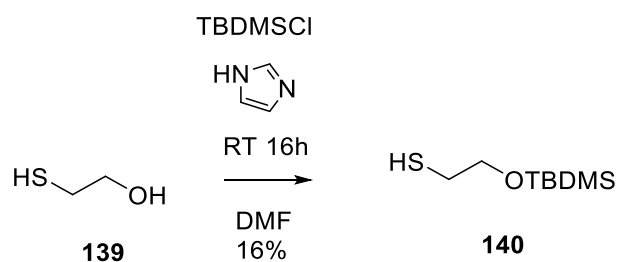
**Figure 112.** Insertion of  $\beta$ -mercaptoethanol at the *para* position of **137**.



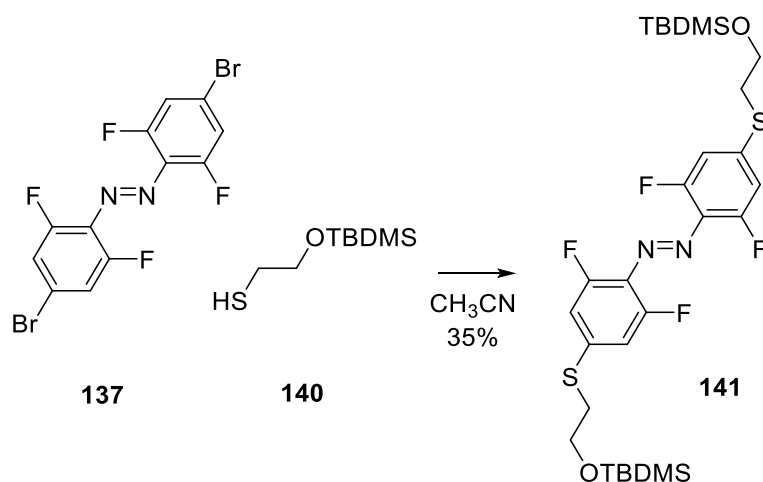
**Figure 113.** UV/visible absorbance spectra of the dark *trans* state (blue) and photostationary light state (red) of **138**.

The addition of sulfur to the *para* position reduced the half-life of *cis*-**138** to 150 minutes at 40 °C in DMSO from 440 minutes for **137** with a bromide at the *para* position. The presence of the sulfur broadened the absorbance peaks so only a single peak is observed for both the  $n-\pi^*$  and  $\pi-\pi^*$  absorbance bands with a longer shoulder resulting from these bands absorbing between 550 and 600 nm. It proved to be difficult to obtain a high resolution mass spectrum, a range of ionisation techniques were used but to no avail.

In order to prove this molecule was formed by reaction through the sulfur atom as intended,  $\beta$ -mercaptoethanol was reacted with chlorotrimethylsilane to mask its alcohol. Unfortunately, the trimethylsilyl protecting group proved to be too unstable for chromatographic separation necessitating use of the more robust *tert*-butyldimethylsilyl ether.<sup>109</sup> This reaction selectively protected the oxygen but the product is very sensitive to oxidation, with the formation of disulfides reducing the isolated yield of the reaction. Therefore, **140** was used immediately after synthesis or stored frozen under argon.



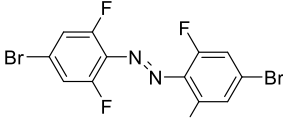
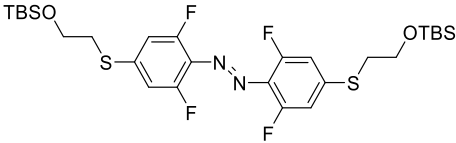
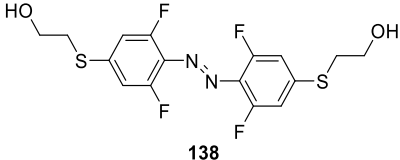
**Figure 114.** Protecting 2-mercaptoethanol with chloro-*tert*-butyldimethylsilane.



**Figure 115.** Reaction of **137** with *tert*-butyldimethylsilyl protected 2-mercaptoethanol.

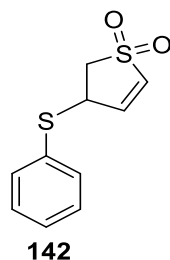
This protected version of **138** yielded satisfactory high resolution mass spectrometry data and once deprotected NMR data matched those of the product of the direct reaction of **137** and  $\beta$ -mercaptoethanol.<sup>109</sup>

**Table 18.** Properties of *para*-sulfur substituted azobenzenes.

Switch	$\pi-\pi^*$ (nm)	$n-\pi^*$ (nm)	% <i>cis</i> at PSS <sub>&gt;530 nm</sub>	Half-life (min) DMSO @ 40 °C
 137	320	440	~ 30 %	420 mins.
 141	420	indistinct	trace	148 mins
 138	400	indistinct	ND	ND

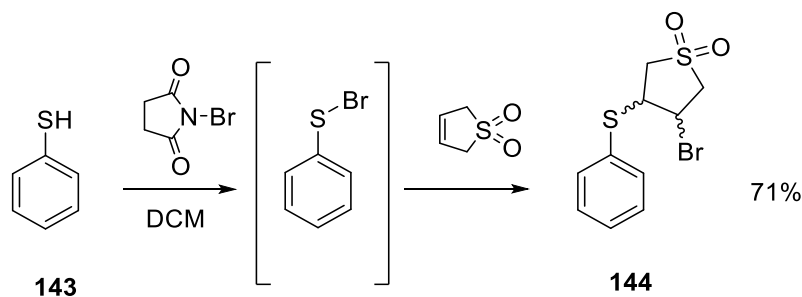
### 3.5 Alternatives to haloacetamides

Unlike amide substitution, a sulfide substituent in the *para* position should retain  $n-\pi^*$  band separation between the *cis* and *trans* isomers due to the lack of effective orbital overlap between the  $sp^3$  atomic orbitals of differing energy levels whilst retaining useful *cis* state half-lives.<sup>64</sup> Despite the disappointing results from exciting the  $n-\pi^*$  band of **138**, the  $\pi-\pi^*$  band is still red-shifted compared to **74**. Thioether-based linkers to attach azobenzenes to peptides could still provide substantial benefits. Several connecting groups and methodologies were investigated. One of the major factors to consider when developing a connecting methodology for attachment to peptides is the rigidity of the connecting group. If the group is too flexible it would lead to an ineffective switch as the shape change of the azobenzene is moderated by the flexible linker rather than being exerted directly on the peptide.



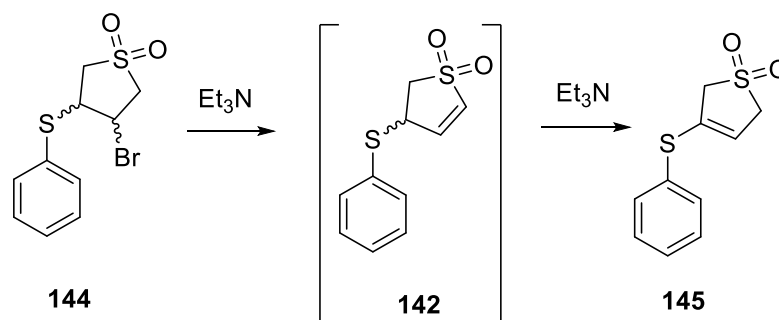
**Figure 116.** An dihydrothiophene sulfoxide thioether.

The addition of  $\alpha,\beta$ -unsaturated sulfoxide such as that in **142** to an azobenzene would add both a highly polar sulfoxide group to increase water solubility and a conjugated alkene as a strong Michel acceptor to react with cysteine sidechains. This would also break the planar nature of the *tetra*-fluorinated azobenzene, further improving solubility.



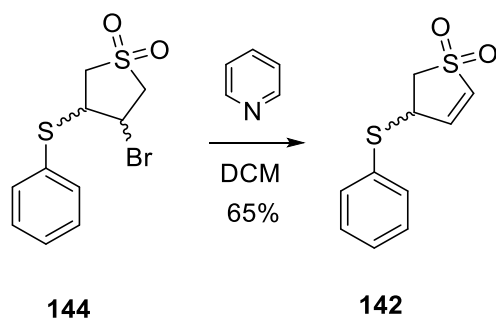
**Figure 117.** Synthesis of a model unsaturated sulfoxide linking group.

Compound **144** was synthesised from thiophenol as a prototype sulfur substituted aromatic ring using *N*-bromosuccinimide to form an intermediate containing a sulfur bromine bond. Addition of this intermediate to butadiene sulfoxide formed **144**. Elimination of hydrogen bromine to yield the Michael acceptor is triggered by the addition of triethylamine to cause an  $E_2$  elimination resulting in compound **142** but triethylamine also proved able to deprotonate the proton  $\alpha$  to the sulfur, causing the double bond to migrate to form the lowest energy isomer **145** (Figure 118).



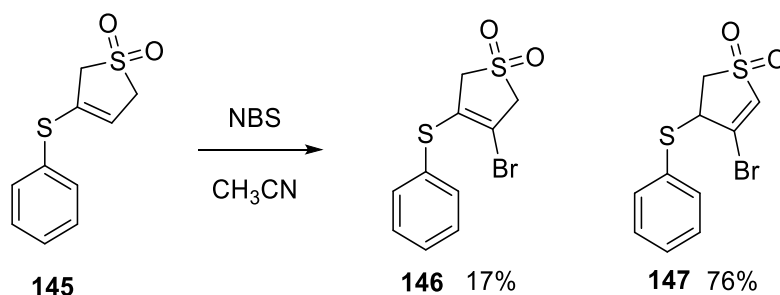
**Figure 118.** Dehydrobromination of compound **144** followed by double bond migration to yield **145**.

The final product shows a very simple <sup>1</sup>H NMR splitting pattern, consistent with the structure of the **145**. The <sup>13</sup>C NMR spectrum of **145** shows two peaks, identified as CH<sub>2</sub> by DEPT experiments, within 1 ppm of each other. Although the splitting pattern of the <sup>1</sup>H NMR is slightly complicated by the ring fixing the orientation of the protons, resonances were unambiguously assigned using 2D correlation spectroscopy which shows the product is **145**. Unfortunately, this constitutional isomer, with the double bond isolated from the sulfoxide results in **145** is not a Michael acceptor.



**Figure 119.** Debromination of **144** resulting in compound **142** with the double bond in the desired position.

Using a milder base, pyridine, prevented the double bond migration and resulted in the desired product with only one CH<sub>2</sub> present in <sup>1</sup>H and <sup>13</sup>C NMR spectra.<sup>110</sup> However, this product rapidly polymerises upon contact with mildly acidic conditions including silica. Model linker **142** therefore proved to be too reactive for use as a Michael acceptor as envisaged but the precursor bromide **144** might be used directly *via* a S<sub>N</sub>2 reaction.

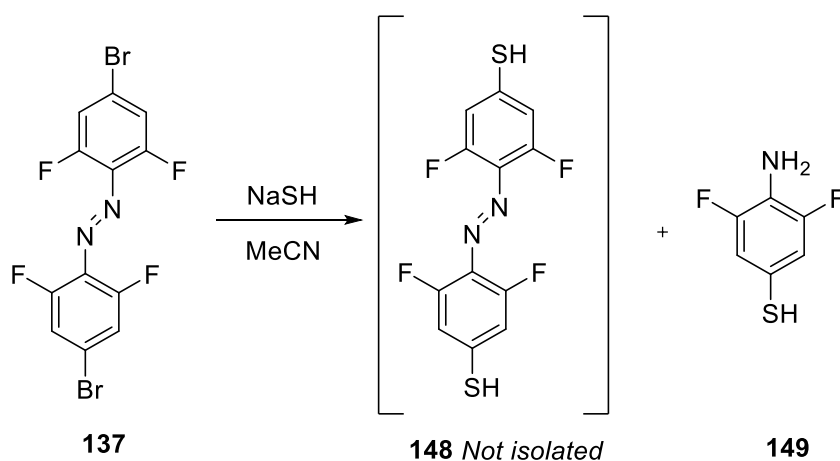


**Figure 120.** Bromination of compound **145**.

In order to create a less sensitive sulfur reactive linker compound **145** was rebrominated with *N*-bromosuccinimide resulting in a mixture of **146** and **147**. This gave primarily the desired product compound **147** so the conditions were not optimised.<sup>110</sup> This compound proved to be much less acid sensitive to acids with the electron withdrawing bromine deactivating the alkene sufficiently to prevent polymerisation. Preliminary liquid chromatography-mass spectrometry (LC-MS) results intriguingly suggested reaction of **147** with cysteine as a minimal peptide mimic occurred in Michael fashion with retention of the bromine, but conclusive evidence was not obtained.

### 3.5.1 Sulfur substitution attempts

To match the sulfoxide linker model system a sulfhydryl in the *para* position of azobenzene was required. Sodium hydrosulfide in acetonitrile was used as a nucleophile for aromatic substitution of the bromine of **137** in the same manner as that of the 2-mercaptoethanol.<sup>111</sup>

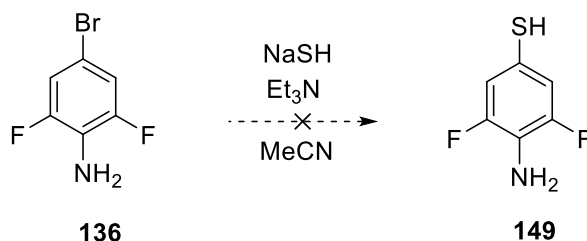


**Figure 121.** Attempted insertion of a *para* sulfur with sodium hydrogen sulfide.

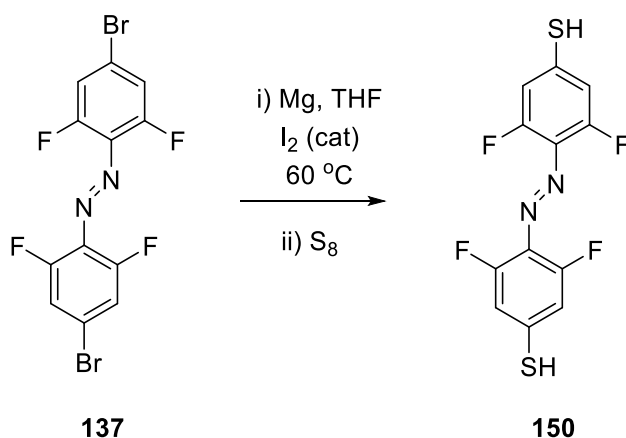
Unfortunately, reduction of the azobenzene to aniline was observed. Reaction in the presence of triethylamine in order to ensure the thiolate was present also showed no



product formation and degradation of the azo bond to aniline. Several different solvents were tried to ensure solubility of all reagents and both anhydrous and anoxic conditions were tested to prevent oxidation of sulfur. When a model nucleophilic aromatic substitution reaction was carried out on 2,6-difluoroaniline **136**, no product was observed even when refluxing. The starting aniline was quantitatively recovered, suggesting any thioaniline product must be derived from reaction of the azobenzene with consequent reduction.



**Figure 122.** Attempted reaction of sodium sulphide with *para*-bromo aniline.

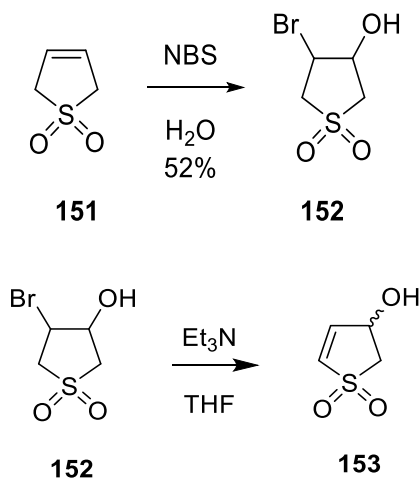


**Figure 123.** Grignard reagent formation for sulfur insertion.

In order to overcome this inherent inactivity of the electron deficient aromatic system, Grignard type reactions as described by Yi-Fan Huang *et al.*<sup>112</sup> using both azobenzene **137** and the aniline **136** were attempted, but neither yielded any recoverable *para* sulfide products. This was possibly again due to the reduction of the azo bond and the free amine quenching the reaction. Attempts to trap the crude product at the end of the reaction with chloroacetyl chloride or using acetic anhydride to trap a free thiol prior to work up all proved unsuccessful.

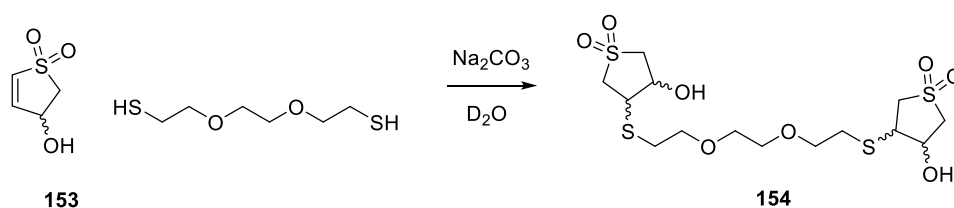
## 3.5.2 Oxygen substitution

Since sulfur proved so difficult to introduce, attempts were made to substitute the *para* position of azobenzene with an oxygen atom. Bromination of the double bond of commercially available **151** in water/tetrahydrofuran formed compound **152**.<sup>13</sup> The low yield is due to internal reaction of the bromohydrin product **152** to an epoxide.



**Figure 124.** Elimination of hydrogen bromide from **152** to form compound **153**.

This was followed by the elimination of hydrogen bromide with triethylamine. Unlike compound **144**, triethylamine does not cause the double bond to migrate due to the oxygen changing the  $pK_a$  of the  $\alpha$ -proton.<sup>13</sup> When tested to ensure that it was an effective Michael acceptor, **153** gratifyingly reacted smoothly with a test thiol when stirred in sodium carbonate buffer (50 mM).



**Figure 125.** Reaction of **153** with a test thiol to form compound **154**.

With **153** demonstrated to react with thiols as desired, all that remained was to attach it to an azobenzene. Unfortunately, deprotonation with sodium hydride to form a nucleophilic alkoxide caused polymerisation of the product so an Ullmann ether synthesis as described by Niu *et al.*<sup>14</sup> was attempted (Figure 126).

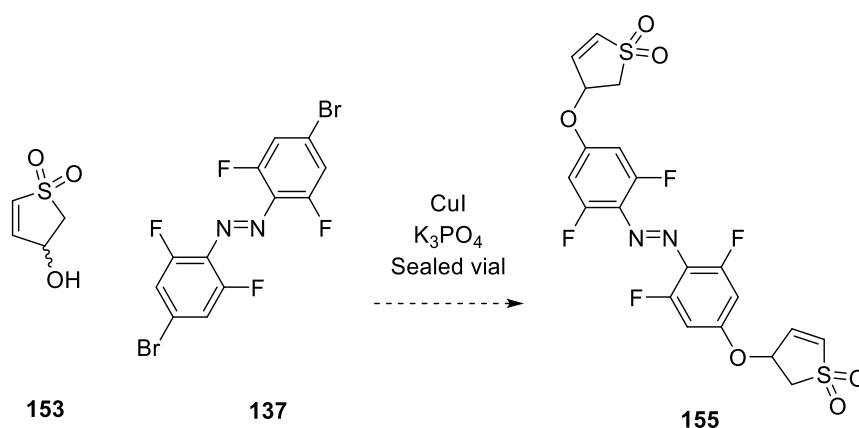


Figure 126. Ullmann ether synthesis.

Niu *et al* recommend 8-hydroxyquinolone to chelate the copper (I) and make it more reactive,<sup>114</sup> but this reaction failed to yield the expected product.

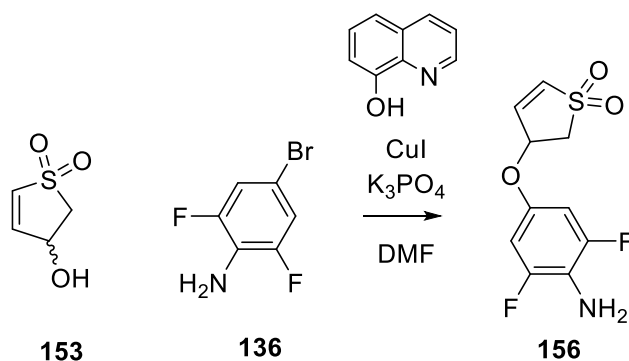
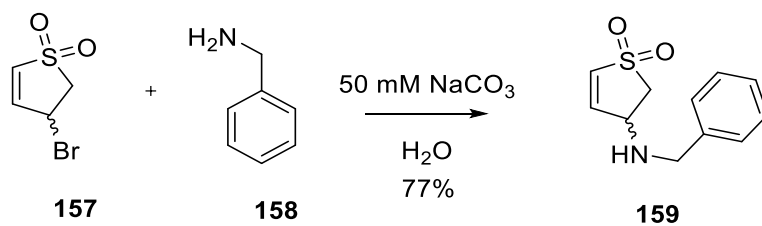
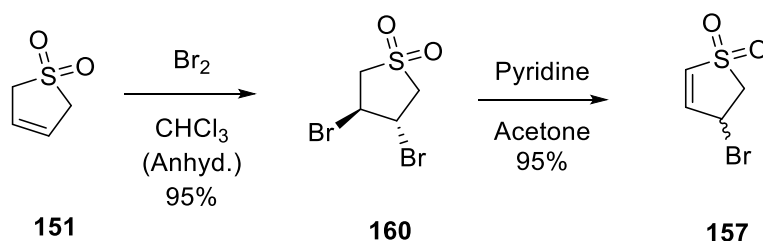


Figure 127. Ullmann ether synthesis.

Substituting the azobenzene system at a late stage having proved much more difficult than expected, it was decided to incorporate an additional amine group at the aniline stage in order to simplify the necessary chemistry in the final step in order to allow conjugation to peptides. In order to avoid the previously described deleterious effects of nitrogen substitution at the *para* position, a benzylamine, rather than aniline system was envisaged. In order to ensure that this is a viable a new crosslinking methodology the system was tested with benzylamine (Figure 128).

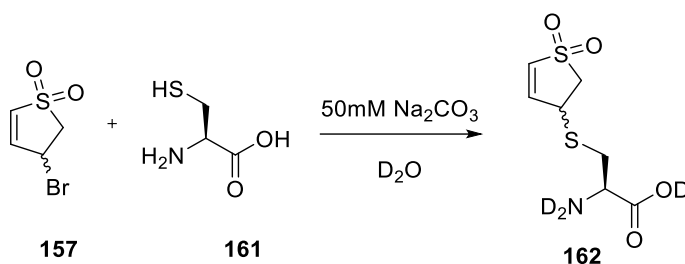


**Figure 128.** Benzylamine reaction with compound **157**.



**Figure 129.** Synthesis of compound **157** by dibromination and elimination.

Butadiene sulfone was reacted with bromine to yield **160** which was recrystallized from acetone.<sup>115</sup> Pyridine in acetone then effected an elimination to the vinyl sulfone.



**Figure 130.** Reaction of **157** with cysteine.

Surprisingly, cysteine reacted with the bromine of **157** in preference to the strong Michael acceptor. This indicated that only the bromide needs to be formed to create a linker that can subsequently be used to attach to the peptide. The fact that the thiol of cysteine reacted faster than the amine suggests this system will be usable with peptides with free amine sidechains if stoichiometries are carefully controlled to avoid unwanted side reactions.

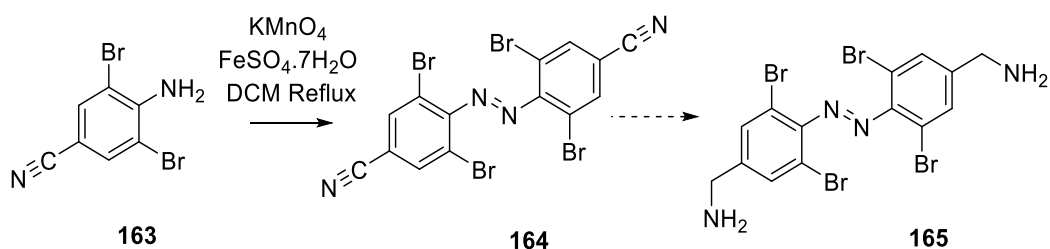


Figure 131. Synthetic pathway toward compound 165.

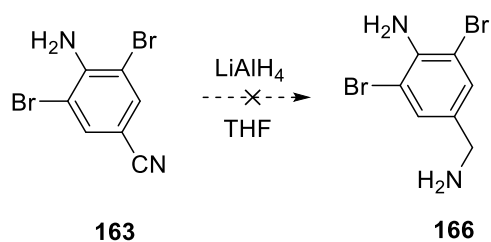
Installation of the methyleneamine at the *para* position of **165** increases the length of the switch and makes it more flexible. This might be advantageous to remove some of the strain in the  $i,i+4$  spacing and also allow the extra extension in the  $i,i+11$  state, but also risks reducing control of the conformation of bound peptides.

The azo bond was then formed as described previously in Section 4.1. Reduction of the nitrile to give the desired benzyl amine was then attempted by a number of methods (Table 19).

Table 19 Conditions used to attempt nitrile reduction.

Reductant	Conditions	Hydrolysis reagent
LiAlH <sub>4</sub>	2 hours, 20 °C	i) HCl ii) NaOH
LiAlH <sub>4</sub>	16 hours, 20 °C	HCl - NaOH
LiAlH <sub>4</sub>	16 hours reflux	HCl - NaOH
LiAlH <sub>4</sub>	16 hours reflux	HCl
LiAlH <sub>4</sub>	16 hours reflux	sodium potassium tartrate
LiBEt <sub>3</sub> H	16 hours reflux	sodium potassium tartrate
Pd/C, H <sub>2</sub> , MeOH	16 hours, 20 °C	n/a

These reactions appeared to proceed by TLC, but no product could be isolated. Attempts to trap amine from the crude reaction mixture by immediately protecting it with a Boc or acetyl group failed, with no amine or protected amine was isolated. It is possible that the lithium aluminium hydride is reducing the azo bond leading to unexpectedly water soluble products.



**Figure 132.** Reduction of compound **163**.

A reduction was also carried out (Figure 132) on compound **163** to test the feasibility of the reaction without the presence of an azo bond, but no identifiable products were isolated.

# Chapter 4

## Refining Photoswitch Properties

## 4 Refining photoswitch properties

### 4.1 Tetra-ortho-chloro-di-para-hydroxymethylenylazobenzene

In order to avoid the intractable diamine **165** and its derivatives, diacid **169** was reduced to alcohol **170**. This could then be esterified with chloroacetyl chloride to create a thiol-reactive ester (**171**) reminiscent of a diester used by Mascarenas, *et al.*<sup>16</sup>

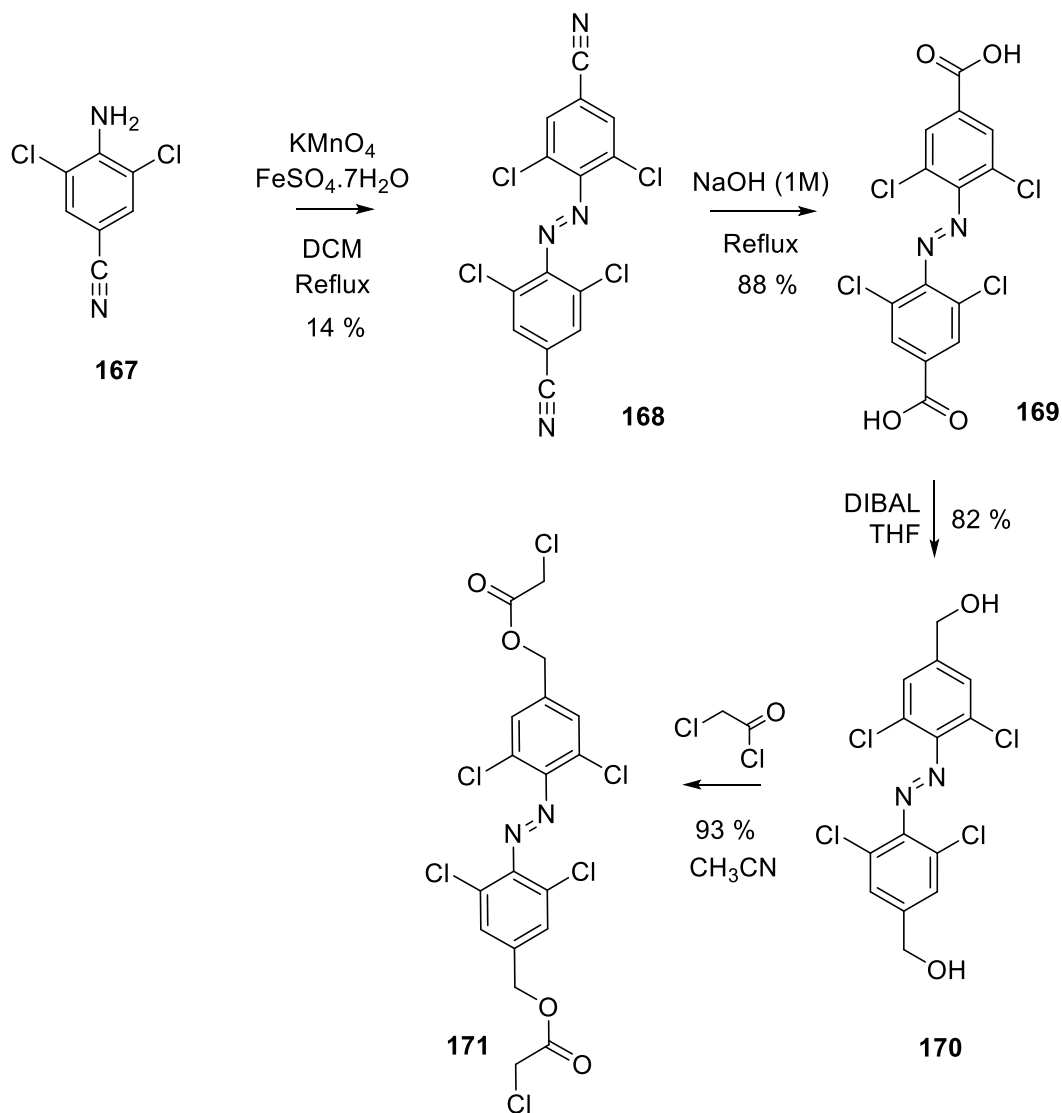
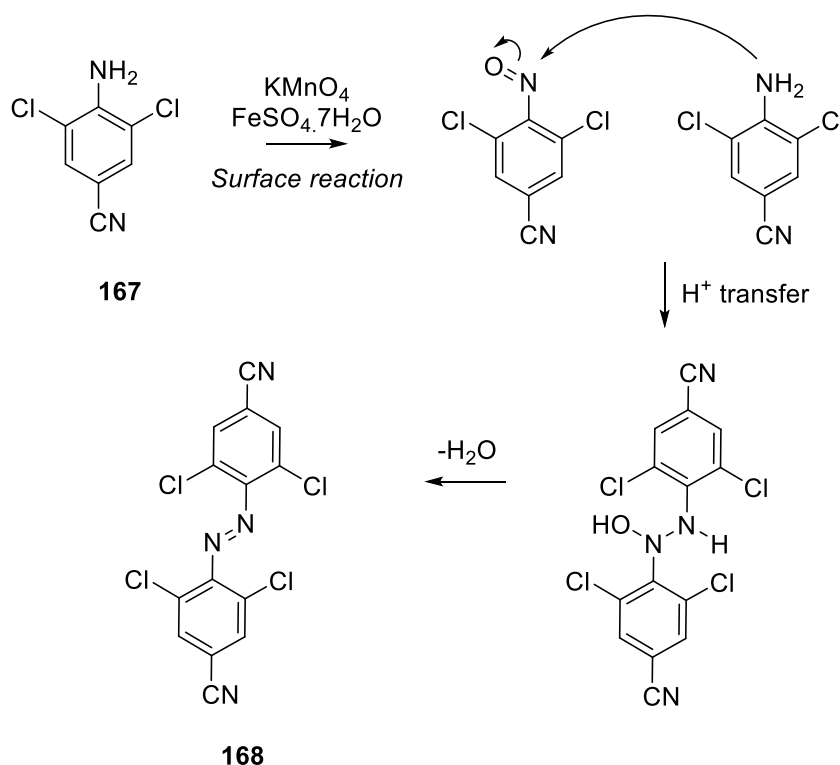


Figure 133. Synthesis of Compound **171** tetra-ortho-chloro benzyl alcohol.

Consulting the literature, the benzyl ester bond was reported to be stable to hydrolysis in the mild conditions of the peptide cross linking reaction. The first route attempted is shown in Figure 133, with the azo bond formed using potassium permanganate and iron sulphate septahydrate.



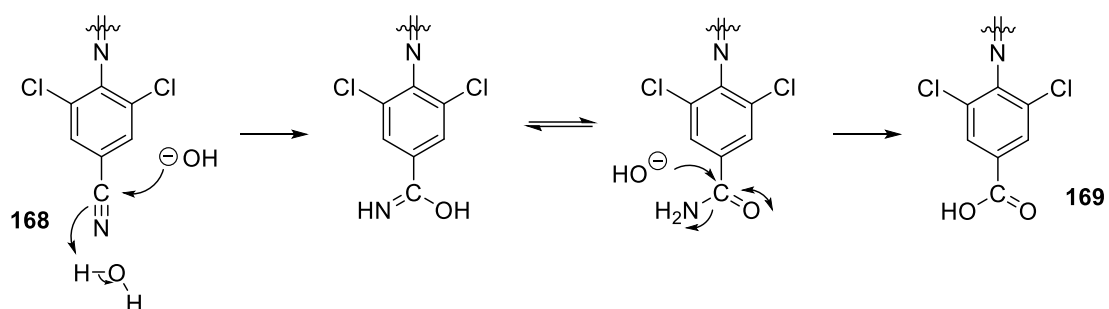


**Figure 134.** Oxidative dimerisation through azo bond formation.

This reaction is believed to proceed by the coordination of the amine to the iron sulphate surface and then the oxidation of the amine to a nitroso species. This can then undergo an attack by a second molecule of amine with consequent elimination of water resulting in the formation of the azo bond. The low yield of this reaction with bulky *ortho* substituents could be due to inefficient condensation leading to competing over-oxidation or because the proportion of azobenzene formed in the *cis* isomer in some of the cases is quite high and this isomer is easier to further oxidise than the *trans* form, leading to loss of product through over oxidation. An over-oxidation product, the nitro derivative of **167**, is one of the major impurities present.

In the case of anilines with electron-withdrawing substituents in the *para* position, the reduced electron density around the nitrogen reduces efficiency further. The oxidant, support surface and the solvent all play important roles in determining the efficiency of the oxidation reaction and were extensively studied by Hecht and co-workers.<sup>118</sup> Changes to the relative ratios of potassium permanganate to iron sulphate were found to cause little difference with no clear trend emerging. In contrast to earlier studies where changes in the solid support could yield up to 8% changes in overall yield, there is little or no effect of changing the support and no reagents increased the yield above 20-35%.<sup>61,118</sup>

Several other reagents have been reported to carry out this oxidative dimerisation reaction including other examples of supported metal catalysis. For example, the use of a gold supported catalyst (Au/meso-CeO<sub>2</sub>) and nitrobenzene starting materials with carbon monoxide as the sole deoxygenate was recently reported.<sup>119</sup> This method obtained high yields for many examples with several over 80% and was also shown to be compatible with halide substituted aromatic rings. A simpler copper (I) bromide reagent has also been shown to give up to 94% conversion of aniline to azobenzene,<sup>120</sup> but copper can react with the aromatic halogens by insertion of the copper into the carbon-halide bond and could in theory lead to a range of undesired products and so was not tested. Another strategy commonly employed to synthesise non-symmetric azobenzenes mixed preformed nitroso and aniline species to join together two similar or dissimilar aromatic rings. However, whilst the yield was poor the potassium permanganate/iron sulfate heptahydrate reaction was sufficiently early in the synthesis to remain viable.



**Figure 135.** Mechanism for base-catalysed hydrolysis of the nitrile.

Base catalysed hydrolysis of a nitrile to a carboxylic acid proceeded smoothly upon refluxing overnight with either sodium or potassium hydroxide (1 M). The resulting dibenzoate **169** was then precipitated by addition of hydrochloric acid to give the corresponding acid in excellent yields (~88 %).

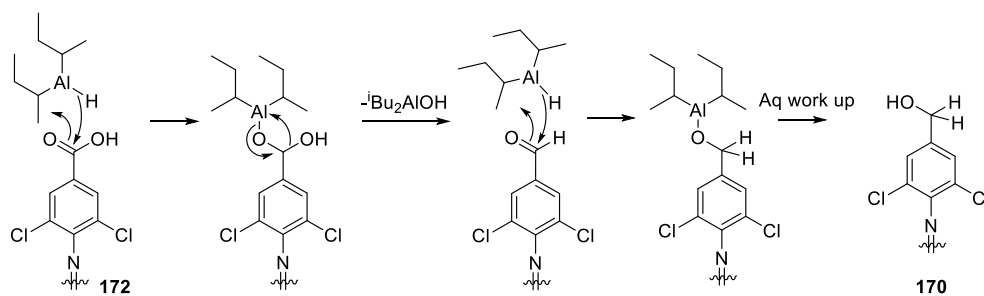


Figure 136. Reduction of ethyl ester.<sup>121</sup>

Reduction of the dibenzoic acid to an alcohol was achieved with diisobutyl aluminium hydride (Figure 136).<sup>121</sup> The azo bond is protected from reduction by the steric bulk of the *ortho* substituents preventing coordination to the azo bond. The resulting alcohol was then reacted with chloroacetyl chloride to give cross linker **171** (Figure 137).

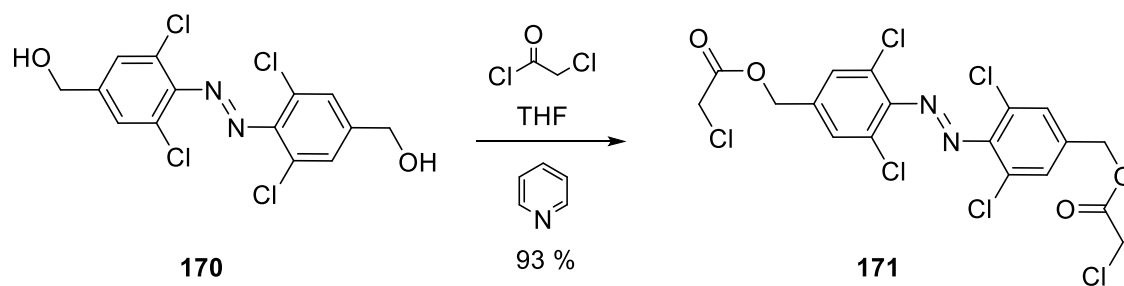


Figure 137. Esterification of **170** to chloroacetate ester **171**.

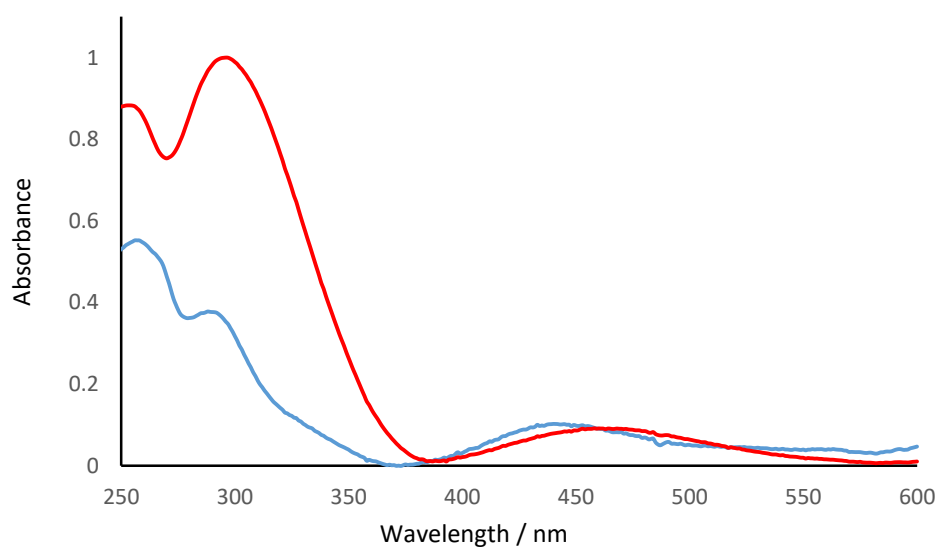
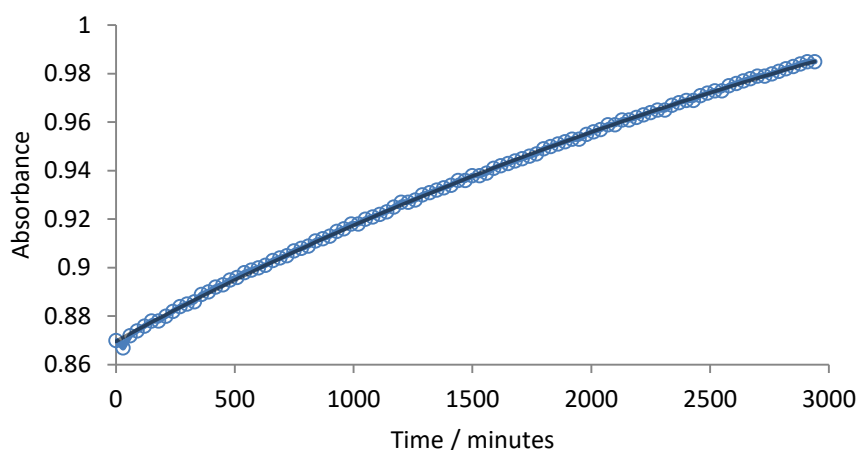


Figure 138. UV/visible absorbance spectra of *cis*-**171** (blue) and *trans*-**171** (red).

UV/visible absorbance spectra for **171** (Figure 138) were obtained using an HPLC equipped with a diode array UV detector. For each peak on the HPLC trace a full UV spectra was recorded, giving spectra for the pure *cis* and *trans* isomers. The spectra showed a significant difference in the  $\pi$ - $\pi^*$  absorbance band and a large  $n$ - $\pi^*$  band shift of 21 nm as described by Hecht *et al.*<sup>60</sup> This band shift allows the photo switching from the *trans* state to the *cis* state using of 530 nm light. Dark-adapted **171** in dimethylsulfoxide was irradiated for 15 minutes through an OG 530 filter (>530 nm) and absorbance at 300 nm was monitored to determine the half-life of the *cis* isomer. At 20 °C no relaxation was evident above noise of the UV/visible spectrometer after 2 days and repeating the experiment at 40 °C for 2 days gave the same result. Carrying the experiment out over 2 days at 60 °C gave data that could be fit to a first order rate equation to give a relaxation time of 3150 minutes or 53 hours.



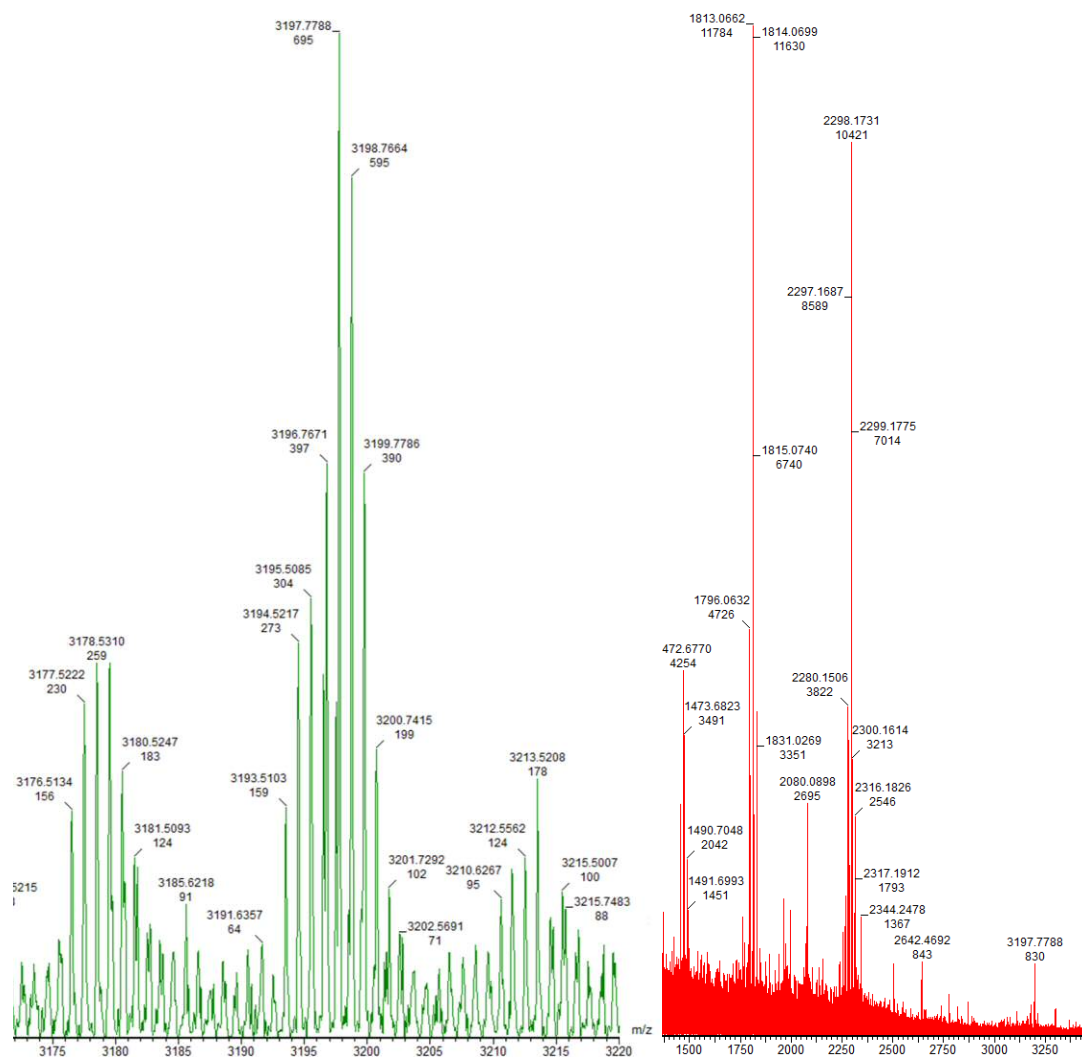
**Figure 139.** Relaxation curve showing absorbance at 300 nm for compound **171** in DMSO at 60 °C.

The half-life of *cis*-**171** is an order of magnitude greater half-life than that of the trifluoromethyl photoswitches, even at an elevated temperature, and approximately three orders of magnitude longer than **71**. Azobenzene **171** was attached (Section 4.6) to a fluorescently labelled peptide (FAM-Bid<sup>i,i+4</sup>) derived from a pro-apoptotic protein similar in nature to Bak (Sections 1.6.4).

	Sequence	Mass
FAM-Bid <sup>i,i+4</sup>	FAM-DIIRNIARHLACVGDICDRSI-NH <sub>2</sub>	3191

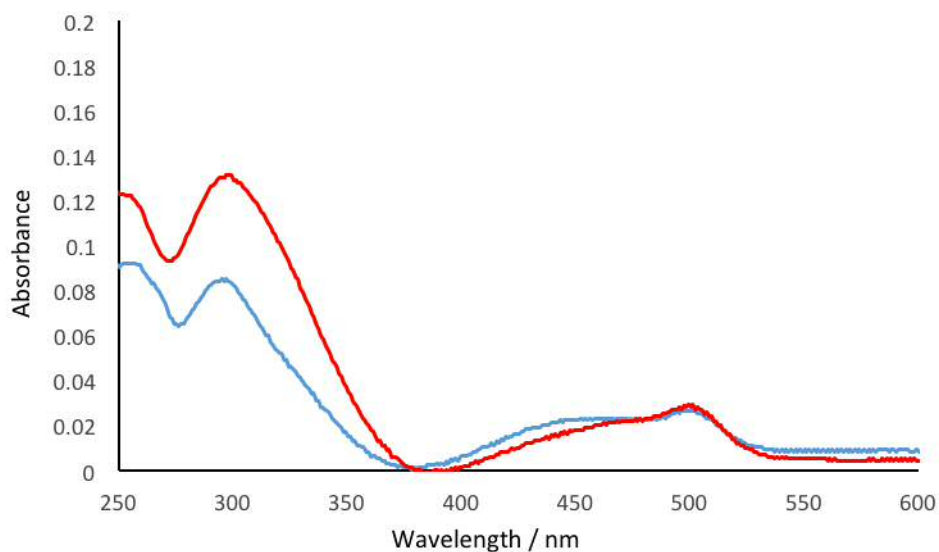
**Table 20** Sequence of FAM-Bid<sup>i,i+4</sup>

This peptide contains two cysteines at a relative spacing of  $i, i+4$  for attachment of the peptide shown in Table 20. This spacing gives attached photoswitches a large degree of control over the peptide  $\alpha$ -helix structure as the extended *trans* isomer strongly disfavours helix formation.<sup>84</sup> This spacing should be ideal for the slightly increased length of the **171** compared to **71**, and the extra flexibility of the switch might help it adopt a stable helix as the *cis* isomer.

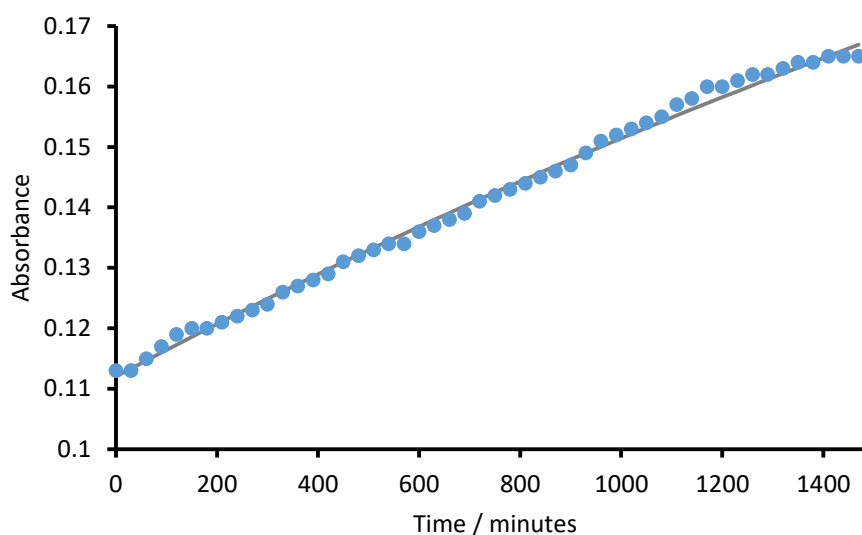


**Figure 140.** ESI (left) and MALDI (right) mass spectra for the crude crosslinked FAM-Bid<sup>*i+4*</sup>-**171** showing molecular ions (calculated monoisotopic mass 3191) and sodium adducts.

The calculated monoisotopic mass of the peptide and crosslinker is 3191 as the peptide has four chlorine atoms this causes the isotopic distribution shown above (Figure 140) and explains the slightly elevated mass observed for the peptide.

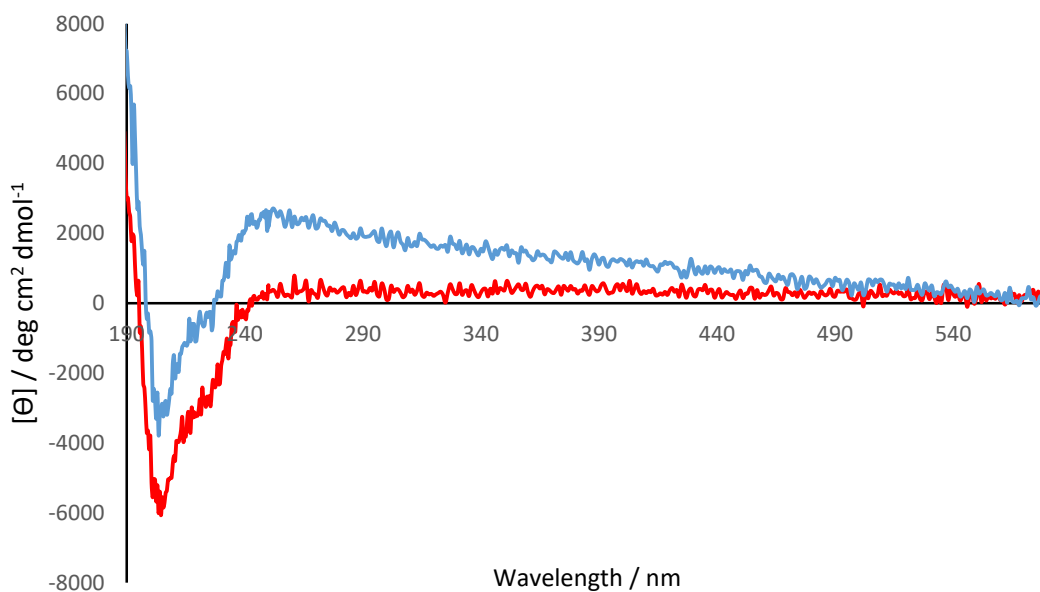


**Figure 141.** UV/visible spectra of crude FAM-Bid<sup>i+4</sup>-171 in the dark adapted states (red) and after irradiation with >530 nm light (blue).



**Figure 142.** Relaxation curve of absorbance at 300 nm for crude FAM-Bid<sup>i+4</sup>-171 in DMSO at 60 °C.

This crosslinked crude peptide showed reversible changes in its UV/visible spectra upon irradiation with >530 nm light with a half-life of 1700 or 28.5 hours for FAM-Bid<sup>i+4</sup>-*cis*-171.



**Figure 143.** CD spectra of crude FAM-Bid<sup>i+4</sup>-171 in the dark adapted states (red) and after irradiation with > 530 nm light (blue).

The CD spectra show a change in structure upon irradiation, as expected for a photoswitch-induced change in  $\alpha$ -helical character. Given this preliminary data for successful peptide cross linking, the effect of the different halogen substitution patterns was investigated to see if azobenzenes could be further tuned to develop long-lifetime, fast-switching azobenzenes with pronounced  $\pi$ - $n^*$  band separation.

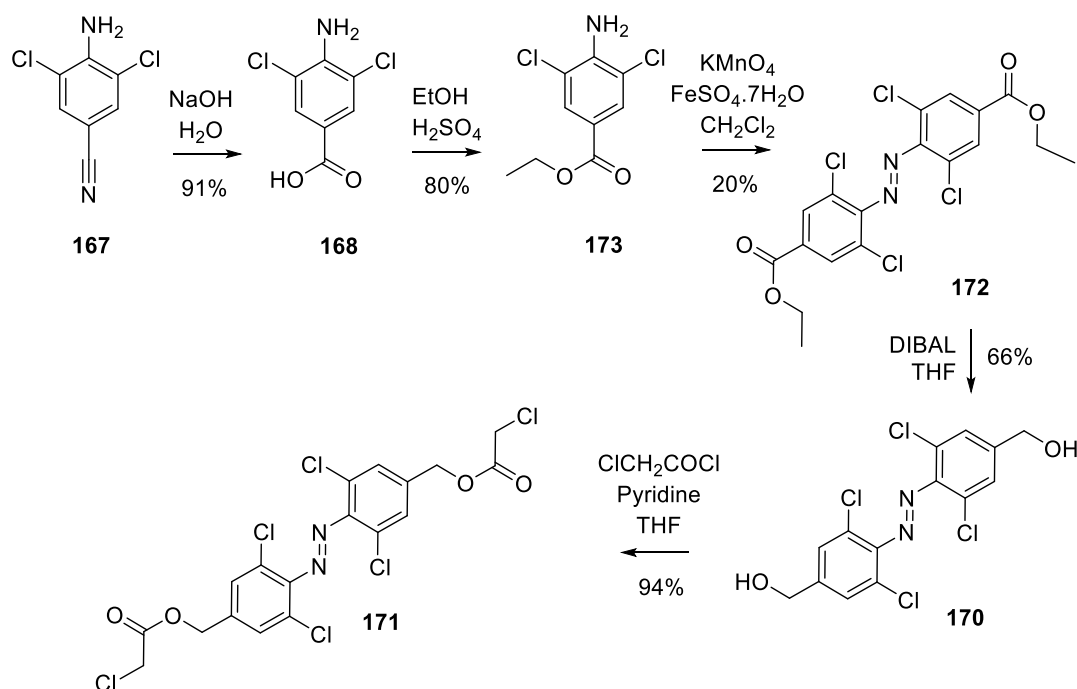
4.2 Synthesis of tetra-*ortho*-haloazobenzenes

Figure 144. Modified synthesis of the tetra-*ortho*-chloro substituted azobenzene **171**.

Slight modifications to the synthetic route improved the ease of synthesis by reducing the need for silica gel chromatography. Compound **171** was synthesised from commercially available 4-amino-3,5-dichlorobenzonitrile as shown above in Figure 144.

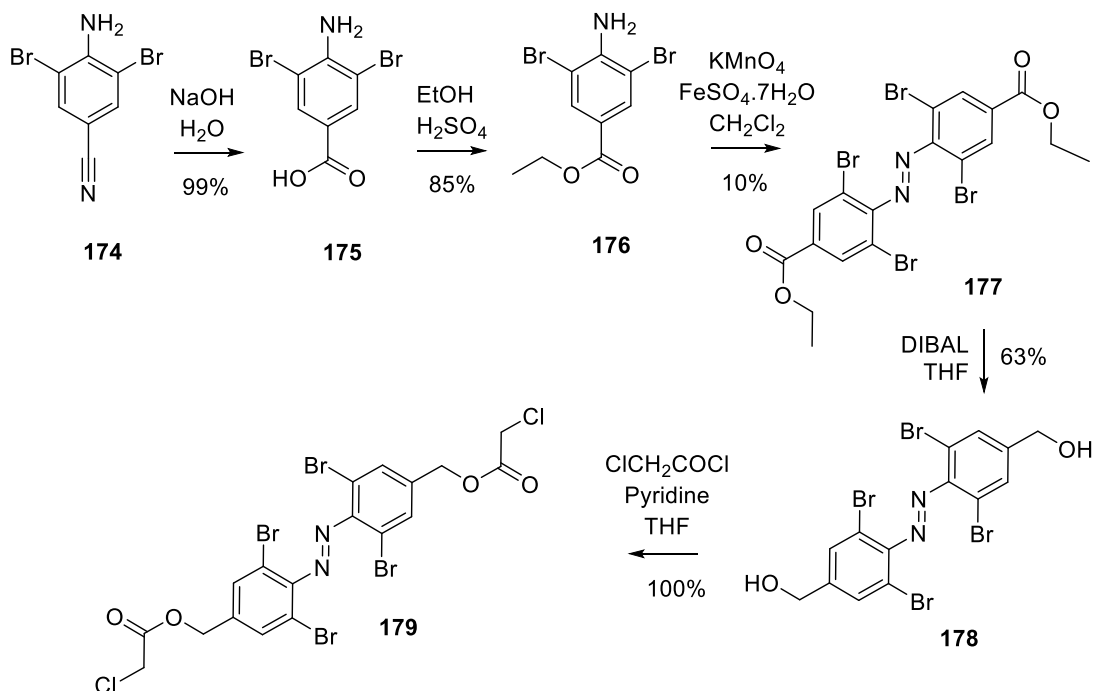


Figure 145. Synthetic scheme for tetra-*ortho*-bromo substituted azobenzene **179**.



The same route was used to prepare compound **179** from commercially available 4-amino-3,5-dibromobenzonitrile **174** (Figure 145), although the key oxidation step was only half as effective with the bulkier bromine substituents. The synthesis of tetra-*ortho*-fluoro photoswitch **185** is slightly longer, starting with commercially available 2,6-difluoroaniline **135**. This is readily brominated in the *para* position with *N*-bromosuccinimide with electronic assistance from the amine.

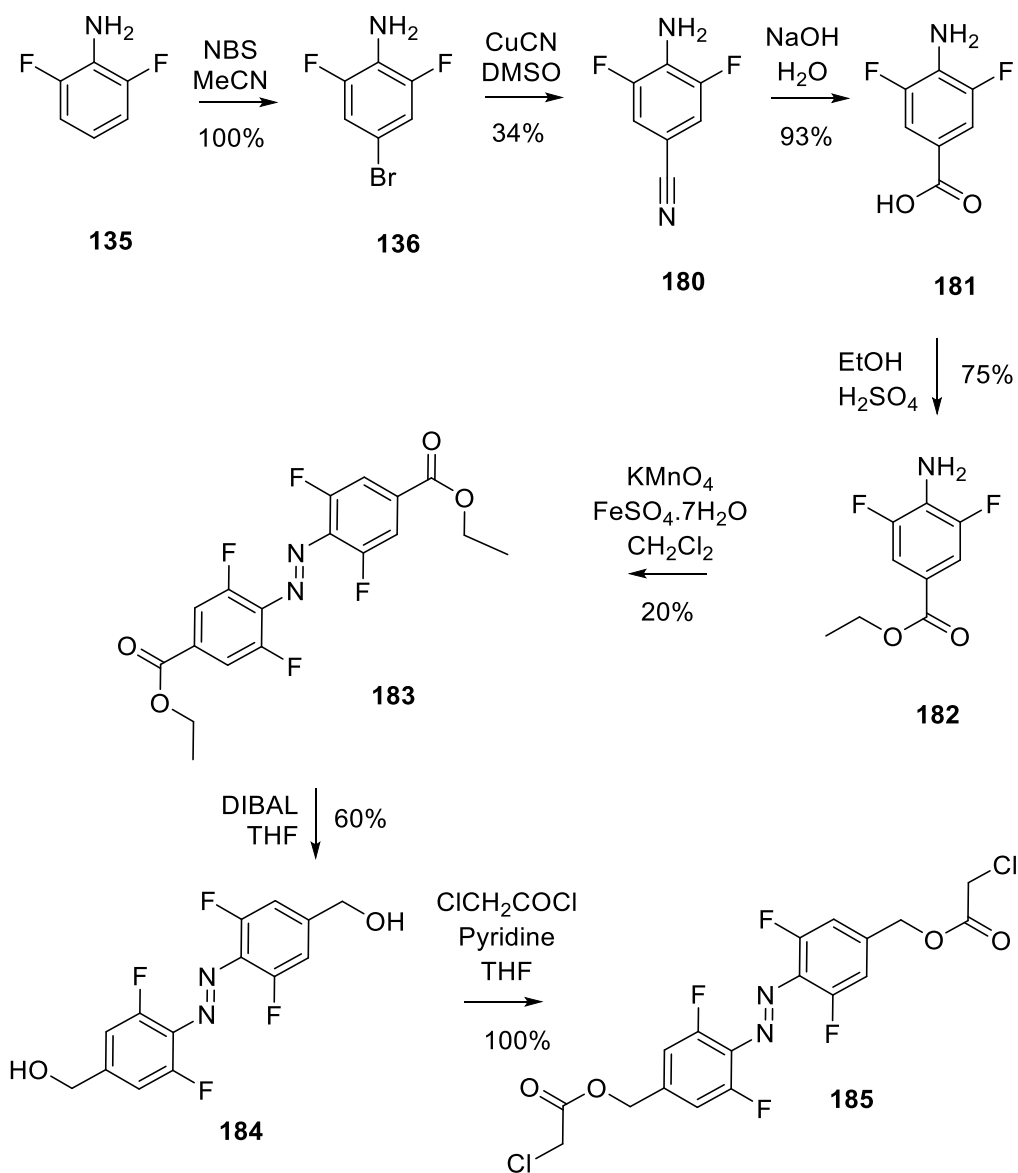


Figure 146. Synthetic scheme for tetra-*ortho*-fluoro substituted azobenzene **186**.

To determine which properties are due to *ortho*-halogens the non-halogenated analogue was synthesised using the same methodology (Figure 147) starting from commercially available 4-aminobenzonitrile to eventually yield **191**.

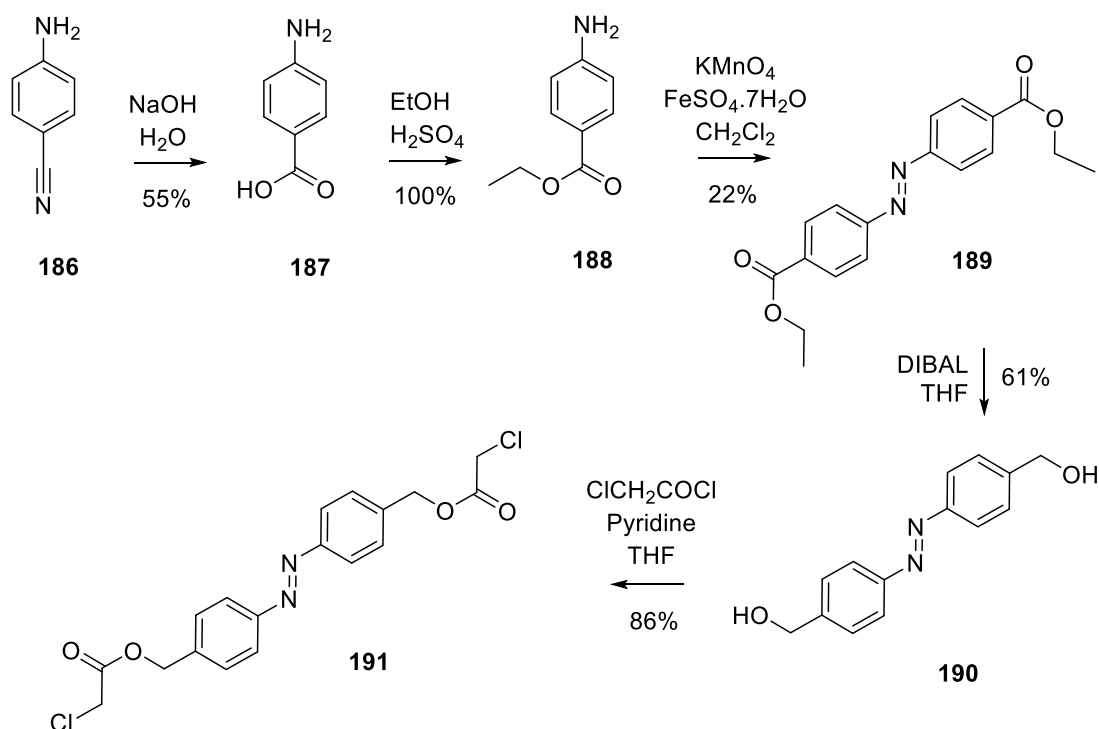


Figure 147. Synthetic scheme for tetra-*ortho*-hydroazobenzene **191**.

### 4.3 UV/visible spectra of tetra halogenated azobenzenes

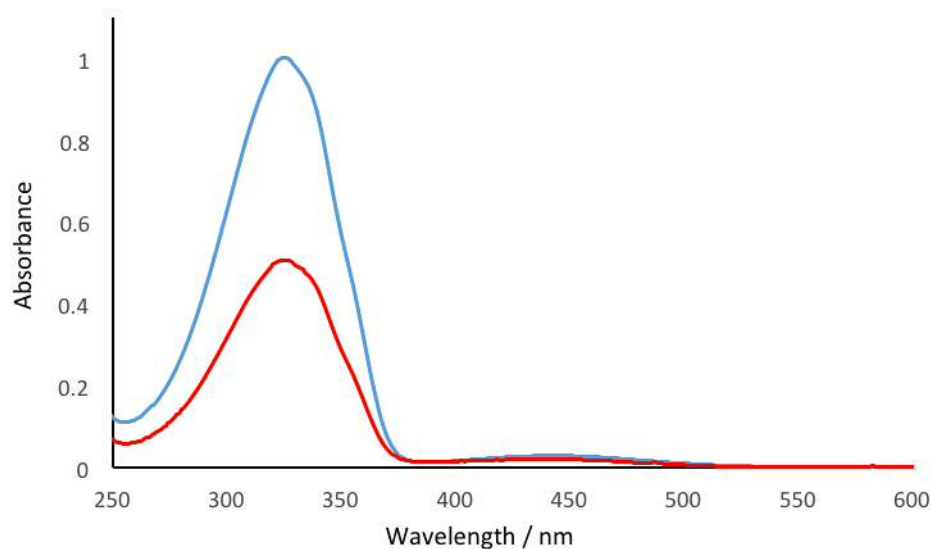
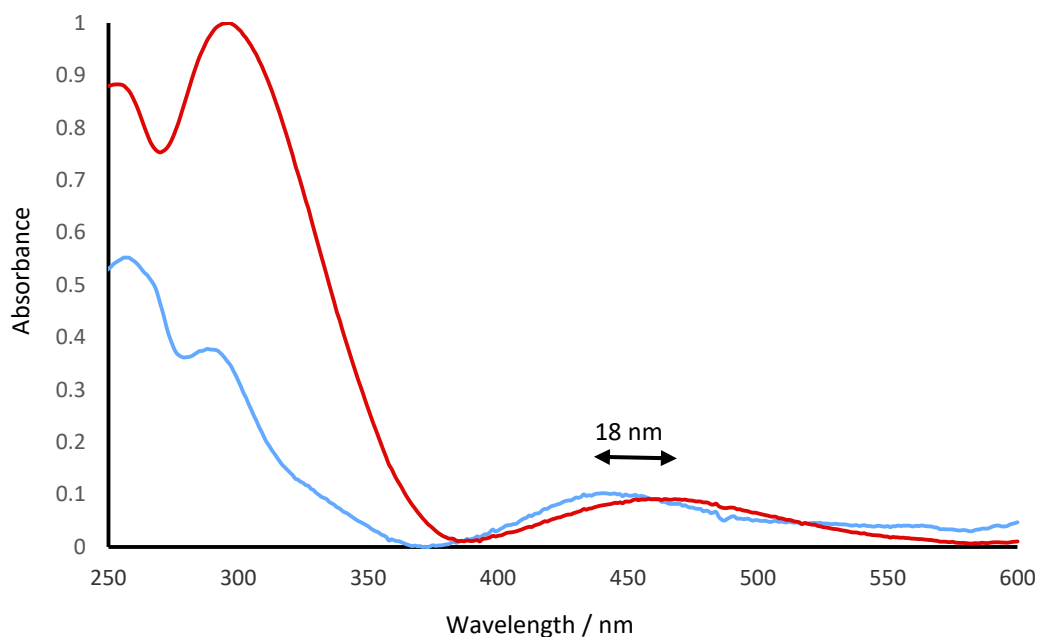


Figure 148. HPLC-DAD UV/visible spectra of *cis*-**191** (blue) and *trans*-**191** (red).

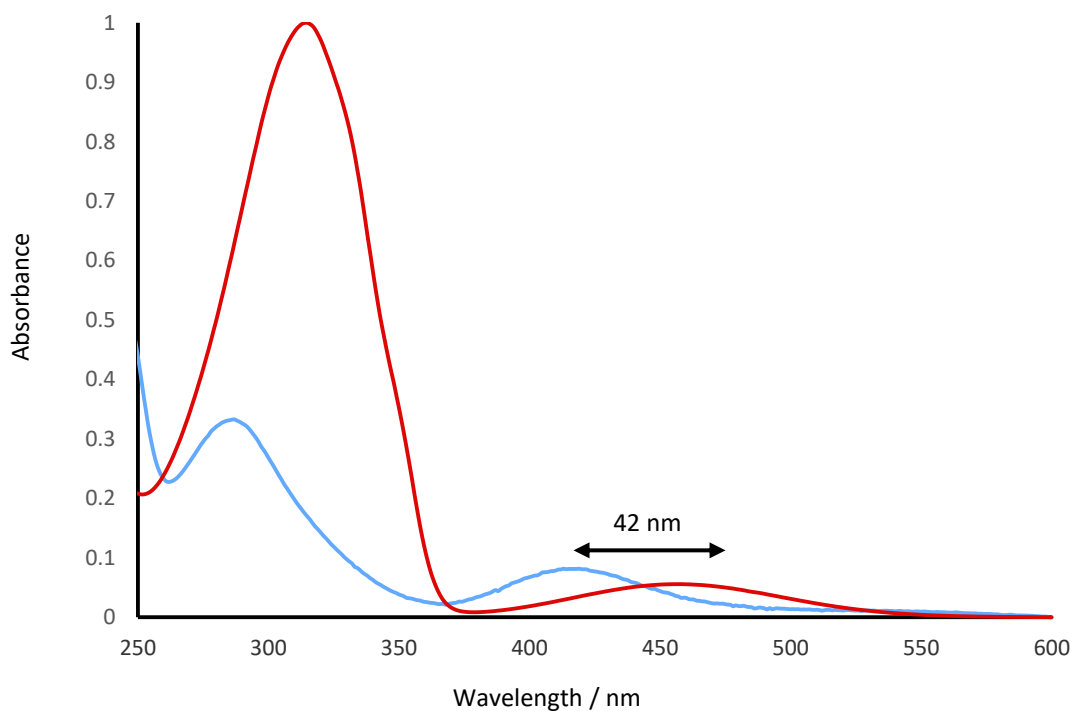
The spectrum of reference photoswitch **191** shows an absorbance band for the  $\pi$ - $\pi^*$  transition of 320 nm. The half-life of *cis*-**191** is only 2 hours at 60 °C in dimethylsulfoxide, an order of magnitude less than tetra-*ortho*-chloro *cis*-**171**.

Azobenzene **191** shows no difference between the maxima of the  $\pi$ - $\pi^*$  or  $n$ - $\pi^*$  absorption bands in the *cis* and *trans* isomers.



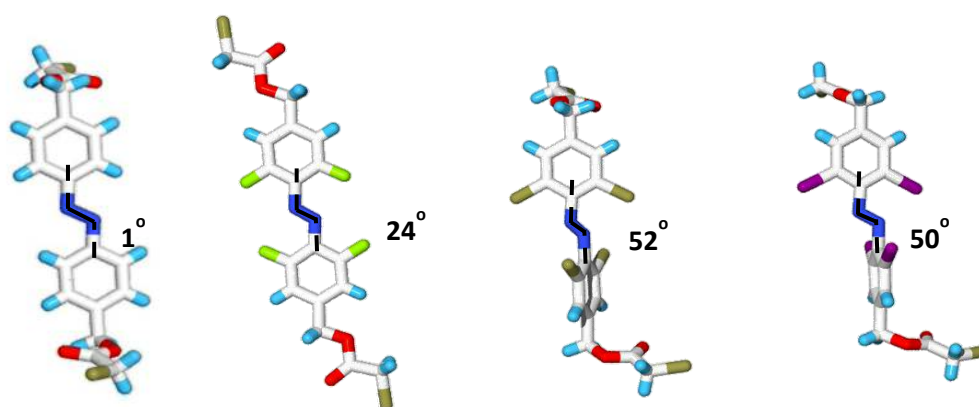
**Figure 149.** HPLC-DAD UV/visible spectra of *cis*-**179** (blue) and *trans*-**179** (red).

The tetra-*ortho*-bromo azobenzene **179** shows a slightly smaller 18 nm  $n$ - $\pi^*$  band separation compared to the chloride equivalent **171**, but the spectra are otherwise very similar as is the half-life of *cis*-**171** at 54 hours at 60 °C.



**Figure 150.** HPLC-DAD UV/visible spectra of *cis*-**185** (blue) and *trans*-**185** (red).

In contrast, tetra-*ortho*-fluoro azobenzene **185** has significantly different absorbance properties. It has an  $n\text{-}\pi^*$  band separation of 42 nm, double that of **179** and **171**. The  $\pi\text{-}\pi^*$  absorbance band of *trans*-**185** is at 320 nm whereas the corresponding tetra-*ortho*-chloro and tetra-*ortho*-bromo *cis*-isomers have their absorbance bands blue-shifted to 300 nm. To investigate the origin of these different properties single crystal X-ray diffraction patterns were recorded and structures solved by Dr Benson Kariuki. Coordinates for these crystal structures can be found in Appendix 7.2.

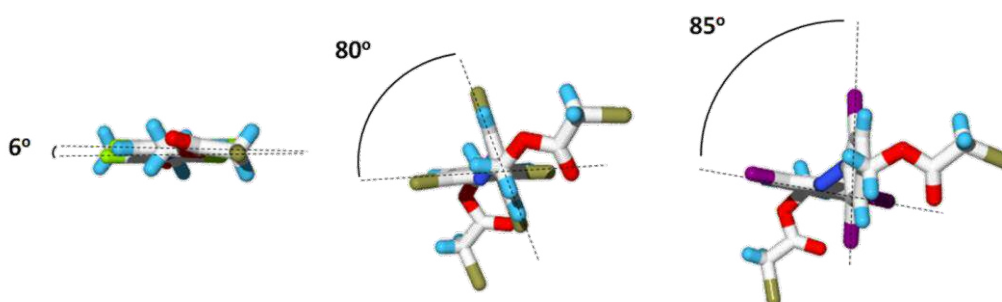


**Figure 151.** Crystal structures of **191**, **185**, **171** and **179** showing the dihedral angle of the azo bonds.

**Table 21.** Comparison of the properties of the tetra-*ortho*-haloazobenzenes.

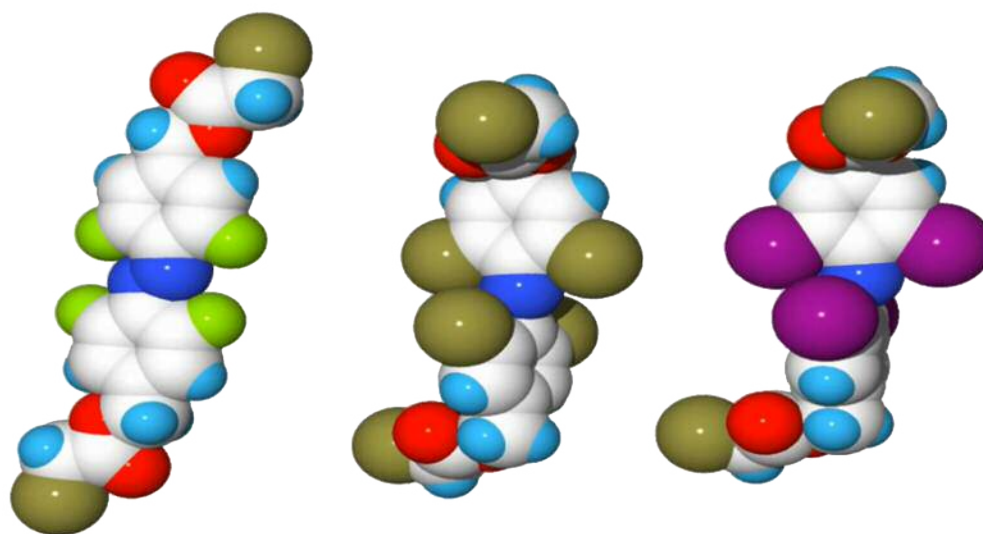
	Dihedral angle	$\pi-\pi^*$ $\lambda_{\max}$ (nm) <i>trans</i>	$n-\pi^*$ $\lambda_{\max}$ (nm) <i>trans</i>	$\Delta(\lambda_{\max} \text{ cis } n-\pi^* - \lambda_{\max} \text{ n}-\pi^* \text{ trans})$ (nm)	Half-life +/- Std Dev (minutes)	Temperature (°C)
<b>191</b> (H <sub>4</sub> )	1.08	323	445	0 nm	110 +/- 10	60
<b>185</b> (F <sub>4</sub> )	24.38	317	459	41 nm	1800 +/- 30	60
<b>171</b> (Cl <sub>4</sub> )	52.39	299	473	21 nm	3170 +/- 40	60
<b>179</b> (Br <sub>4</sub> )	50.49	297	467	18 nm	3300 +/- 90	60

Some of these properties can be explained by the dihedral angle of the azo bond; in the case of the tetra-*ortho*-fluoro the dihedral angle is much smaller allowing better electronic overlap between the two rings. This overlap is broken as the rings rotate to minimise steric clashes in the *cis* isomer, resulting in a larger difference in conjugation, and hence UV/visible spectra between the two isomers. This effect is even more clearly illustrated by the angles between the planes of the benzene rings.

**Figure 152.** Crystal structures of **185**, **171** and **179** showing the angles of the aromatic rings.

Looking down the axis of the azo bond shows **185** has a nearly planar structure allowing extended conjugation between the aromatic rings, but **171** and **179** show a twist in the rings in the *trans* states to almost 90° because of relatively larger chlorine and bromine substituents. This extra steric bulk partly accounts for the longer relaxation from the

*cis* isomers by hindering the rotation around the N-N bond in the excited state, although **185** appears relatively unhindered (Figure 153).



**Figure 153.** Space-filling model of the compounds **185**, **171**, and **179** illustrating the increased steric demands of tetra-*ortho*-chloro- and tetra-*ortho*-bromo-substituents.

#### 4.4 Di-*ortho*-haloazobenzenes

Since the change in planarity appears to be an important factor determining the shifts in  $n-\pi^*$  band shifts and hindered rotation caused by steric bulk affect the half-life of a series of di-*ortho*-halogenated azobenzenes were synthesised to determine if they possessed useful band shifts whilst retaining the improved half-lives of tetra-*ortho*-haloazobenzenes.

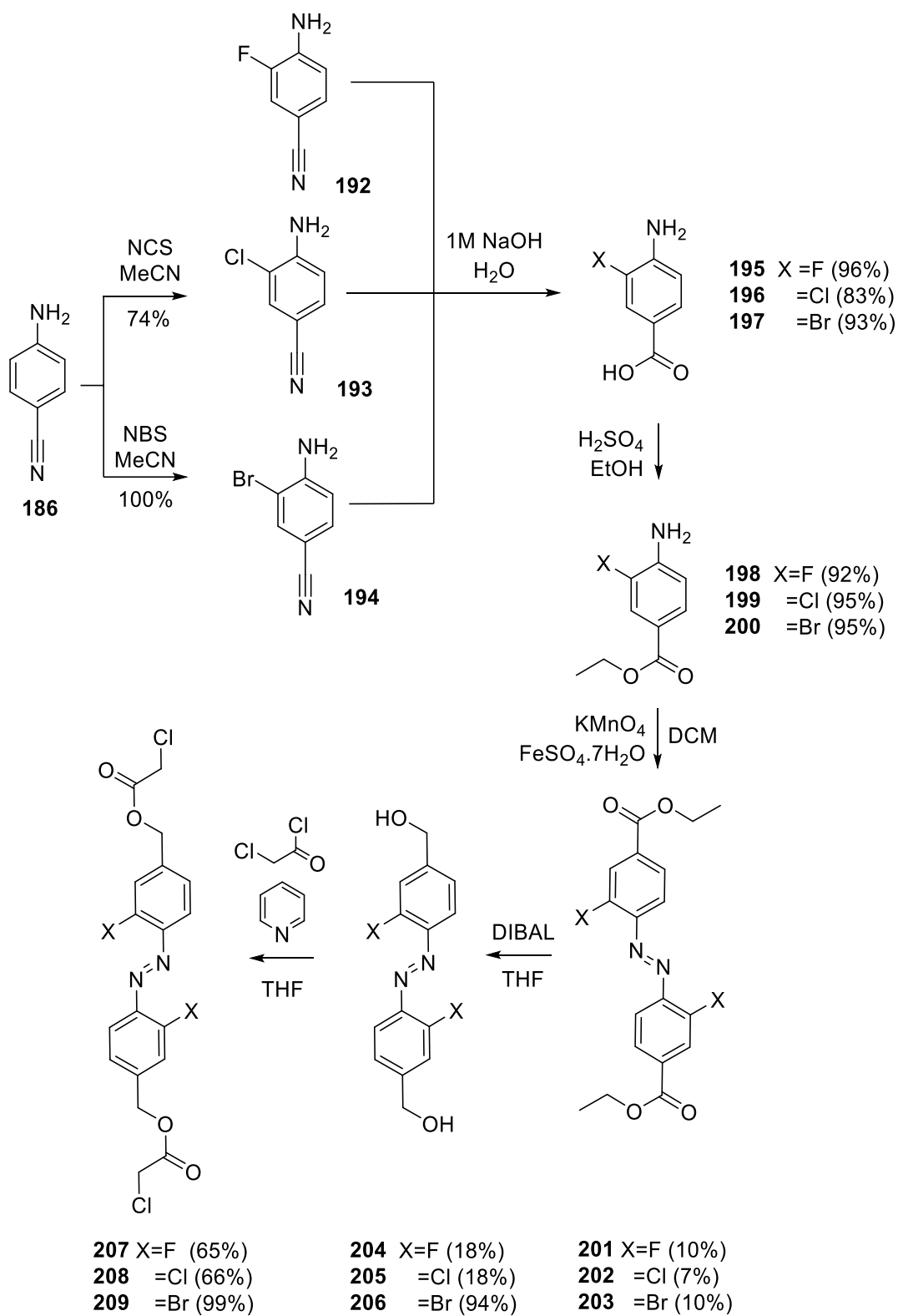


Figure 154. Synthetic scheme for di-*ortho*-haloazobenzenes 207, 208, and 209.

The first step of the syntheses of dichloro and dibromo photoswitches are halogenation of one of the *ortho* positions of 4-aminobenzonitrile. 4-Amino-2-fluorobenzonitrile was purchased from Fluorochem for the difluoro switch synthesis.

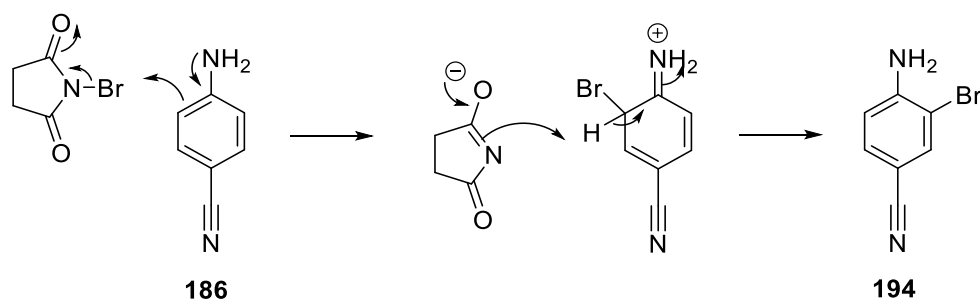
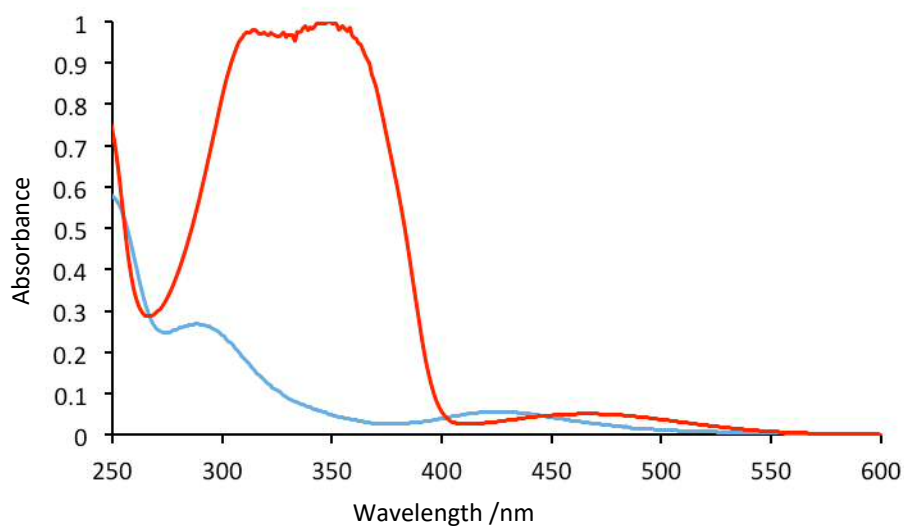


Figure 155. Monobromination of 4-aminobenzonitrile **186**.

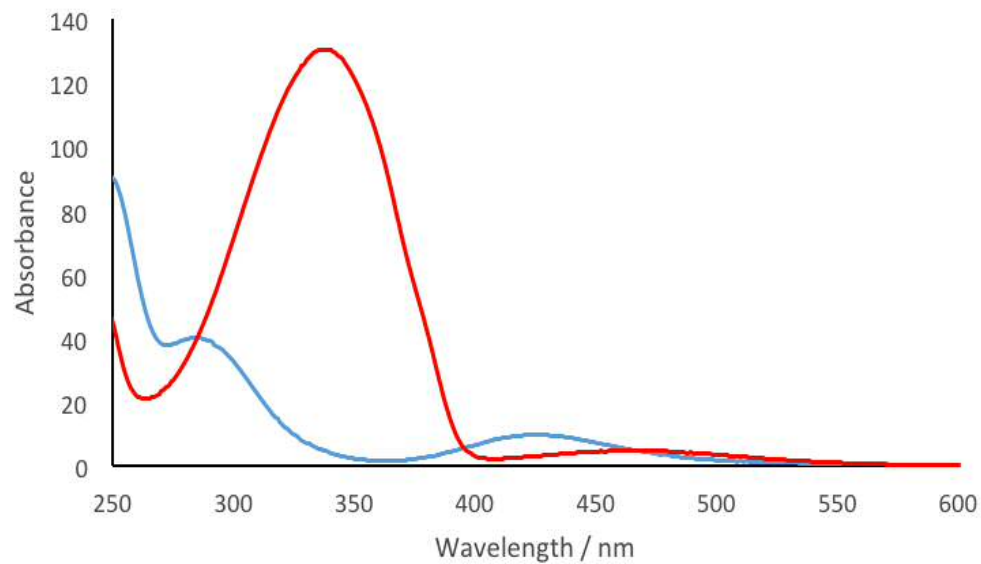
The *N*-halosuccinimide acts as a source of electrophilic bromine and chlorine atoms so that the aromatic ring attacks to add eg. a bromine at the *ortho* position directed by the presence of the electron donating amine (Figure 155). A proton is then extracted by the succinimide to restore aromaticity, leaving the appropriate *ortho* halogen. Only one equivalent of halogenating agent is added and selectivity is achieved as the electron withdrawing effects of the first substitution make the second substitution less favoured than the first. These nitriles are then hydrolysed, esterified, dimerised, reduced and esterified as previously described. The yields of the oxidation reactions to form the azo bond are greatly reduced compared to those of the tetra-halo or tetra-hydro analogues. This is probably due to less effective screening of the azo bond from further oxidation than in the tetra-*ortho*-halo versions. A lack of protection from reduction also reduces



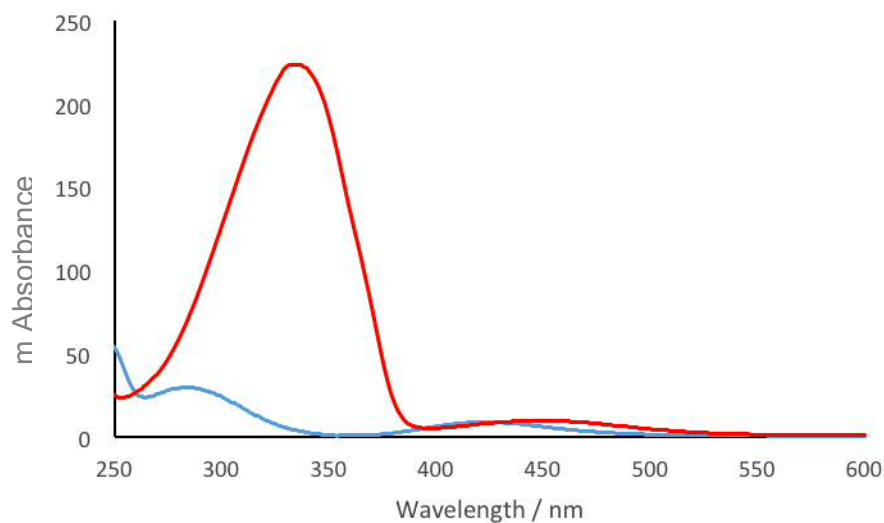
the yield of the DIBAL reductions compared to the corresponding yields of tetra-halo photoswitches.



**Figure 156.** HPLC-DAD UV/visible spectra of *cis*-209 (blue) and *trans*-209 (red).



**Figure 157.** HPLC-DAD UV/visible spectra of *cis*-208 (blue) and *trans*-208 (red).

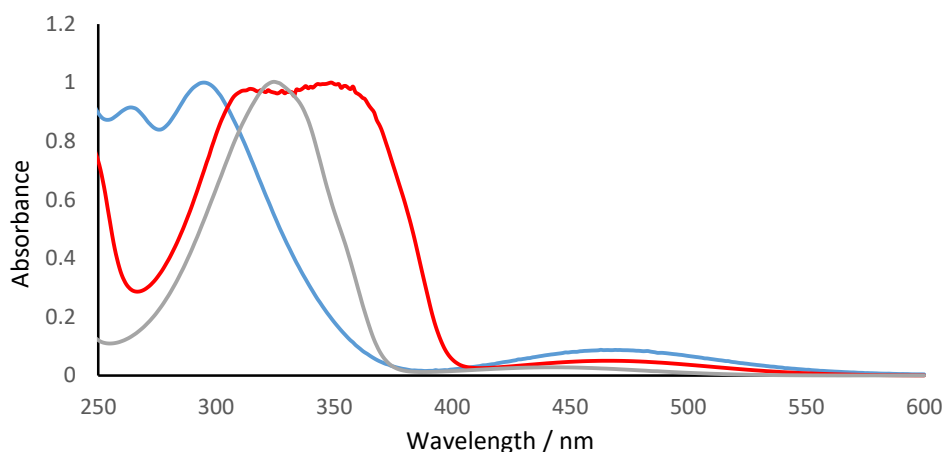


**Figure 158.** HPLC-DAD UV/visible spectra of *cis*-**207** (blue) and *trans*-**207** (red).

The UV/visible spectra of the dihalogenated switches all have similar  $\pi$ - $\pi^*$  absorbance maxima between 330 and 340 nm. Compounds **208** and **209** show large separations (42 nm and 43 nm respectively) between  $n$ - $\pi^*$  absorbance maxima of their *trans* and *cis* isomers which means the photostationary state upon irradiation of light > 530 nm favours formation of the *cis* isomer.

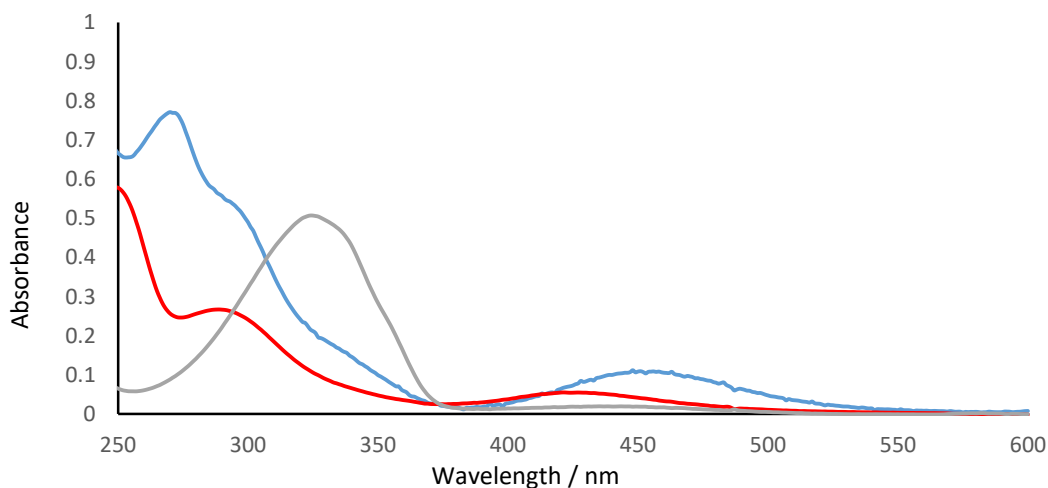
**Table 22** Optical and physical properties of the di-*ortho*-halo azobenzenes.

	Dihedral angle	$\pi$ - $\pi^*$	$n$ - $\pi^*$	$\Delta(\lambda_{\max})$	Half-life +/- Std Dev (minutes)	Temperature (°C)
		$\lambda_{\max}$ (nm) trans	$\lambda_{\max}$ (nm) trans	$\lambda_{\max}$ $n$ - $\pi^*$ <i>trans</i> (nm)		
<b>207</b> (F <sub>2</sub> )	n/d	339	450	25 nm	460 +/- 30	60
<b>208</b> (Cl <sub>2</sub> )	1.10	333	468	42 nm	210 +/- 10	60
<b>209</b> (Br <sub>2</sub> )	n/d	340	468	43 nm	230 +/- 10	60



**Figure 159.** UV/visible spectra of *trans*-**191** (unsubstituted, grey) *trans*-**209** (dibromo, red) and *trans*-**179** (tetrabromo, blue).

**191** shows a typical UV/visible absorbance spectra with an  $\pi$ - $\pi^*$  transition absorbance at 340 nm. Di-*ortho*-bromo substitution in **209** caused a red shift of ~25 nm to ~375 nm whereas the tetra-*ortho* substituted compound **179** showed a blue shift to 300 nm. This could be due to compound **209** being planar but the additional bulk once the other two *ortho* positions are halogenated **179** mean it is no longer planar, leading to a large shift of the  $\pi$ - $\pi^*$  band to 300 nm.



**Figure 160.** UV/visible spectra of *cis*-**191** (unsubstituted, grey) *cis*-**209** (dibromo, red) and *cis*-**179** (tetrabromo, blue).

The UV/visible spectra of the *cis* isomers show a large change in the position of the  $n$ - $\pi^*$  absorbance band to longer wavelengths with successive substitution.

**Table 23.** Comparison between non halogenated di bromide and tetra bromide substituted azobenzenes.

	Dihedral angle	$\pi-\pi^*$ $\lambda_{\max}$ (nm) trans	$n-\pi^*$ $\lambda_{\max}$ (nm) trans	$\Delta(\lambda_{\max}$ <i>cis</i> $n-\pi^*$ - $\lambda_{\max}$ $n-\pi^*$ <i>trans</i> ) (nm)	Half-life +/- Std Dev (minutes)	Temperature (°C)
<b>191</b> (H <sub>4</sub> )	1.08	323	445	0 nm	110 +/- 10	60
<b>209</b> (Br <sub>2</sub> H <sub>2</sub> )	1.10	333	468	42 nm	210 +/- 10	60
<b>179</b> (Br <sub>4</sub> )	50.49	297	467	18 nm	330 +/- 90	60

Comparing the non-halogenated **191** to di-*ortho*-bromo **209** and tetra-*ortho*-bromoazobenzene **179** it is evident that increased steric bulk of *ortho* bromo substituents increases the stability of the *cis* isomer, but also that the extra bromide substituents in **179** compared to **209** reduced the  $n-\pi^*$  band separation lowering the proportion of *cis*-**209** present in the photo stationary state and the increasing the time taken to reach it.

The change in N=N dihedral angle again correlates with this increase in the band separation. In the sterically constrained species where the bulk of the large aromatics prevents a planar orientation that would allow orbital overlap the ground state energy is increased. These planar di-*ortho*-halo species have a half-life that is one tenth of that of the tetra-*ortho*-chloro compound **171** and tetra bromo compound **179**, but the di-*ortho*-fluoro has a longer life than the dichloro or dibromo equivalents, implying electronic effects must also play a role in addition to steric effects.

#### 4.5 Mixed halogens

In order to keep the planar structure which causes the greatest splitting between the absorption maxima of the  $n-\pi^*$  band in *cis* and *trans* isomers and obtain the increased half-lives characteristic of tetra-*ortho*-substituted azobenzenes, mixed dichlorodifluoro and dibromodifluoro halogen azobenzenes were synthesised.

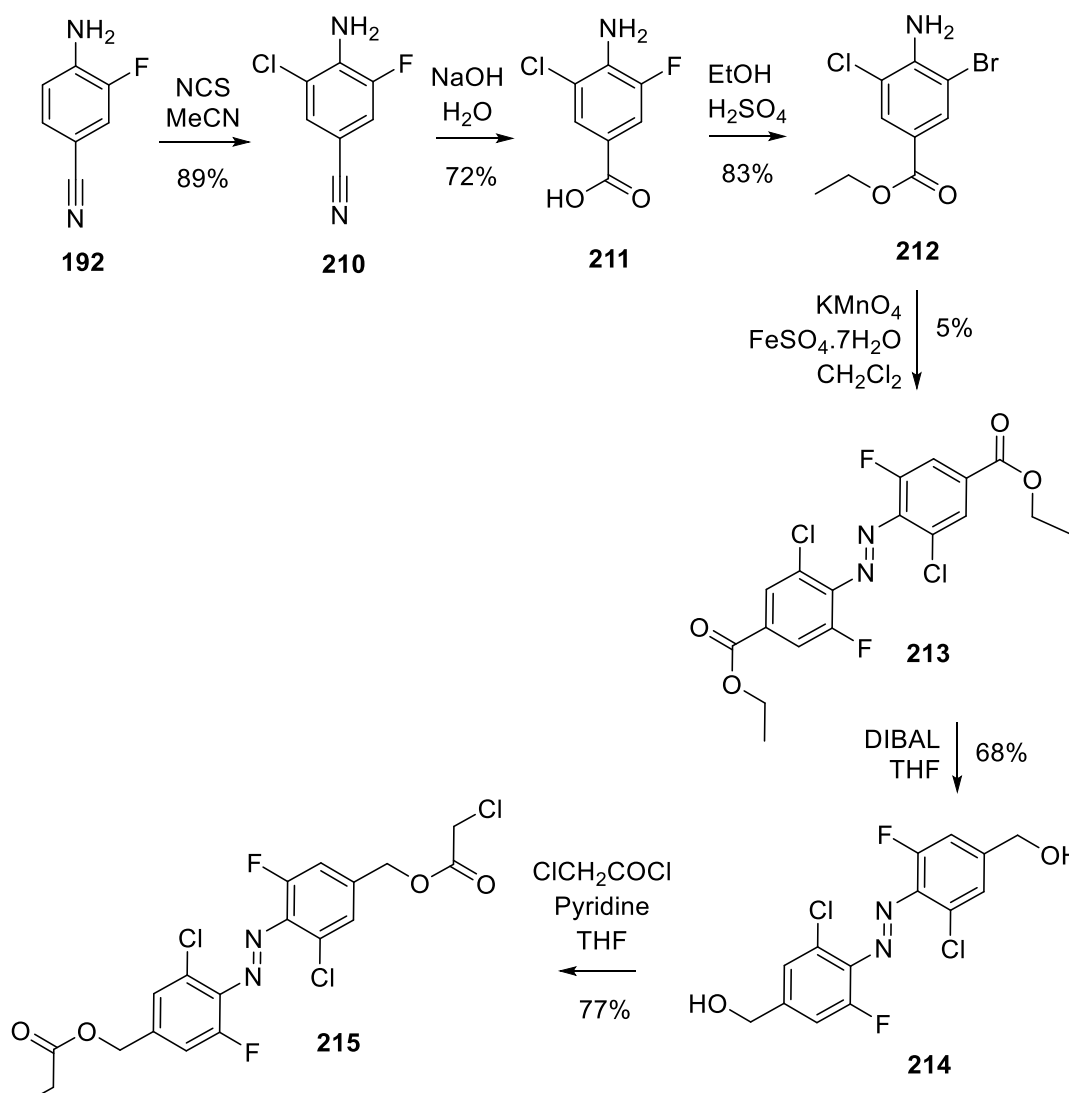


Figure 161. Synthesis of dichlorodifluoroazobenzene 215.

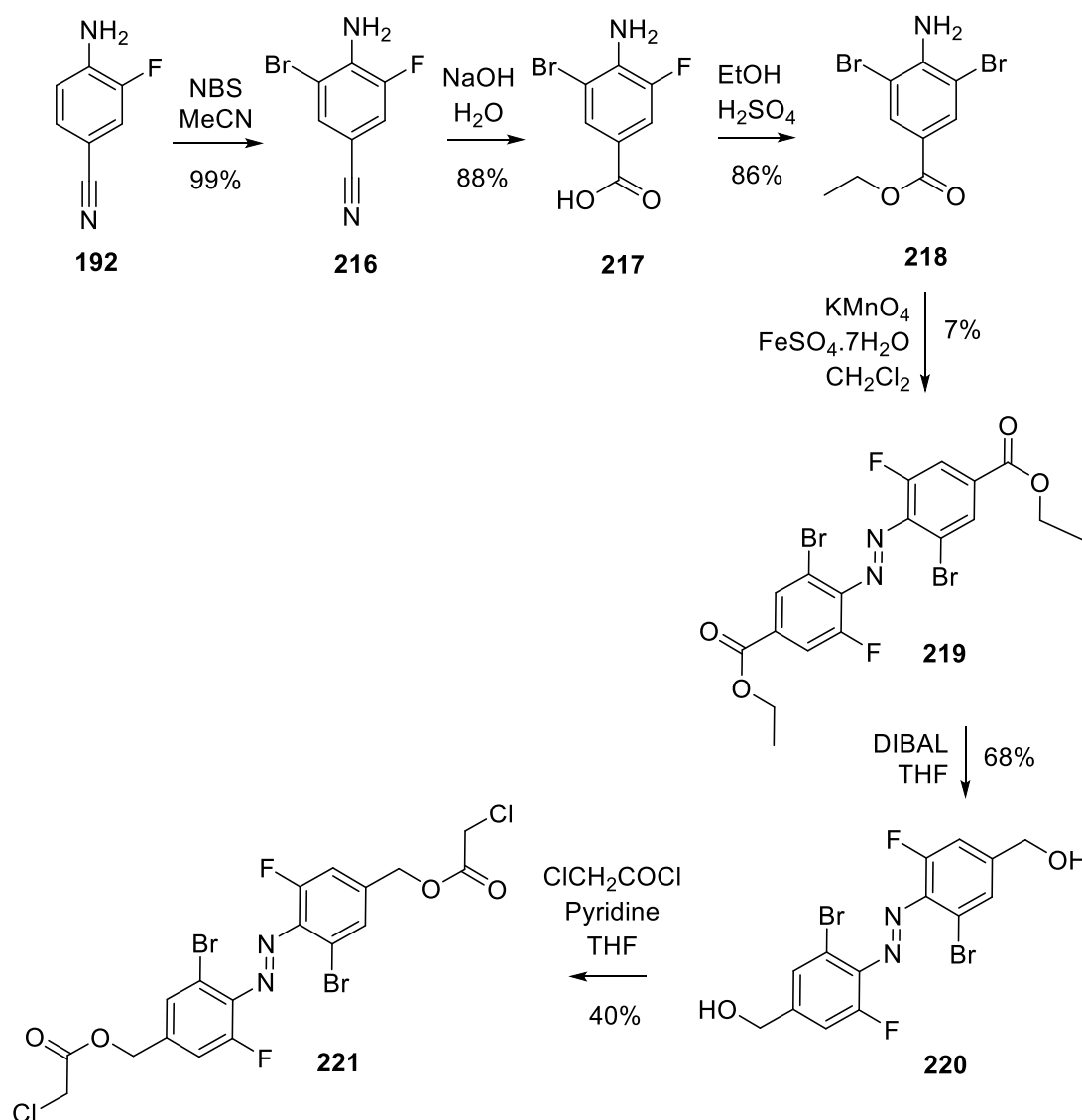
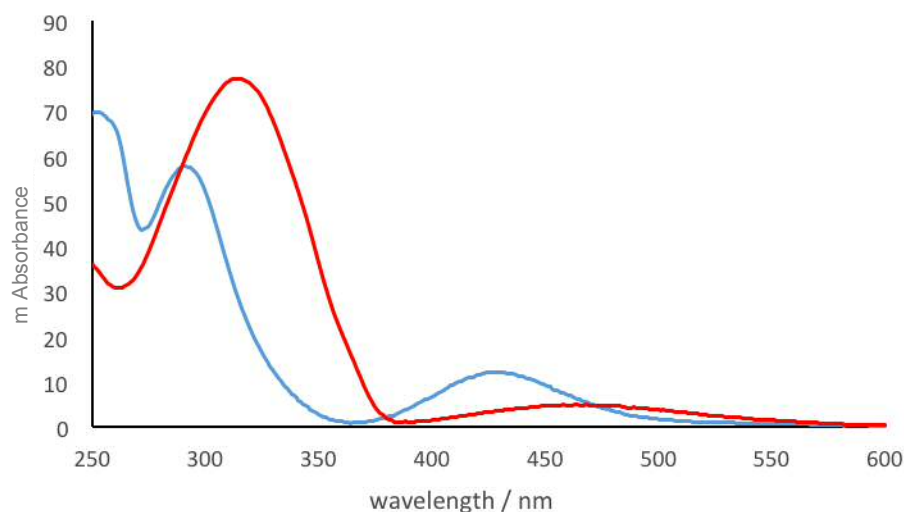
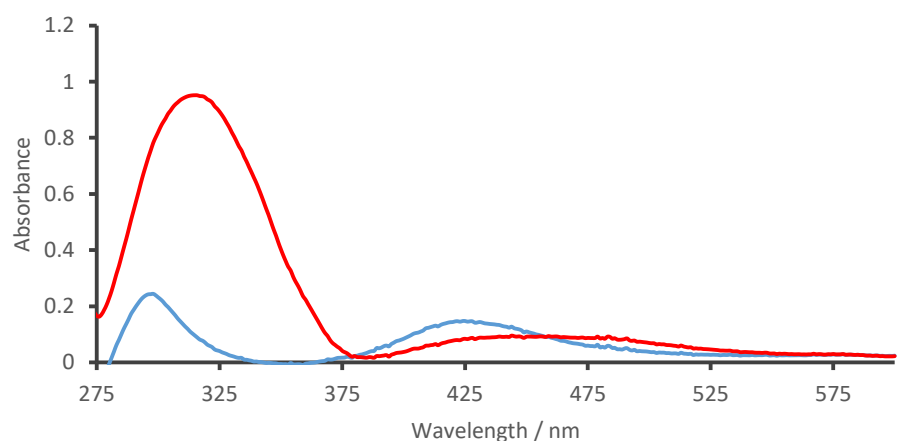


Figure 162. Synthesis of dibromodifluoroazobenzene **221**.

The syntheses of the mixed halogenated azobenzenes **215** and **221** are similar to that of the dihalogenated azobenzenes, albeit with lower yielding azo bond formation; this is due to increased ease of oxidation of either the aniline starting material or the azobenzene bond once formed. Despite the low yield, enough material was produced to test the properties of these compounds.

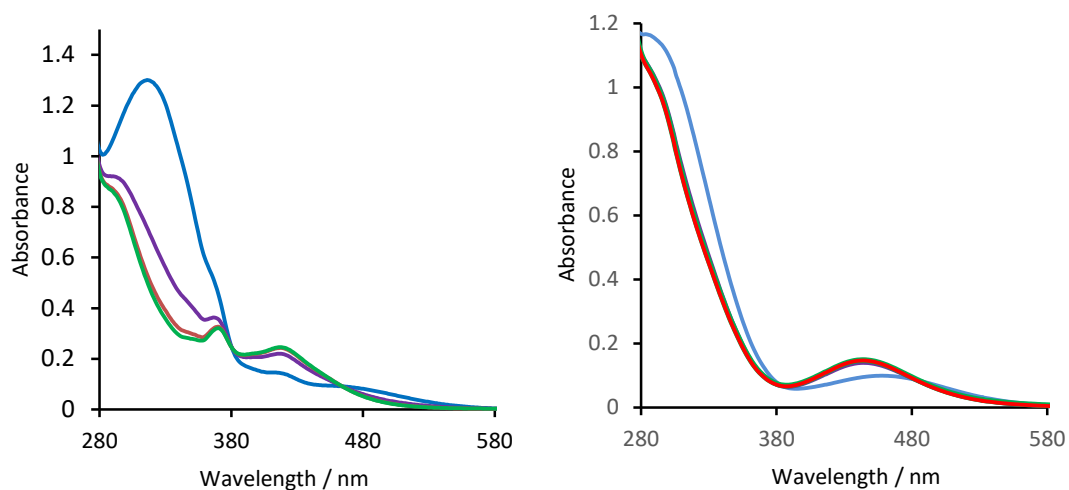


**Figure 163.** UV/visible absorbance spectra of *cis*-215 (blue) and *trans*-215 (red).



**Figure 164.** UV/visible absorbance spectra of *cis*-221 (blue) and *trans*-221 (red).

The UV/visible absorbance spectra of **215** and compound **221** show the largest band shifts, 43 and 47 nm respectively, of any molecules in this series. This large band separation and less hindered azo bonds of these hybrid switches allow much more efficient switching from *trans*-**215** to the *cis*-**215**, with irradiation with >530 nm light producing a photostationary state with approximately 65 % *cis* based on the loss of absorbance at 340 nm which is entirely due to *trans* state.



**Figure 165.** Comparison of photo switching of dichloro,difluoro **215** (left) and tetrachloro **171** (right) photoswitches after irradiation with light through a >530 nm filter for 0 (blue), 5 (purple), 10 (red) and 15 (green) minutes.

**Table 24.** Photoswitching of **215**.

Irradiation time > 530 nm (minutes)	Abs <sub>340 nm</sub>	Percentage switching <sup>a</sup>
0	0.414	0
5	0.208	50
10	0.156	62
15	0.149	64

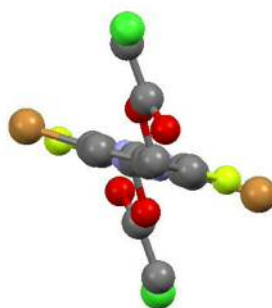
<sup>a</sup> Based on 340 nm absorbance.

These molecules show fast and efficient switching upon irradiation with >530 nm light in marked contrast with, for example, the tetra-*ortho*-chloro switch **171** whose conversion between *trans* and *cis* isomers is too small to be reliably established.



**Table 25.** Spectroscopic and half-life data for the mixed halogenated azobenzenes.

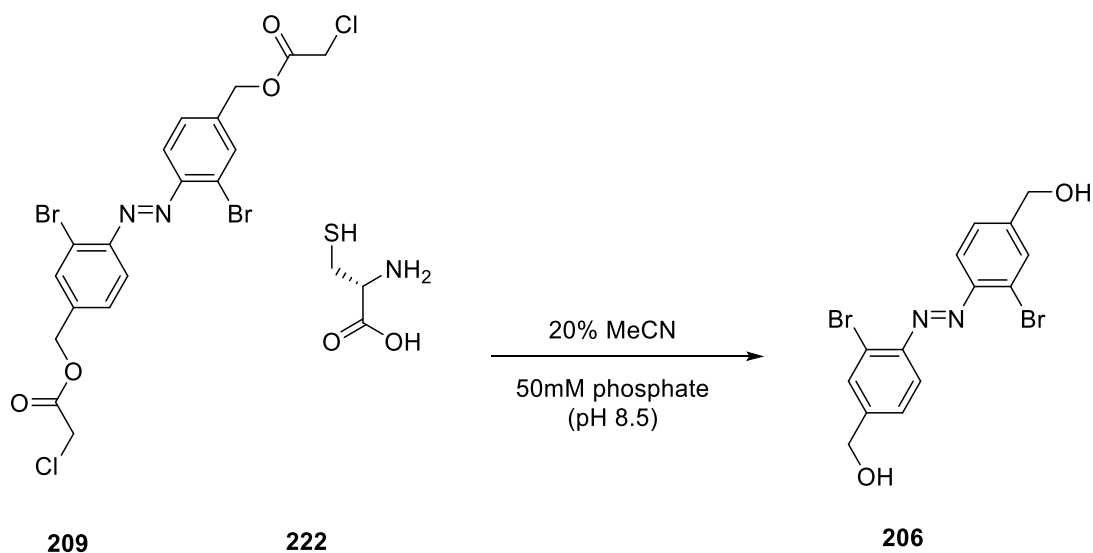
	Dihedral angle	$\pi$ - $\pi^*$ $\lambda_{\max}$ (nm) trans	$n$ - $\pi^*$ $\lambda_{\max}$ (nm) trans	$\Delta(\lambda_{\max} \text{ cis } n$ - $\pi^* - \lambda_{\max} \text{ n-}\pi^*$ $\text{trans})$ (nm)	Half-life +/- Std Dev (minutes)	Temperature (°C)
<b>215</b> (Cl <sub>2</sub> F <sub>2</sub> )	0	320	469	44	3000 +/- 220	60
<b>221</b> (Br <sub>2</sub> F <sub>2</sub> )	n/d	318	475	47	2300 +/-70	60

**Figure 166.** Crystal structure of *trans*-**215** showing the coplanarity of the aromatic rings.

The crystal structure of *trans*-**215** shows the desired planer structure. Not only do these mixed halogen photoswitches **215** and **221** give rise to the largest  $n$ - $\pi^*$  band separations, approaching 50 nm, they also have long half-lives of approximately 48 hours even at 60 °C which makes them viable for *in vitro* and *in vivo* experiments. Both achieve photostationary states where greater than 65 % of the photoswitch is in the *cis* state under > 530 nm irradiation.

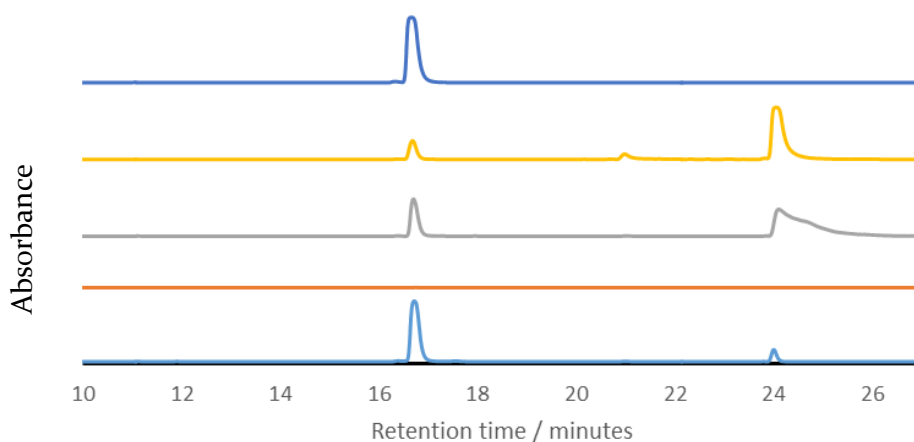
#### 4.6 Attaching to a peptide

Unfortunately, these crosslinkers are poorly soluble in water so solvent mixtures were screened in order to find the best conditions for the cross linking reaction. Crosslinkers were reacted with two equivalents of cysteine in phosphate buffer (50 mM) at pH 8.5.

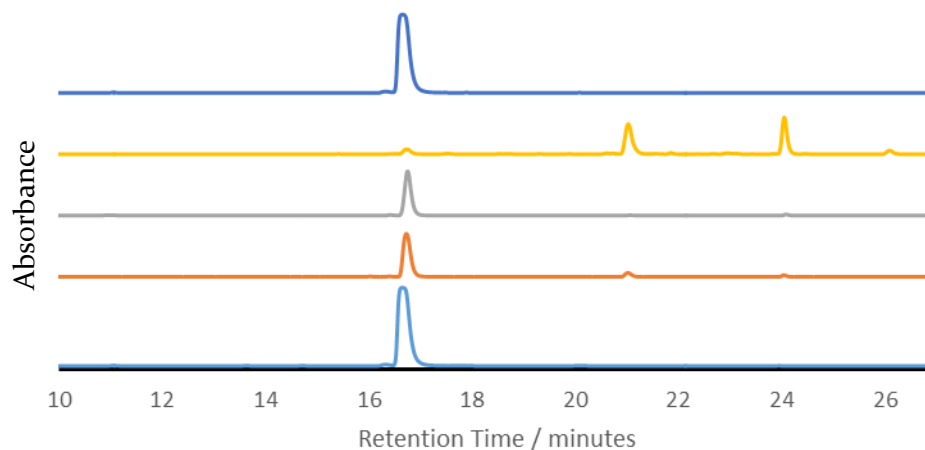


**Figure 167.** Attempted crosslinking of compound **209** with cysteine.

A range of water miscible cosolvents (dimethylsulfoxides, methanol, tetrahydrofuran and acetonitrile) and two temperatures (40 and 60 °C) were screened. Most conditions initially appeared to harbour reactions at various rates giving rise to an intermediate peak at 21 minutes by HPLC assumed to be the single cysteine adduct and the peak at 16.6 minutes representing the cross linker attached to two cysteines (Figure 168).

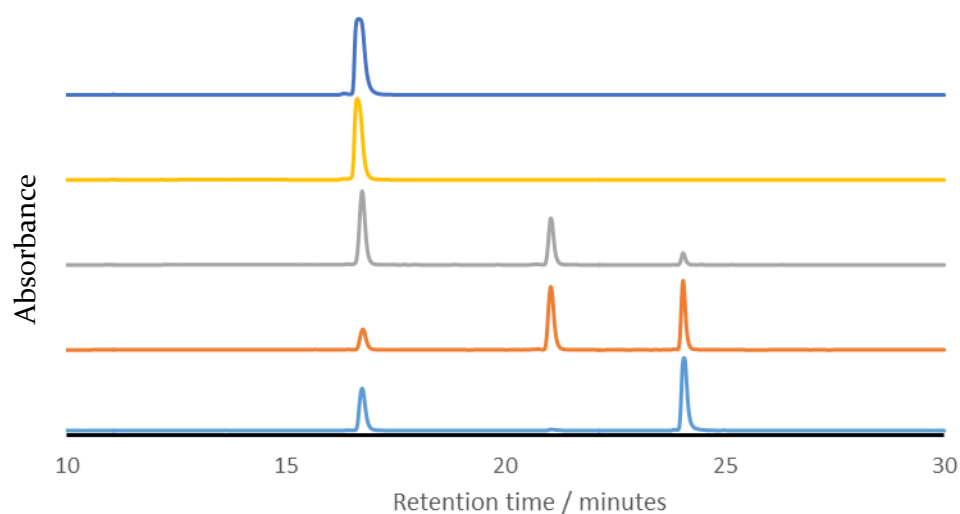


**Figure 168.** Screening of crosslinking conditions reacting overnight at 40 °C. benzyl alcohol (blue), THF (yellow), methanol (grey), acetonitrile (orange), DMSO (light blue)



**Figure 169.** Screening of crosslinking conditions reacting overnight at 60 °C. benzyl alcohol (blue), THF (yellow), methanol (grey), acetonitrile (orange), DMSO (light blue)

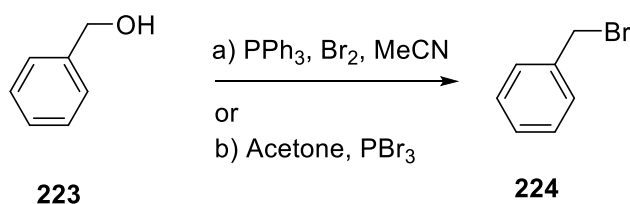
Unfortunately, mass spectrometry and coinjection of reaction mixtures and the benzyl alcohol precursor **206** showed that the products were actually due to ester hydrolysis reforming **206**. The reaction was then repeated in the buffer with dimethylformamide as a cosolvent but without cysteine and injections made to the HPLC at varying reaction times which also showed degradation of the photoswitch **209** to the precursor compound **206**.



**Figure 170.** HPLC traces of the degradation of compound **206** to benzyl alcohol compound **206**: Initial (light blue), 1 hour (orange), 2 hours (grey), overnight (yellow), and benzyl alcohol (blue).

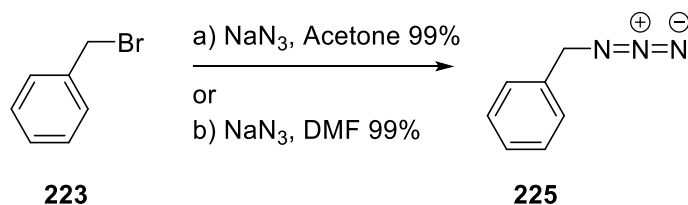
#### 4.7 Preventing hydrolysis

In order to prevent hydrolysis the ester bond needed to be replaced with either an amide or a carbon-nitrogen bond. This was achieved in a test reaction by converting benzyl alcohol **223** to benzylbromide **224** (Figure 171).



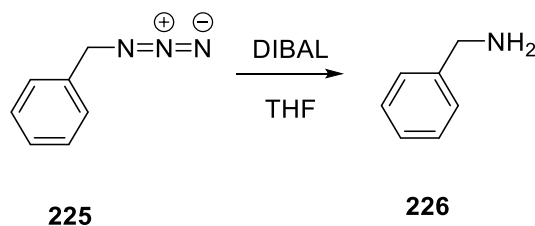
**Figure 171.** Bromination of benzyl alcohol.

Both bromination reaction conditions worked well, with the phosphorous tribromide reaction conveniently requiring no purification. Yields are not shown in Figure 171 because the benzyl bromide is volatile and when a vacuum line was used to remove remaining solvent the product was also removed. Reaction conditions to introduce an azide group by nucleophilic substitution were then tested (Figure 172).



**Figure 172.** Nucleophilic aromatic substitution of benzyl bromide with sodium azide.

Again, both reactions worked well giving quantitative conversion to the azide. However, it was more difficult to remove traces of dimethylformamide from the product, making the acetone reaction superior.



**Figure 173.** DIBAL reduction of benzyl azide to benzylamine.

The reduction of benzyl azide was then attempted using diisobutylammonium hydride as described by Xie *et al*,<sup>122</sup> but the product proved problematic to isolate or unequivocally detect. The work up used a tartrate solution to release the aluminium species from the amine and although this layer was extensively extracted with organic solvents and even freeze dried, no **226** was isolated.

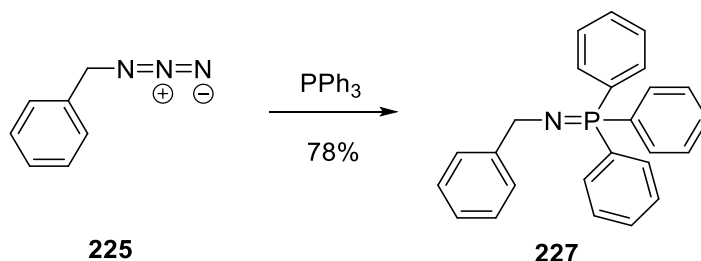


Figure 174. Imminophosphane formation from benzyl azide.

Azide **225** was instead reacted with triphenyl phosphine to give a Staudinger reduction type imminophosphane intermediate.  $^{31}\text{P}$  NMR showed the expected shifts for product formation, but mass spectrometry was inconclusive as rapid formation of benzylamine and phosphine oxide in the presence of water complicated analysis. This property was therefore harnessed for *in situ* amide formation, obviating the need to isolate the difficult amine intermediate and furnishing a hydrolysis-resistant linker.

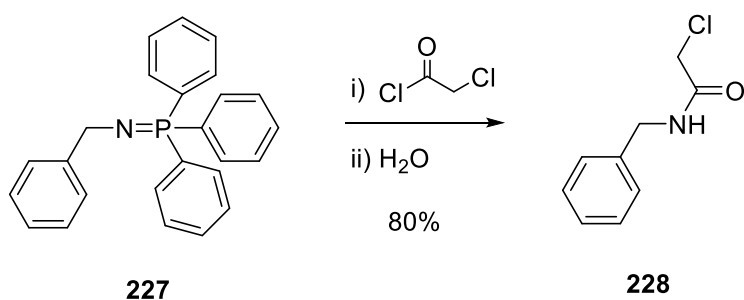
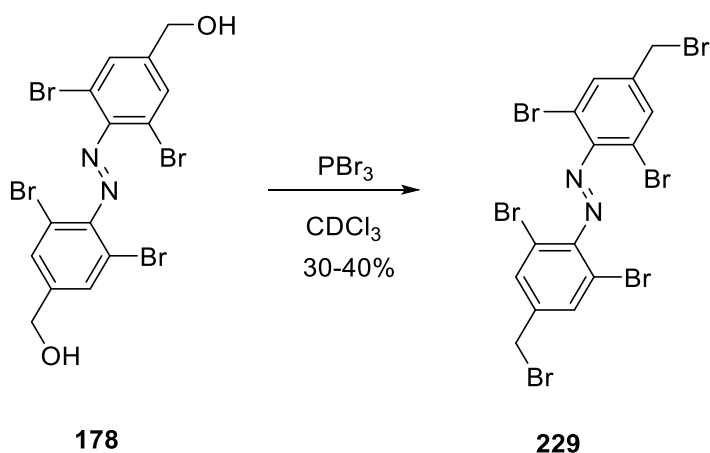


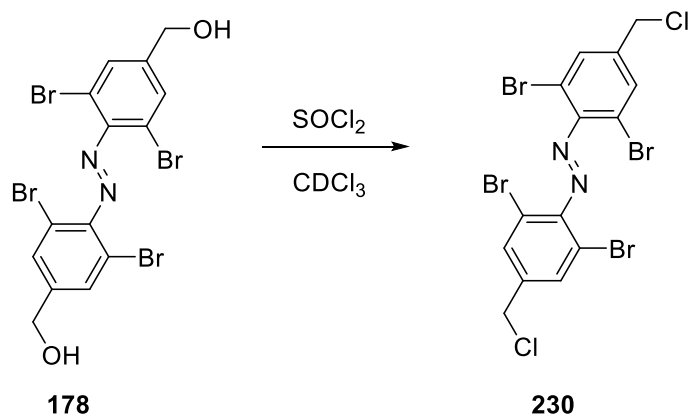
Figure 175. Chloroacetylation of compound **227**.

The tetra-*ortho*-bromo azobenzene alcohol **178** was brominated as shown in Figure 176.



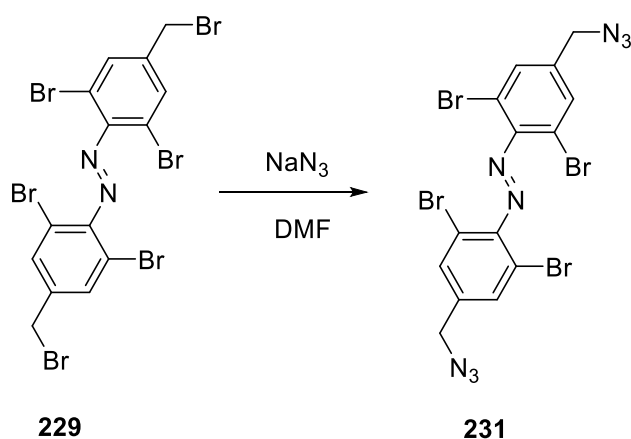
**Figure 176.** Bromination of to form hexabromide **229**.<sup>123</sup>

The isolated yield reaction of this reaction was disappointing (30-40%) despite the reaction proceeding to completion by NMR to form the hexabromide **229**. The acidic conditions during silica gel chromatography were found to result in hydrolysis of the benzylic bromide and yields near 90% were obtained using crude product without further purification. Since hexabromide **229** could act as a crosslinker in its own right<sup>80</sup> a test reaction was carried out with peptide in dimethylsulfoxide but unfortunately only crosslinker hydrolysis products were detected.



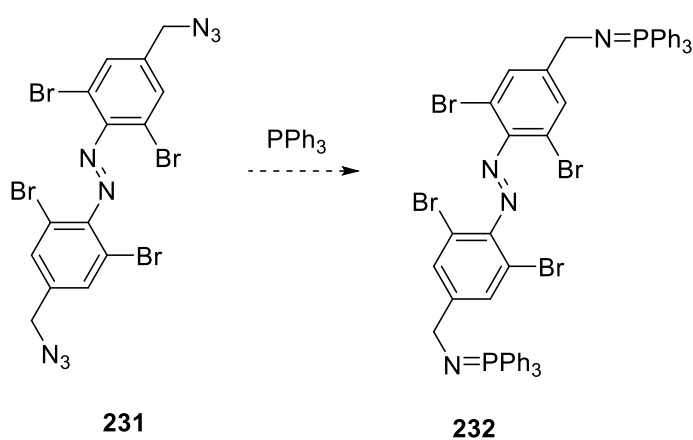
**Figure 177.** Chlorination of **178** to tetrabromodichloride **230**.

This reaction was also attempted with the benzyl chloride equivalent **230** with unfortunately identical results.



**Figure 178.** Formation of benzylic azide **229** from hexabromide **231**.

When the displacement reaction was carried out with five equivalents of sodium azide it led to a red solid that was not soluble in any organic solvent. The reaction was therefore carried out with only two equivalents of azide, but again an insoluble red solid was formed. The IR spectrum of this solid showed the characteristic presence of a azide peak, so the solid was heated with triphenyl phosphine in dimethylformamide to attempt to form an imminophosphrane, but no reaction was detected by  $^{31}\text{P}$  NMR.



**Figure 179.** Reduction of crude azide displacement product postulated to contain **231** to imminophosphorane **232**.

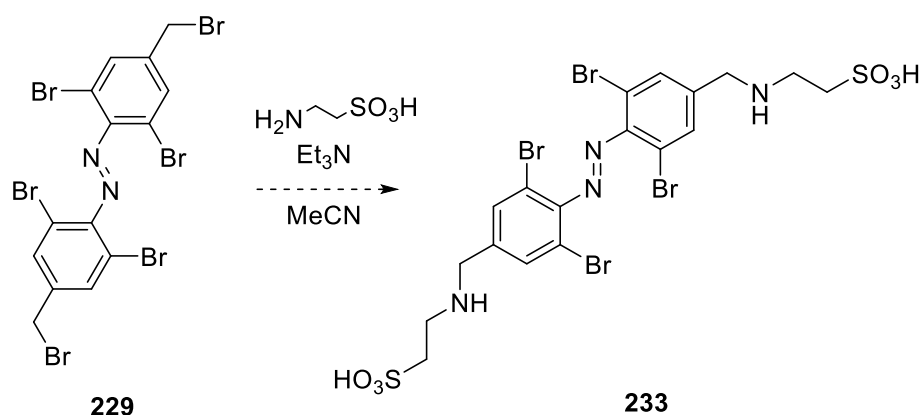


Figure 180. Formation of compound 233.

In light of these results, hexabromide **229** was instead reacted with a polar amine nucleophile, taurine, in an attempt to create a secondary amine suitable for reaction with chloroacetyl chloride whilst providing the water-soluble properties of the sulfonic acid. Unfortunately it proved difficult to find a solvent that dissolved both reactants as taurine is insoluble in acetonitrile and neither taurine or **229** are soluble in tetrahydrofuran. Attempted reaction in dimethylformamide also failed.

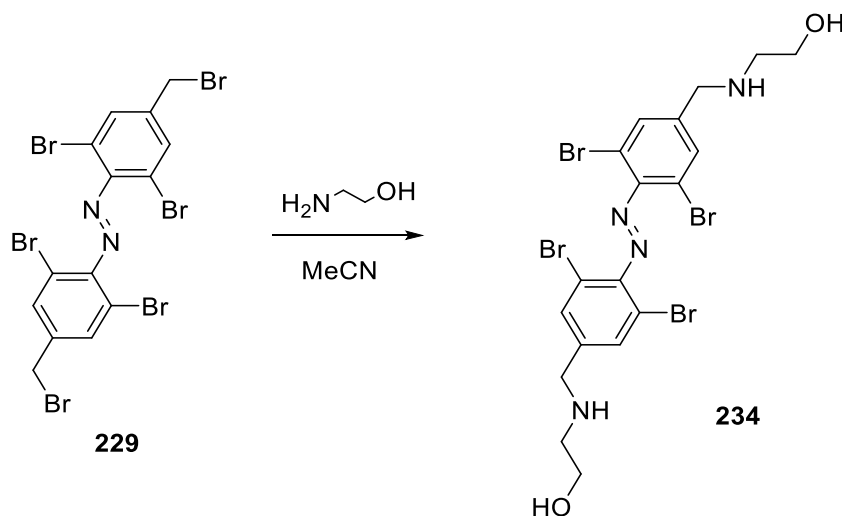


Figure 181. Formation of compound 234.

Switching taurine for ethanolamine was more productive, giving compound **234** in quantitative yields and requiring only filtration through a plug of silica to remove excess ethanolamine. This product was reacted with chloroacetyl chloride to yield **235**. Only two equivalents of chloroacetyl chloride were used in order to prevent ester formation with the primary alcohol (Figure 182).



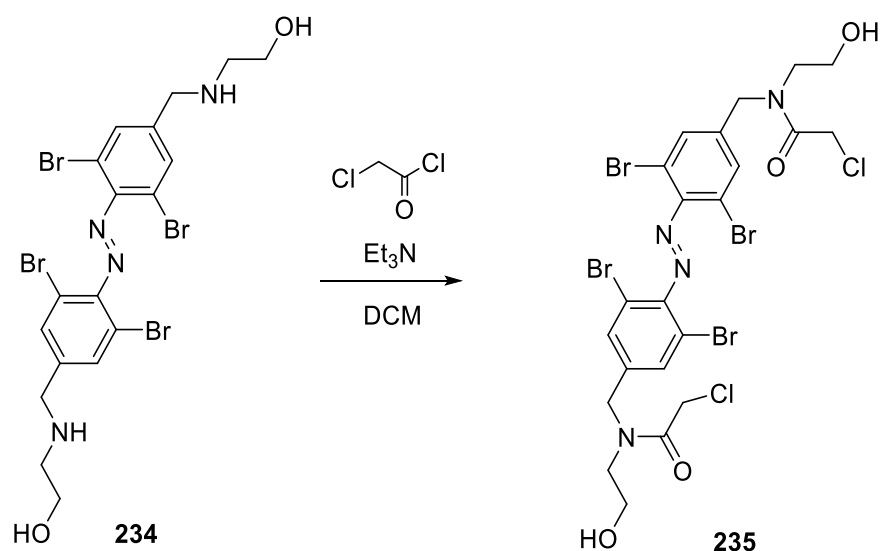


Figure 182. Formation of chloroacetamido azobenzene 235.

The use of a secondary amine in this fashion gives great scope for the introduction of other groups that can be used to orthogonally incorporate molecules containing, e.g. groups for click chemistry or for the attachment of cell penetrating tags. Some examples of potential crosslinkers are shown below (Figure 183).

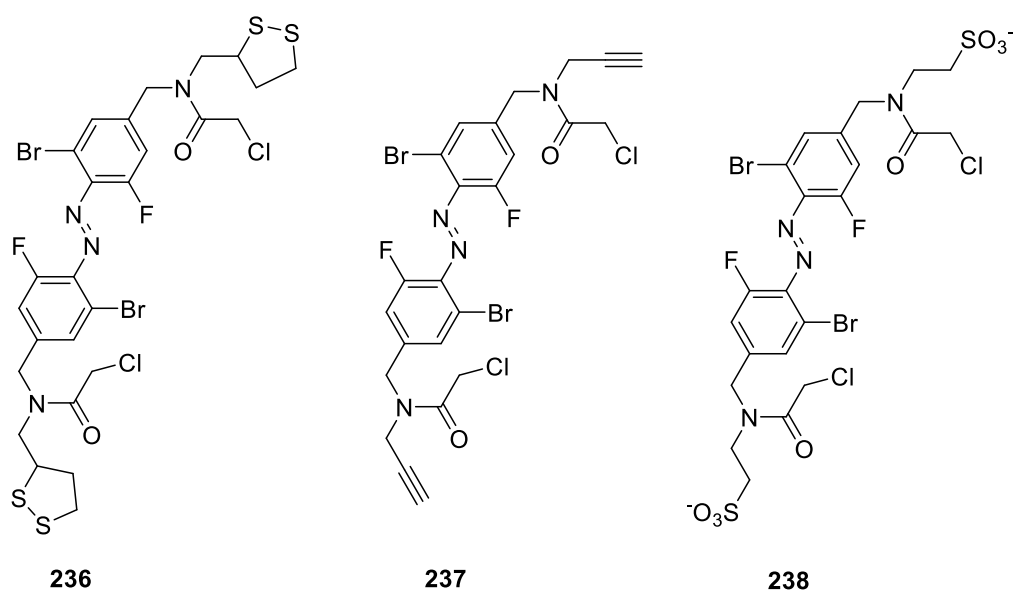
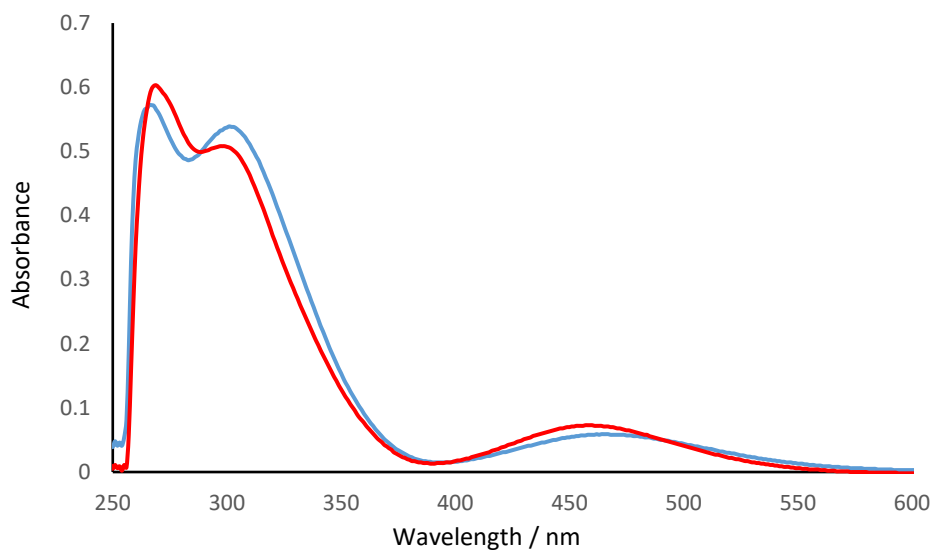


Figure 183. Possible azobenzene products from introducing functionalised primary amines.



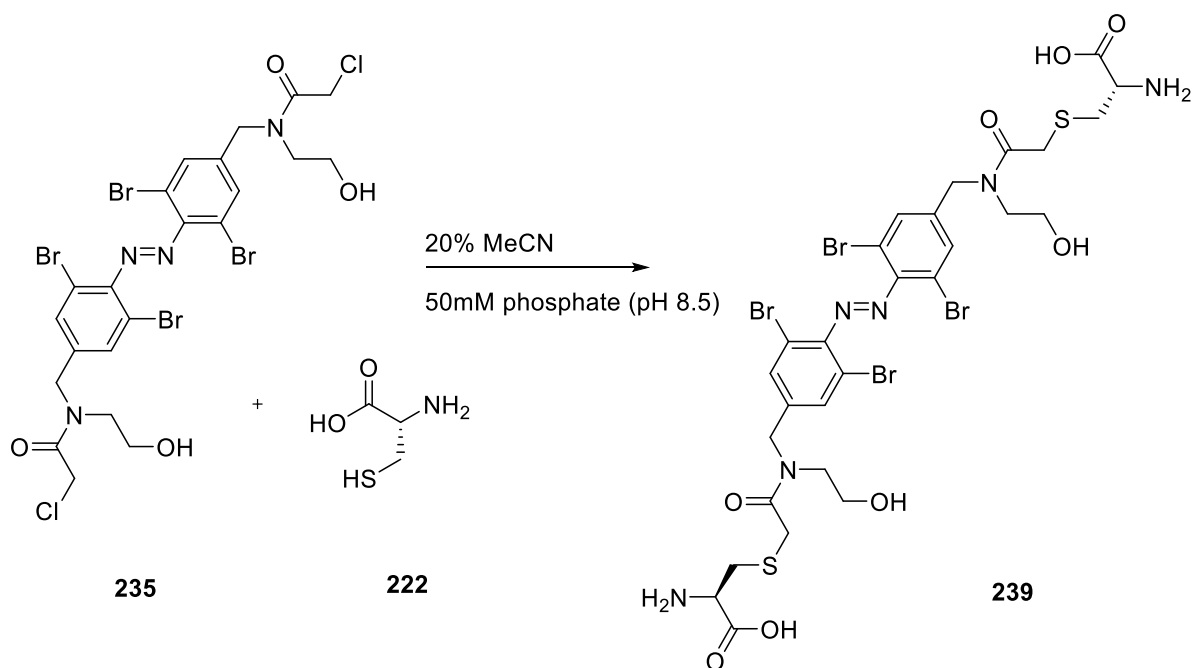
**Figure 184.** UV traces of compound of *trans* (blue) and *cis* (red) **235**.

The UV/visible absorbance spectra of **235** show a reduced band splitting between the *cis* and *trans*  $n-\pi^*$  bands as well as a drastically reduced half-life in comparison to the tetra-*ortho*-bromo **179**. This could be caused by a number of factors, including the hydroxyl group stabilizing or destabilizing transition states with the formation of iminium ion formation although this is not directly conjugated to the aromatic ring.

**Table 26.** Half-lives of *cis*-**235** at different temperatures.

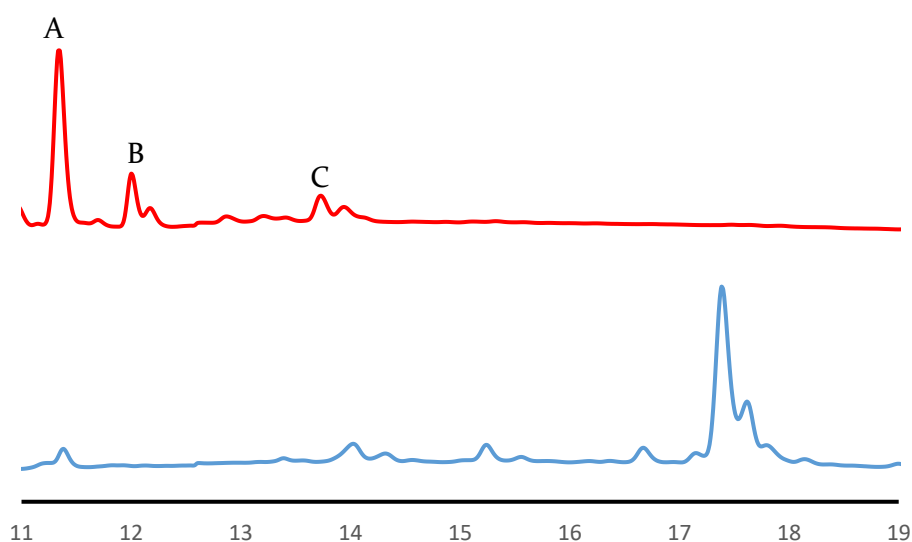
Half-life +/- Std Dev (minutes)	Temperature (°C)
438.7 +/- 49.2	20
41.6 +/- 0.5	40
6.4 +/- 0.4	60

## 4.8 Attachment to cysteine



**Figure 185.** Test of crosslinking reaction conditions in the presence of cysteine.

Chloroacetamide **235** was reacted with cysteine in 50 mM phosphate (pH 8.5) buffer with 20% acetonitrile as a cosolvent as described by Sudhir *et al.*<sup>124</sup> and monitored by HPLC (Figure 186). Complete conversion to the bis-cysteine was observed with no hydrolysis products observed.



**Figure 186.** HPLC UV traces of crosslinked cysteine (red) in both the *cis* and *trans* states (peaks A and B) with a small amount of single cysteine attached (peak C) and compound **235** (blue) for comparison.

Both major peaks A and B were confirmed by high resolution mass spectrometry to contain both the cross linker and cysteine. UV/visible spectra showed peaks A and B to be *cis-235* and *trans-235* respectively. This methodology could now be applied to crosslink peptides.

# Chapter 5

## Conclusions

## 5 Conclusions

Experiments with peptides derived from zinc fingers alkylated with BBSCA (**71**) showed that the conformation and binding affinity of the a peptide derived from the key DNA-binding  $\alpha$ -helix of the zinc finger transcription factor slug can be controlled with light. The binding affinity of the Slug Finger 1-**71** for its target DNA sequence was affected by the change of conformation caused by isomerisation of the photoswitch. The success of Slug Finger 1-**71** shows the feasibility of a minimalist design, but Slug Finger 2-**71** and Slug Double Finger-**71,71** showed that this approach has limitations. Whilst the lack of control of DNA-binding affinity for Slug Finger 2-**71** may be a result of the peptide design, the incorporation of multiple switches in Slug Double Finger-**71,71** highlights fundamental shortcomings in using **71**. The half-life of *cis*-**71** is simply too short for this application and combined with the requirement for damaging UV light, this motivated the investigation of other photosensitive molecules for the control of protein conformation. Initial experiments with thioindigo were aborted when the synthesis of a water-soluble version proved problematic. Inability to detect the formation of a cyclisation product under any of the conditions used suggests that the presence of a sulfonamide group makes the benzene ring too electron deficient for the subsequent ring closure by radical reaction or nucleophilic attack on the acyl chloride from the ring.

The use of *ortho*-substitution to increase the half-lives of the azobenzene switches and generate separation between *cis* and *trans* absorbance bands was investigated. Synthesis of *ortho*-trifluoromethyl substituted azobenzenes allowed comparison of the effects of  $\sigma$ -electron withdrawing trifluoromethyl and  $\sigma$ -withdrawing but  $\pi$ -donating effect of *ortho* halogen substituents. The most apparent change in the UV absorbance spectra is a general blue-shift in the absorbance of the  $\pi$ - $\pi^*$  orbital from 360 nm for the di-*ortho*-trifluoromethyl substituted azobenzenes to ~300 nm for the tetra-*ortho*-halogen substituted azobenzenes. The photoswitching properties of these compounds were examined by recording spectra before and after illumination. The tetra-*ortho*-trifluoromethyl substituted azobenzenes were difficult to switch to the *cis* state and did not show a great improvement in relaxation time over **71**.

The tetra-*ortho*-fluoro switch developed by Hecht *et al.* was used to investigate the effect of incorporation of a heteroatom substituents at the *para* position, but reactions to incorporate both sulfur and oxygen were unsuccessful. However, azobenzenes with the tetra-*ortho*-fluoro substitution pattern showed a greatly increased half-life and band separation between the *cis* and *trans* states.

A tetra-*ortho*-chloroazobenzene photoswitch was synthesised with a *para* hydroxymethyl which was reacted with chloroacetate to provide a means for attachment to a peptide. A series of azobenzenes with different halogen substituents were synthesised and crystal structures showed that the *trans* states occupy different conformations according to the steric bulk of the substituents; in particular the rings of the tetra-*ortho*-chloro and tetra-*ortho*-bromo substituted azobenzenes show markedly different twisted conformations compared to that of the planar tetra-*ortho*-fluoro azobenzene. This suggests that the tetra-*ortho*-bromo and tetra-*ortho*-chloro *cis* half-lives should be extended by constraints upon rotation, but in fact they are comparable to the tetra-*ortho*-fluoroazobenzene. Di-*ortho*-chloro and bromoazobenzenes also adopt a planar conformation as do mixed dichloro/difluoro and dibromo/difluoro species.

All of the tetra-*ortho*-haloazobenzenes possessed improved properties with an increased band separation compared to di-*ortho*-trifluoromethylazobenzenes and orders of magnitude superior *cis* isomer half-lives compared to **71**, along with the ability to switch with red light. However, under the crosslinking conditions the ester-based linker was shown to hydrolyse. To overcome this several modifications were attempted including converting the ester bond to an amide. This modification slightly decreased the relaxation time, but did show extended stability to hydrolysis.

Future work should focus on the di-*ortho*-chloro, di-*ortho*-fluoroazobenzene as it is easier to switch than the tetra-*ortho*-bromoazobenzene and its planar *trans* isomer results in a wider band separation. The di-*ortho*-chloro di-*ortho*-fluoro substitution should be preferred to the di-*ortho*-bromo di-*ortho*-fluoroazobenzene due to the lower reactivity of the ring substituents relative to reactive groups on the benzylic carbon.

# Chapter 6

## Methods



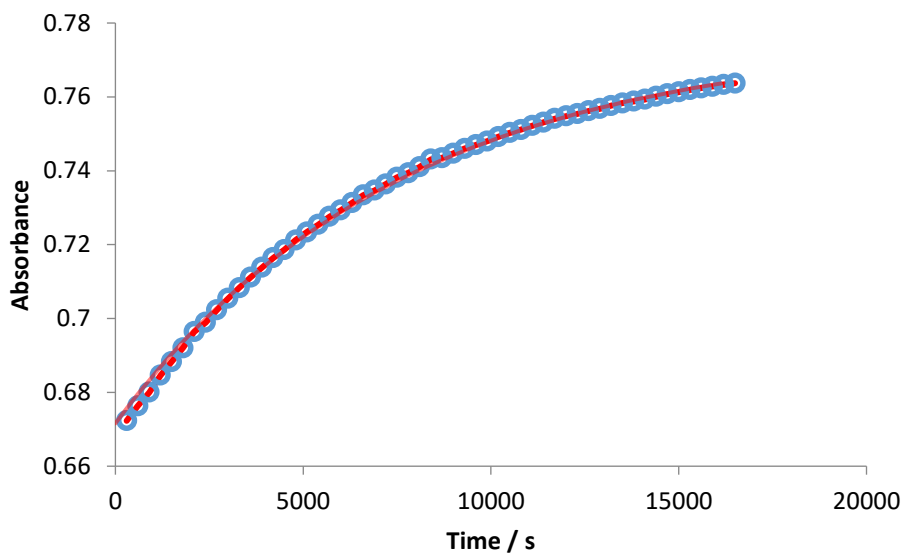
## 6 Methods

### 6.1 General chemical methods

#### 6.1.1 Half-life Calculations

Unless otherwise stated all half-lives were determined in DMSO and calculated in triplicate. Spectra from a vial containing only solvent were subtracted from each experiment. Solutions of compounds were made up to an absorbance of approximately one to ensure optimal equipment performance. Samples were then equilibrated to the experimental temperature inside the Shimadzu UV-2600 fitted with a CPS 240 temperature control unit. Samples were then removed and irradiated, depending on the wavelength required, with: a Panacol-Elosol UV point light source fitted either with a OG filter 530 nm or a  $360 \pm 5$  nm band pass filter, a 400 nm Bivar UV<sub>3</sub>TZ-400-15 LED with a 15° diffuser, or a Luxeon Rebel Royal Blue 455 nm LED. Samples were then transferred back to the UV detector and full spectra recorded at fixed intervals.

Results from a single sample were used to choose the appropriate monitoring wavelength and length of time for three further fresh samples which were again equilibrated to the experimental temperature inside the spectrophotometer. The samples were then moved and irradiated with the stated wavelength of light for 5 minutes unless otherwise stated. Thermal reversion of each sample was recorded and these data were then plotted as shown below.



**Figure 187.** Relaxation curve example. Raw data (blue) and the calculated data based on optimised variables (red)

An initial estimate was calculated based on the following equations:

<i>min</i>	minimum absorbance reading
<i>delta</i>	change in absorbance over experiment
<i>k</i>	rate constant
<i>t</i>	time in seconds

This is then fitted to the raw data using the Solver feature of Microsoft Excel to perform a non-linear regression that minimises the differences between the calculated values and the raw data by altering the *min*, *delta* and the rate constant *k* of the above equation. Solutions were plotted to check that the data fit the correct shape (Figure 187). The calculated rate for the reaction that is then converted to a half-life based on the following equation:

$$t_{1/2} = \frac{\ln 2}{k}$$

Fits were carried out independently for each sample and an average value is reported with the standard deviation of the half-lives calculated from each of the three experiments.

### 6.1.2 Crosslinking of Slug

Freeze dried peptide previously purified by reverse phase HPLC was dissolved in Tris buffer (50 mM, pH 8.3) containing TCEP (2 mM) to a concentration of 5 mg / mL. One equivalent of crosslinker in the same buffer was added in three aliquots over 2 hours at 4 °C. the resulting solution was stirred at 4 °C overnight or until all non-crosslinked peptide was visible by analytical HPLC.

### 6.1.3 HPLC

Crosslinking reactions were monitored by HPLC using a Dionex Acclaim 120 C18 (3  $\mu$ m) 150 mm x 4.6 mm column eluting with a gradient of 0 to 100% acetonitrile (0.1% TFA):water (0.1% TFA) over 50 minutes at 1 ml/min. This gradient was also used for semi-preparative HPLC for using the same system with a flow rate of 5 ml/min over a Phenomenex Gemini C18 (10  $\mu$ m) 250 mm x 10 mm column. Fractions containing the desired material were combined, freeze dried and stored at -20 °C until needed.

Chemicals were purchased from Sigma Aldrich, Fisher Scientific, Alfa Aesar, or Fluorochem. Unless otherwise indicated all chemicals were used as received without further purification.

When necessary, solvents were dried by standing over oven-dried 3 Å molecular sieves for 48 hours under argon and stored in this state until required.<sup>126</sup> Dry solvents were examined by NMR for traces of water before use.

NMR spectra were recorded at room temperature unless otherwise stated and are referenced to trimethylsilane as 0 ppm. Spectra were acquired on a Bruker Fourier 300 MHz, Avance III 400 MHz or an Avance III 600 MHz with a cryogenically-cooled probe with preamplifiers.

### 6.1.4 Solid Phase Synthesis

Solid phase chemistry allows easy removal of excess reagents simply by washing the resin with an appropriate solvent. A Rink Amide Resin solid support modified with an acid-labile linker was first swelled in a mixture of dimethylformamide (DMF) and dichloromethane (DCM). After half an hour this mixture of solvents was drained off

and a solution containing the first Fmoc-protected amino acid, *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU) and di-*iso*-propylethylamine (DIPEA) was added. In the presence of DIPEA, HBTU reacts with the carboxylate group of the Fmoc-protected amino acid to form an isourea which is subject to nucleophilic attack by either the free amine of the linker to form a peptide bond between the amino acid and the resin or hydroxybenzotriazole to form an active ester. The resin was then washed with DMF then DCM to remove excess reagents and urea byproducts. The Fmoc group was then removed with piperidine and the resin washed once again. Repetitions of this coupling/deprotection cycle allow the peptide chain extended.

#### 6.1.5 Circular Dichroism Spectroscopy

Spectra were recorded between 400 and 250 nm in 1 nm intervals and between 250 to 180 nm in 0.5 nm intervals at 15 °C. Samples were dissolved in potassium phosphate buffer (20 mM, pH 7) in a 0.1 mm length cuvette. All concentrations were determined using a NanoDrop ND-1000 spectrophotometer using the published extinction coefficient of azobenzene **71**.

#### 6.1.6 Fluorescence Anisotropy

Fluorescence anisotropy experiments were performed at 15 °C for ready comparison with previously published results. The fluorescently labelled partner (5 nM) in buffer (3 mL) was titrated with the unlabelled partner to final concentrations of 10 to 3000 nM. For titrations involving light-state Slug peptides, the stock solutions of the peptides were irradiated with UV light for 60 seconds prior to additions and the cuvette itself was also then irradiated with UV light for a further 60 seconds. A minimum of ten anisotropy readings were recorded for each concentration and the results were averaged. Anisotropy was calculated using the following formula.

---

Where GF is the grating factor,  $I_{vh}$  is the intensity with the gratings vertical and horizontal,  $I_{vv}$  is the intensity with the gratings vertical and vertical.

#### 6.1.7 Binding Calculations

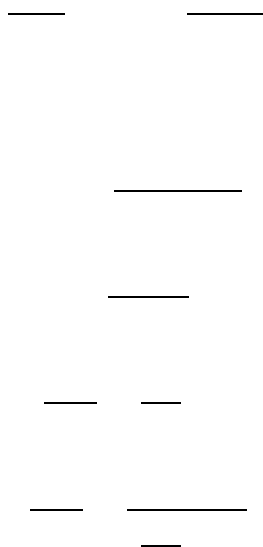
The binding of the crosslinked Slug peptides to the target DNA was calculated from data obtained from anisotropy titrations described above using the following formula.

Where:

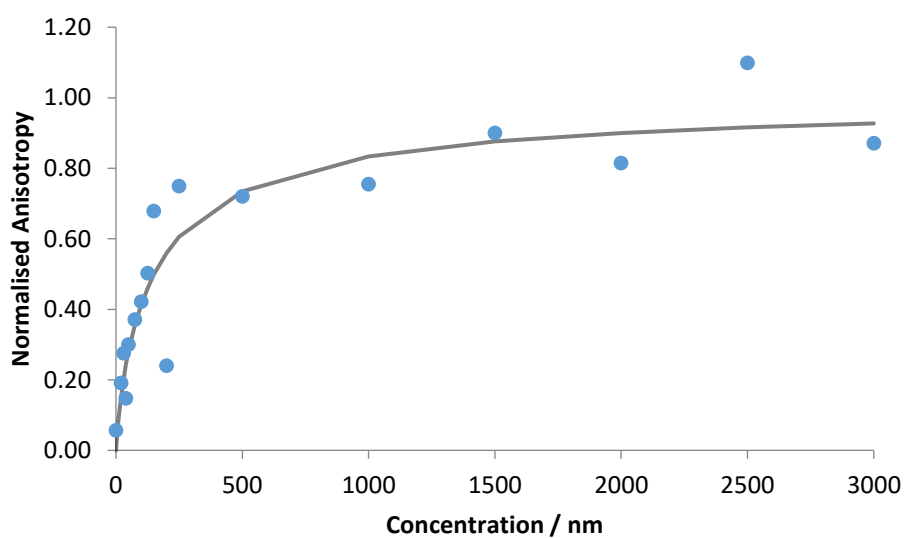
[A] is the concentration of the peptide

[B] is the concentration of the labelled DNA

[AB] is the concentration of the bound complex



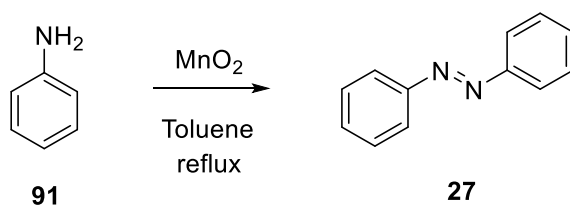
The binding of the peptide to the DNA is calculated by first taking the average of ten anisotropy measurements for each concentration then plotting them against the peptide concentration.



**Figure 188.** Example binding plot from fluorescence anisotropy experiment.

The binding curve shown above in grey is plotted based nonlinear regression fitting using the equation below where  $n$  is the number of sites and  $C$  is the concentration of the peptide. The values for the  $K_d$  are plotted with the use of Solver to minimise the net root mean square difference between the observed and the theoretical data.

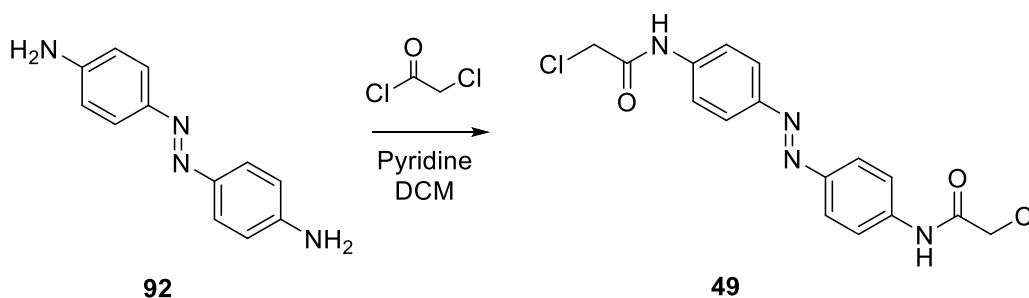
### 6.1.8 Azobenzene (27)



Aniline (1.0 g, 10.7 mM, 1 eq.) and manganese dioxide (5.2 g, 60.0 mM, 5.6 eq.), were suspended in toluene (50 mL), and refluxed under a Dean-Stark trap for 12 hours. The reaction was filtered through celite, and the solvent was removed under reduced pressure. The resulting solid was recrystallised from ethanol to give 27 (0.32 g, 32 %).

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.97 (dd,  $J$  = 8.1, 1.4 Hz, 2H), 7.65-7.21 (m, 3H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 152.7, 131.0, 129.1, 122.9. LRMS (EI +ve) 182.08 (99%), 152.06 (27%), 135.05 (88%), 130.99 (34%), 105.04 (75%), 92.03 (50%), 85.95 (66%), 83.95 (67%), 69.00 (34%), 64.03 (32%), 63.03 (30%). HRMS (ESI +ve) 182.0842 (calculated 182.0844 for C<sub>12</sub>H<sub>10</sub>N<sub>2</sub>).

### 6.1.9 Bis-4,4'-chloroacetamidoazobenzene (49)

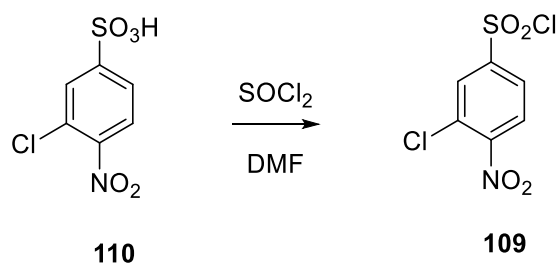


To a suspension of *bis*-4,4'-diaminoazobenzene (50 mg, 0.24 mM, 1 eq.) and pyridine (95  $\mu$ L, 1.18 mM, 4.9 eq.) in dichloromethane (20 mL) was added dropwise chloroacetyl

chloride (145 mg, 1.18 mM, 4.9 eq.). The resulting solution was stirred for 3 days, then water was added. The layers were separated and the aqueous layer was extracted with dichloromethane. The combined organic layers were washed with sodium bicarbonate, dried over sodium sulfate and the solvent was removed under reduced pressure. The resulting residue was purified by silica gel chromatography eluting with ethyl acetate/hexane mixtures (50%:50%) to yield **54** as a brown solid (54.5 mg, 63 %).

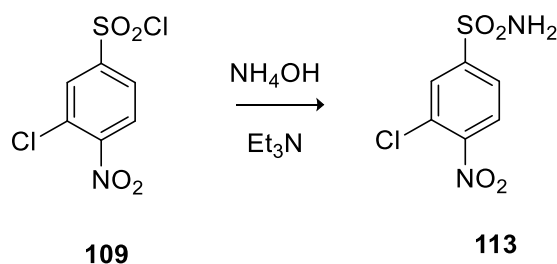
$^1\text{H NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  11.52 (s, 1H), 8.86-8.60 (m, 4H), 5.18 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  166.4, 149.3, 142.5, 124.9, 120.9, 45.00. **LRMS** (EI + ve) 366.05 (50%), 364.05 (77%), 196.02 (28%), 170.02 (32%), 168.02 (100%). **HRMS** (ESI +ve) 364.0496 (calculated 364.0494 for  $\text{C}_{16}\text{H}_{14}\text{N}_2\text{O}_2\text{Cl}_2$ ).

#### 6.1.10 4-Chloro-3-nitrobenzene sulfonate (**109**)



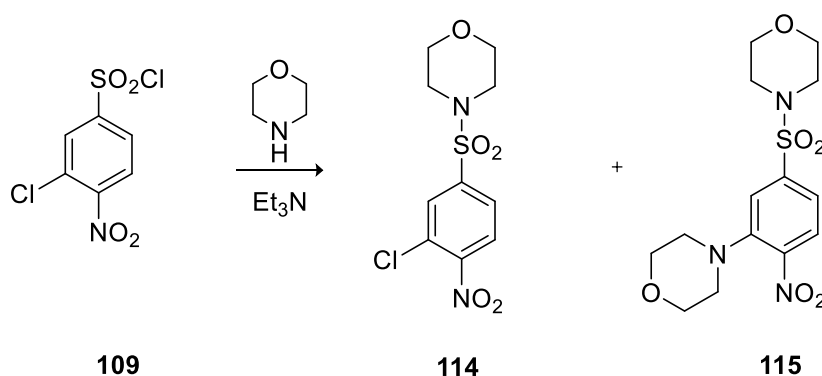
4-Chloro-3-nitrobenzene sulfonate (1.00 g, 4.21 mM, 1 eq.) was dissolved in 1,2-dichloroethane (10 mL) and thionyl chloride (1.00 g, 8.41 mM, 2 eq.) and dimethylformamide (0.16 mL, 2.01 mM, 0.48 eq.) were added. The reaction was refluxed for 6 hours, then cooled to room temperature and the solvent removed under reduced pressure. The residue was dissolved in ether and washed with sodium carbonate and brine then dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **110** (0.9 g, 83.5 %) as an off white solid.

$^1\text{H NMR}$  (250 MHz  $\text{CDCl}_3$ )  $\delta$  8.57 (d, 1H,  $J = 2.5$  Hz), 8.18 (dd, 1H  $J = 2.5$  Hz, 8.5 Hz), 7.88 (d, 1H  $J = 8.5$  Hz).  $^{13}\text{C NMR}$  (75 MHz  $\text{CDCl}_3$ )  $\delta$  147.9, 143.2, 134.8, 133.9, 130.8, 124.5. **LRMS** (ESI +ve) 254.92 (65%  $\text{M}^+$ ), 219.94 (100%  $[\text{M}-\text{Cl}]^+$ ), 155.98 (30%  $[\text{M}-\text{SO}_2\text{Cl}]^+$ ), 109.99 (25%  $[\text{M}-\text{SO}_2\text{Cl}-\text{NO}_2]^+$ ), 75.02 (40%  $[\text{M}-\text{SO}_2\text{Cl}-\text{NO}_2-\text{Cl}]^+$ ). **HRMS** (EI +ve) 254.9160 (calculated 254.9160 for  $\text{C}_6\text{H}_3\text{NO}_4\text{SCl}_2$ ) **MP** 102-104 °C.

6.1.11 4-Chloro-3-nitrobenzenesulfonamide (**113**)

4-Chloro-3-nitrobenzenesulfonyl chloride (0.25 g, 0.98 mM, 1 eq.) was dissolved in tetrahydrofuran (5 mL), ammonium hydroxide solution (35 % 0.35 mL, 3.49 mM, 3.5 eq.) and triethylamine (1.07 mL, 1.36 mM, 1.4 eq.) were added and the resulting solution was stirred for 3 hours. The solution was then extracted with dichloromethane and the organic layer was washed with dilute hydrochloric acid solution (1 M), saturated sodium hydrogen carbonate and brine, then dried over magnesium sulfate, filtered and the solvent was removed under reduced pressure to yield **113** (0.10 g, 43 %) as an orange oil.

$^1\text{H NMR}$  (250 MHz  $\text{CDCl}_3$ )  $\delta$  8.45 (d, 1H,  $J = 2.0$ ), 8.10 (dd, 1H,  $J = 8.5$  Hz, 2.0 Hz), 8.01 (d, 1H,  $J = 8.5$  Hz), 3.45 (s, 2H),  $^{13}\text{C NMR}$  (75 MHz  $\text{CDCl}_3$ )  $\delta$  123.5, 129.3, 131.1, 133.4, 144.3, 147.6. **LRMS** (EI +ve) 235.97 (45%  $\text{M}^+$ ) 219.95 (33%  $[\text{M}-\text{NH}_2]^+$ ) 205.95 (100%), 142.03 (48%), 126.01 (92%), 90.03 (69%), 71.05 (38%). **HRMS** (EI +ve) 235.9660 (calculated 235.9660 for  $\text{C}_6\text{H}_5\text{N}_2\text{O}_4\text{SCl}$ ).

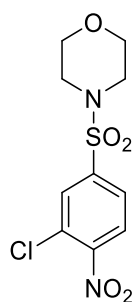
6.1.12 4[(3-Chloro-4-nitrophenol) sulphonyl] morpholine (**114** and **115**)

4-Chloro-3-nitrobenzenesulfonyl chloride (**108**, 5.0 g, 19.5 mM 1 eq.) was dissolved in dichloromethane (50 mL) and the solution was cooled to 5 °C in an ice bath. Triethylamine (6.6 mL, 48mM 2.5 eq.) and morpholine (1.60 mL, 18.4 mM 0.9 eq.) were added and the reaction was stirred for 3 hours at room temperature. The solution was

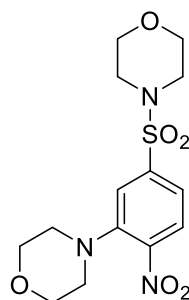


then washed with dilute hydrochloric acid (1 M), saturated sodium hydrogen carbonate and brine. The organic layer was dried over magnesium sulfate, filtered and the solvent was removed under reduced pressure to yield an oily orange mixture of two products. This residue was purified over silica gel eluting with 1% methanol in dichloromethane to yield **114** (0.73 g, 13.7 %) as a yellow oil and **115** (1.81 g, 26 %) as a dark green oil.

An improved procedure was developed whereby **109** (0.15 g, 0.56 mM 1 eq.) was dissolved in dichloromethane (50 mL) and cooled to 5 °C. Diisopropylethylamine (0.2 mL, 1.76 mM 3 eq.) and morpholine (46 mg, 0.53 mM 0.9 eq.) were added dropwise over 30 minutes and the reaction was stirred for 1.5 hours at 5 °C. The reaction was then quenched with saturated sodium hydrogen carbonate and the organic fraction was washed with brine, dried over magnesium sulfate, filtered and the solvent was removed under reduced pressure to yield **114** (180 mg, 99 %) as an orange solid.

**114**

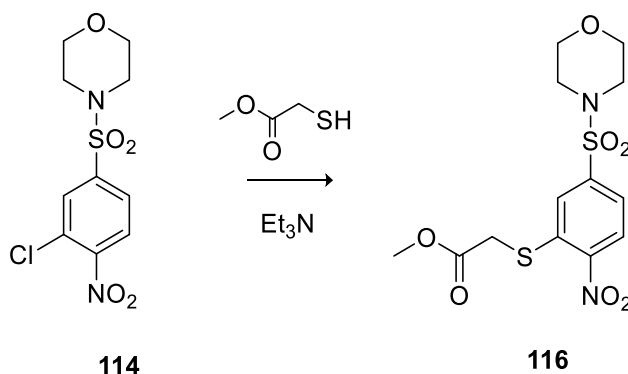
<sup>1</sup>H NMR (250 MHz CDCl<sub>3</sub>) δ 8.45 (d, 1H, *J* = 2.0), 7.82 (dd, 1H, *J* = 8.0 Hz, 2.0 Hz), 7.69 (d, 1H, *J* = 8.0 Hz), 7.10 (t, 4H, *J* = 5.0 Hz), 2.98 (t, 4H, *J* = 5.0 Hz). <sup>13</sup>C NMR (75 MHz CDCl<sub>3</sub>) δ 147.8, 135.0, 133.2, 132.2, 130.3, 124.8, 65.2, 45.7. LRMS (EI +ve) 306.01 (40% [M]<sup>+</sup>), 262.99 (53% M+2 -NO<sub>2</sub>), 219.95 (28% [M-(NC<sub>4</sub>H<sub>8</sub>O)]<sup>+</sup>), 155.99 (20%), 109.99 (26%), 86.04 (100%), 75.02 (30%). HRMS (EI +ve) 306.0076 (calculated 306.0077 for C<sub>10</sub>H<sub>11</sub>N<sub>2</sub>O<sub>5</sub>SCl). MP 148-150 °C.



115

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.16 (d,  $J = 2.3$  Hz, 1H), 7.78 (dd,  $J = 8.8, 2.3$  Hz, 1H), 7.16 (d,  $J = 8.8$  Hz, 1H), 3.93-3.82 (m, 4H), 3.82-3.72 (m, 4H), 3.27-3.15 (m, 4H), 3.10-2.96 (m, 4H).  $^{13}\text{C NMR}$  (63 MHz,  $\text{CDCl}_3$ )  $\delta$  207.0, 148.4, 136.5, 132.6, 127.0, 120.0, 66.3, 66.0, 50.9, 46.0. **LRMS** (APCI +ve) 518.09 (30%), 482.33 (29%), 441.31 (25%), 358.11 ( $\text{M}^+$ , 100%), 242.29 (35%), 219.18 (40%), 117.09 (32%). **HRMS** (EI +ve) 357.0998 (calculated 357.0995 for  $\text{C}_{14}\text{H}_{19}\text{N}_3\text{O}_6\text{S}$ ).

#### 6.1.13 Methyl 2-((5-(morpholinosulfonyl)-2-nitrophenyl)thio)acetate (116)



114

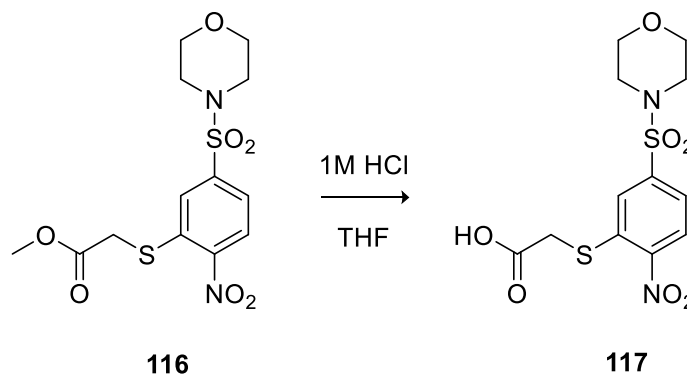
116

4-[(3-Chloro-4-nitrophenyl)sulfonyl]morpholine (1.0 g, 3.26 mM, 1 eq.) was dissolved in acetonitrile (100 mL) and methyl thioglycolate (0.185 mL, 2.93 mM, 0.9 eq.) was added. The mixture was refluxed for 5 hours then cooled to room temperature. The solution was washed with saturated sodium bicarbonate solution, then dried over magnesium sulfate and filtered. The solvent was removed under reduced pressure to give **116** (0.95 g, 77 %).

$^1\text{H NMR}$  (300.MHz  $\text{CDCl}_3$ )  $\delta$  8.58 (d, 1H,  $J = 3.0$  Hz), 7.89 (dd, 1H,  $J = 9.0$  Hz, 3.0 Hz), 7.68 (d, 1H,  $J = 9.0$  Hz), 3.83 (s, 2H), 3.79 (s, 3H), 3.76 (t, 4H,  $J = 6.0$  Hz), 3.04 (t, 4H,  $J = 6.0$  Hz).  $^{13}\text{C NMR}$  (75 MHz  $\text{CDCl}_3$ ) 168.9, 145.0, 143.0, 132.6, 131.9, 127.2, 125.7, 66.0, 53.3, 45.9, 34.7. **LRMS** (EI +ve) 376.05 (64%  $\text{M}^+$ ), 342.03 (20%), 289.98 (20%), 201.97 (20%),

137.00 (24%), 83.95 (100%). **HRMS** (EI +ve) 376.0391 (calculated 376.0399 for  $C_{13}H_{16}N_2O_7S_2$ ). **MP** 140-141 °C.

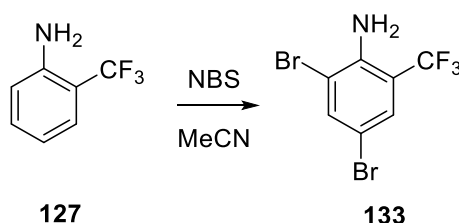
#### 6.1.14 Methyl 2-((5-(morpholinylsulfonyl)-2-nitrophenyl)thio)acetic acid (117)



Methyl 2-((5-(morpholinylsulfonyl)-2-nitrophenyl)thio)acetate (0.25 g, 0.66 mM, 1 eq.) was dissolved in tetrahydrofuran (15 mL) to which was added dilute hydrochloric acid (2 M, 15 mL) and the reaction was stirred for 16 hours at 55 °C. The solution was then cooled and neutralised with saturated sodium bicarbonate solution. The tetrahydrofuran solvent was removed under reduced pressure and the resulting aqueous phase was extracted with ethyl acetate. The organic layer was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **117** (0.15 g, 63 %).

<sup>1</sup>H NMR (300.MHz CDCl<sub>3</sub>) δ 8.51 (d, 1H, *J* = 4.5 Hz), 7.92 (dd, *J* = 9.0 Hz, 4.5 Hz, 1H), 7.50 (d, *J* = 9.0 Hz, 1H), 3.76-3.88 (m, 6H), 3.04 (t, *J* = 4.5 Hz, 4H). **LRMS** (ESI -ve) 363.02 (17%), 362.02 (22%), 361.01 (100% M<sup>-</sup>), 317.03 (19%), 287.02 (23%). **HRMS** (EI +ve) 361.0173 (calculated 361.0164 for  $C_{12}H_{13}N_2O_7S_2$ ).

#### 6.1.15 2,4-Dibromo-6-trifluoromethylaniline (133)

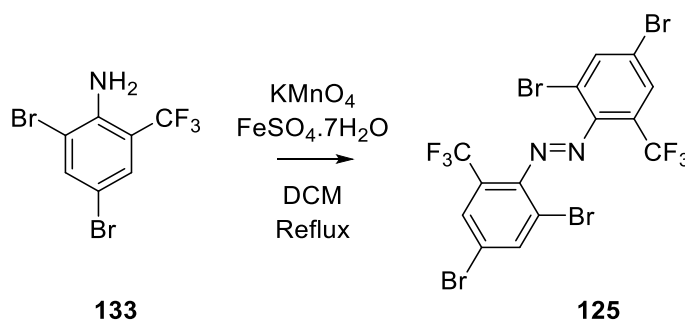


To a solution of 2-trifluoromethyl aniline (1.00 g, 6.21 mM, 1 eq.) in acetonitrile (30 mL) was added *N*-bromosuccinimide (4.86 g, 27.3 mM, 4.4 eq.) and the mixture was stirred overnight. The solvent was then removed under reduced pressure and the solid residue

was partitioned between hexane and water. The organic fraction was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **133** (1.80 g, 90%) as a purple solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.72 (d,  $J = 2.1$  Hz, 1H), 7.54 (d,  $J = 2.1$  Hz, 1H), 4.72 (s, 2).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  141.2, 138.1, 128.8, (d,  $J = 1.0$  Hz), 125.2, 115.5 (d,  $J = 1.0$  Hz), 111.5, 108.2.  $^{19}\text{F NMR}$  (282.24 MHz,  $\text{CDCl}_3$ )  $\delta$  -63.50. **LRMS** (ESI -ve) 321.86 (5%), 318.96 (16%), 317.86 (6%), 276.90 (28%), 274.90 (100%), 272.90 (98%), 254.89 (100%), 252.89 (98%), 227.90 (23%), 217.94 (56%). **HRMS** (APCI +ve) 317.8735 (calculated 317.8735 for  $\text{C}_7\text{H}_5\text{NBr}_2\text{F}_3$ ).

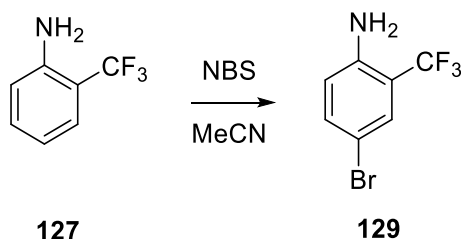
#### 6.1.16 Tetra-2,2',4,4'-bromo-6,6'-di(trifluoromethyl)-azobenzene (**125**)



To a suspension of freshly ground potassium permanganate (2.00 g) and iron (II) sulfate heptahydrate (2.00 g) in dichloromethane (25 mL) was added 2,4-dibromo-6-trifluoromethylaniline (1.00 g, 3.14 mM, 1 eq.). The resulting suspension was refluxed for 18 hours, then cooled and filtered through celite. The solvent was removed under reduced pressure and the residue purified over silica gel eluting with 10% ethyl acetate in hexane to yield **125** (0.20 g, 20%) as a dark red oil.

$^1\text{H NMR}$  (250 MHz,  $\text{CDCl}_3$ )  $\delta$  8.09 (d,  $J = 2.0$  Hz, 1H), 7.92 (d,  $J = 2.0$  Hz, 1H),  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  147.1, 140.7, 129.5, (d,  $J = 1.0$  Hz), 127.5 (d,  $J = 3$  Hz), 123.44 (d,  $J = 2.0$  Hz), 120.49, 114.57.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -58.13. **LRMS** (APCI +ve) 638.70 (18%), 636.70 (64%), 634.70 (100%), 632.71 (66%), 630.71 (19%). **HRMS** (APCI +ve) 630.7077 (calculated 630.7085 for  $\text{C}_{14}\text{H}_5\text{N}_2\text{Br}_4\text{F}_6$ ).

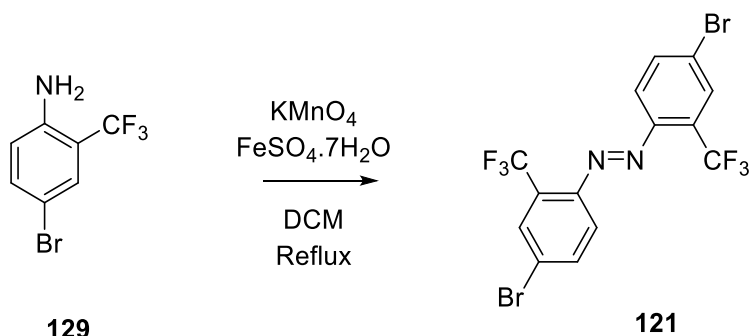
## 6.1.17 4-Bromo-2-trifluoromethylaniline (129)



To a solution of 2-trifluoromethylaniline (5.00 g, 31 mM, 1 eq.) in acetonitrile (30 mL) was added *N*-bromosuccinimide (5.50 g, 31 mM, 0.98 eq.) portion wise and the solution was stirred overnight. The solvent was removed under reduced pressure and the residue partitioned between hexane and water. The organic fraction was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **129** (7.08 g, 99%) as a red oil.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.45 (d, *J* = 2.3 Hz, 1H), 7.28 (dd, *J* = 8.7, 2.3, Hz 2H), 6.54 (d, *J* = 8.7 Hz, 2H), 4.05 (s, 2H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 143.5, 135.6, 129.3, (q, *J* = 1.0 Hz), 125.8, 120.1 (d, *j* = 1.7 Hz), 115.4 (q, *J* = 1.0 Hz), 108.9. <sup>19</sup>F NMR (282.24 MHz, CDCl<sub>3</sub>) δ -77.00. LRMS (ESI +ve) 492.87 (20%), 490.87 (35%), 477.88 (52%), 475.88 (100%), 473.88 (51%), 412.95 (82%), 410.95 (82%), 325.17 (24%), 311.16 (23%), 141.92 (20%). HRMS (APCI +ve) 239.9630 (calculated 239.9630 for C<sub>7</sub>H<sub>5</sub>NBrF<sub>3</sub>).

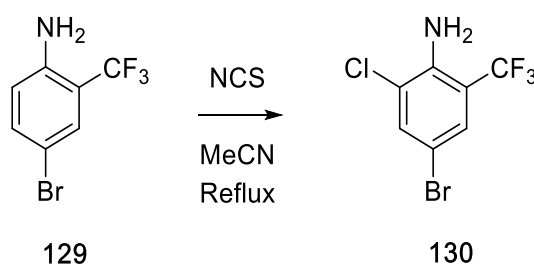
## 6.1.18 Di-4,4'-bromo-2,2'-di(trifluoromethyl)-azobenzene (121)



To a suspension of freshly ground potassium permanganate (1.00 g) and iron (II) sulfate heptahydrate (2.00 g) in dichloromethane (30 mL) was added 4-bromo-2-(trifluoromethyl)aniline (0.50 g, 2.08 mM, 1 eq.). This suspension was refluxed for 2 days, then filtered through celite and the solvents removed under reduced pressure. The residue was purified over silica gel eluting with 5% ethyl acetate in hexane to yield **121** (181 mg, 36.5%) as dark red crystals.

$^1\text{H NMR}$  (250 MHz,  $\text{CDCl}_3$ )  $\delta$  8.00 (d,  $J = 1.9$  Hz, 1H), 7.70-7.90 (m, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  196.2, 147.7, 136.1, 130.1, 126.2, 123.9, 118.2.  $^{19}\text{F NMR}$  (282.24 MHz,  $\text{CDCl}_3$ )  $\delta$  -67.50. **LRMS** (APCI +ve) 478.88 (33%), 476.88 (65%), 474.89 (34%), 398.97 (98%), 396.98 (100%). **HRMS** (APCI +ve) 474.8870 (calculated 474.8875 for  $\text{C}_{14}\text{H}_6\text{N}_2\text{Br}_2\text{F}_6$ ).

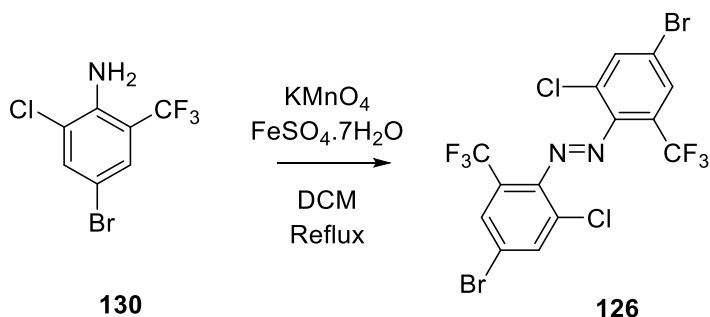
#### 6.1.19 4-Bromo-2-chloro-6-trifluoromethylaniline (130)



To a solution of 4-bromo-2-trifluoromethylaniline (2.00 g, 8.33 mM, 1 eq.) in acetonitrile (30 mL) was added *N*-chlorosuccinimide (1.34 g, 10.08 mM, 1.2 eq.) and the mixture was refluxed overnight. The solvent was removed under reduced pressure and the residue was partitioned between hexane and water. The organic fraction was dried over magnesium sulphate, filtered and the solvent removed under reduced pressure to yield **130** (2.09 g, 91%) as a red oil.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.57 (d,  $J = 2.2$  Hz, 1H), 7.47 (d,  $J = 2.2$  Hz, 1H), 4.67 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  140.2, 135.1, 128.1 (d,  $J = 1$  Hz), 125.3, 121.6 (d,  $J = 2$  Hz), 115.7 (d,  $J = 3$  Hz), 107.8.  $^{19}\text{F NMR}$  (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -65.00. **LRMS** (APCI +ve) 277.92 (25%), 275.92 (100%), 273.92 (100%), 231.97 (33%), 229.97 (51%). **HRMS** (APCI +ve) 273.9241 (calculated 273.9239 for  $\text{C}_7\text{H}_5\text{NBrClF}_3$ ).

#### 6.1.20 4,4'-Dibromo-2,2'-dichloro-6,6'-di(trifluoromethyl)-azobenzene (126)

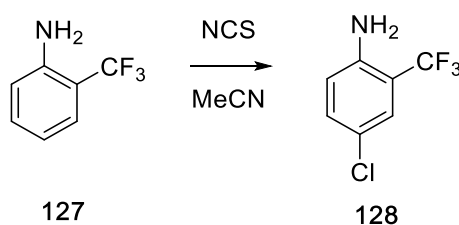


To a suspension of freshly ground potassium permanganate (2.00 g) and iron (II) sulfate heptahydrate (2.00 g) in dichloromethane (30 mL) was added 4-bromo-2-chloro-6-trifluoromethylaniline (1.0 g, 3.67 mM, 1 eq.) and this mixture was refluxed for

18 hours. The suspension was then filtered through celite and the solvent removed under reduced pressure. The residue was purified over silica gel eluting with 5% ethyl acetate in hexane to yield **126** (0.36 g, 18%) as a dark red oil as a mixture of isomers.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.99-7.83 (m, 2H), 7.78 (dd,  $J = m$ , 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ ) 140.8, 135.5, 126.0 (d,  $J = 1.0$  Hz), 125.36, 121.9, (d,  $J = 1.0$  Hz), 115.3, 111.2,  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -58.31. **LRMS** (APCI +ve) 548.80 (27%), 546.80 (88%), 544.81 (100%), 542.81 (36%), 500.86 (22%), 492.86 (51%), 490.87 (80%), 488.87 (47%), 471.13 (22%). **HRMS** (APCI +ve) 542.8086 (calculated 542.8095 for  $\text{C}_{14}\text{H}_5\text{N}_2\text{Br}_2\text{Cl}_2\text{F}_6$ ).

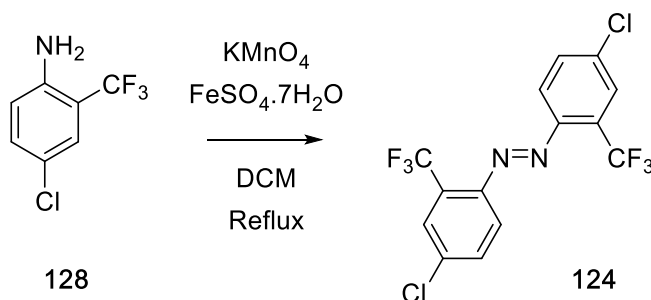
#### 6.1.21 2-Chloro-6-trifluoromethylaniline (**128**)



To a solution of 2-trifluoromethylaniline (2.00 g, 12.41 mM, 1 eq.) in acetonitrile (30 mL) was added *N*-chlorosuccinimide (1.66 g, 12.41 mM, 1 eq.) and the mixture was stirred at room temperature for 4 days. The solvent was removed under reduced pressure and the residue was partitioned between hexane and water. The organic fraction was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **128** (1.69 g, 56%).

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.42-7.40 (m, 1H), 7.29-7.19 (m, 1H), 6.68 (d,  $J = 8.7$  Hz, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  143.0, 132.8, 126.3, 122.4, 118.5.  $^{19}\text{F NMR}$  (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -63.03. **LRMS** (APCI +ve) 198.01 (35%), 196.01 (100%), 176.00 (11%). **HRMS** (APCI +ve) 196.0135 (calculated 196.0135 for  $\text{C}_7\text{H}_6\text{NClF}_3$ ).

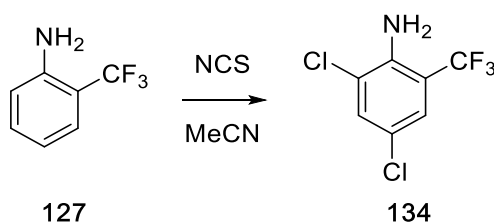
#### 6.1.22 4,4'-Dichloro-2,2'-di(trifluoromethyl)-azobenzene (**124**)



To a suspension of freshly ground potassium permanganate (0.7 g) and iron (II) sulfate heptahydrate (0.7 g) in dichloromethane (30 mL) was added 4-chloro-2-(trifluoromethyl)aniline (0.35 g, 1.79 mM, 1 eq.). The resulting suspension was refluxed for 16 hours then filtered through celite and the solvent removed under reduced pressure. The residue was purified over silica gel eluting with 10% ethyl acetate in hexane to yield **124** (30.5 mg, 9 %) as dark red crystals.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.82-7.72 (m, 2H), 7.57 (dd,  $J = 8.7, 2.1$  Hz, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  147.4, 139.0, 133.0, 131.0 (d,  $J = 1$  Hz), 130.5 (d,  $J = 2$  Hz), 127.0, 118.0.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -57.72. **LRMS** (APCI +ve) 390.98 (10%), 388.99 (6%), 387.99 (13%) 386.99 (100%), 355.02 (14%), 353.03 (40%). **HRMS** (APCI +ve) 386.9885 (calculated 386.9885 for  $\text{C}_{14}\text{H}_6\text{N}_3\text{Cl}_4\text{F}_6$ ).

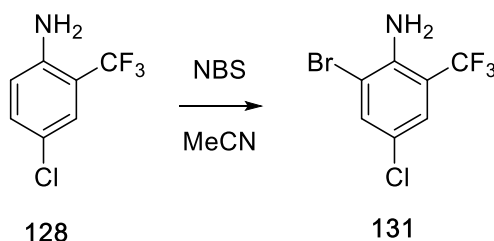
#### 6.1.23 2,4-Dichloro-6-trifluoromethylaniline (**134**)



To a solution of 2-trifluoromethylaniline (2.00 g, 12.41 mM, 1.0 eq.) in acetonitrile (30 mL) was added *N*-chlorosuccinimide (3.65 g, 27.44 mM, 2.2 eq.) and the resulting solution was refluxed for 9 days. The solvent was removed under reduced pressure and the residue was then partitioned between hexane and water. The organic fraction was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **134** (2.42 g, 84%) as a red oil.

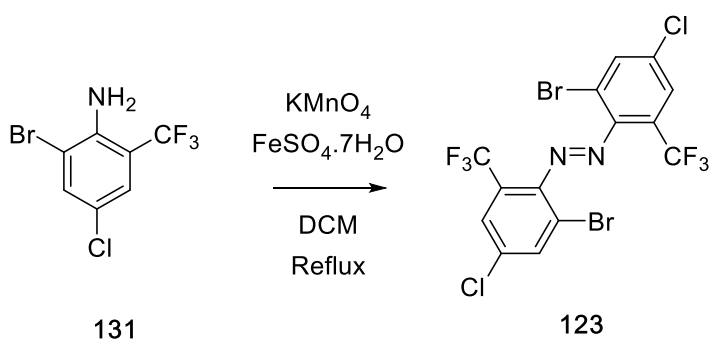
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.43 (d,  $J = 2.4$  Hz, 1H), 7.36 (d,  $J = 2.4$  Hz, 1H), 4.65 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  139.8, 132.44, 125.3, 125.2, 121.5 (d,  $J = 2$  Hz), 115.5, 115.1.  $^{19}\text{F NMR}$  (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -64.00. **LRMS** (APCI +ve) 233.97 (12%), 231.97 (67%), 229.97 (100%). **HRMS** (APCI +ve) 229.9746 (calculated 229.9746 for  $\text{C}_7\text{H}_5\text{NBrCl}_2\text{F}_3$ ).



6.1.24 2-Bromo-4-chloro-6-trifluoromethylaniline (**131**)

To a solution of 4-chloro-2-(trifluoromethyl)aniline (1.00 g, 5.13 mM, 1 eq.) in acetonitrile (20 mL) was added *N*-bromosuccinimide (0.92 g, 5.13 mM, 1 eq.) and the resulting solution was stirred at room temperature for 3 days. The solvent was then removed under reduced pressure and the resulting solid partitioned between water and hexane. The organic layer was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **131** (1.10 g, 79%) as red crystals.

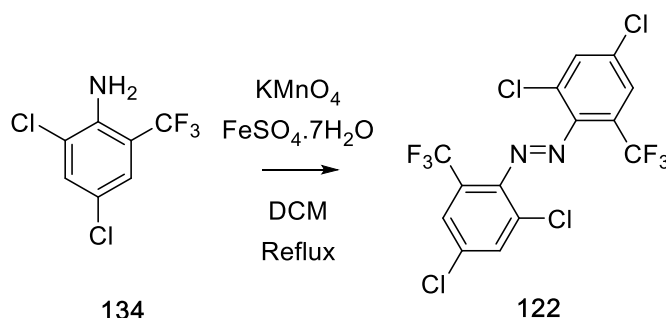
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.36 (d, *J* = 2.3 Hz, 1H), 7.29 (d, *J* = 2.3 Hz, 1H), 4.57 (s, 2H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 140.8, 135.5, 126.0 (d, *J* = 1 Hz), 125.4, 121.9 (d, *J* = 1 Hz), 115.3, 111.2. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -58.31. LRMS (APCI +ve) 277.92 (26%), 275.92 (100%), 273.92 (77%), 229.974 (5%). HRMS (APCI +ve) 273.9241 (calculated 273.9241 for C<sub>7</sub>H<sub>5</sub>NBrClF<sub>3</sub>).

6.1.25 2,2'-Dibromo-4,4'-dichloro-6,6'-di(trifluoromethyl)-azobenzene (**123**)

To a suspension of freshly ground potassium permanganate (1 g) and iron (II) sulfate heptahydrate (1 g) in dichloromethane (30 mL) was added 2-bromo-4-chloro-6-trifluoromethylaniline (0.53 g, 1.94 mM, 1 eq.). The resulting suspension was refluxed for 18 hours then filtered through celite. The solvent was removed under reduced pressure and the residue was purified with silica flash chromatography eluting with 19:1 hexane:ethyl acetate to yield **123** 80 mg 15% as a dark red oil as a mixture of isomers.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.91-7.75 (m, 1H), 7.70 (m, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  146.0, 135.4, 134.7, 127.6 (d,  $J = 4$  Hz), 125.8, 123.9, 120.2.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -58.19, 58.23. **LRMS** (APCI +ve) 548.80 (31%), 546.80 (80%), 544.81 (100%), 542.81 (38%), 445.74 (46%). **HRMS** (APCI +ve) 542.8095 (calculated 542.8086 for  $\text{C}_{14}\text{H}_4\text{N}_2\text{Br}_2\text{Cl}_2\text{F}_6$ ).

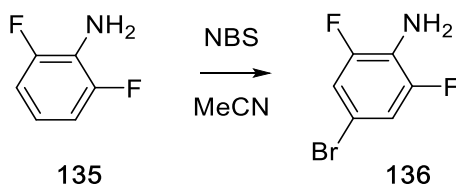
#### 6.1.26 2,2',4,4'-Tetrachloro-6,6'-di(trifluoromethyl)-azobenzene (122)



To a suspension of freshly ground potassium permanganate (2.00 g) and iron (II) sulfate heptahydrate (2.00 g) in dichloromethane (30 mL) was added 2,4-dichloro-6-trifluoromethylaniline (1.00 g, 3.67 mM, 1 eq.) and the mixture was refluxed for 14 hours. The suspension was then filtered through celite and the solvent removed under reduced pressure. The residue was purified over silica gel eluting with hexane to yield **122** (159 mg, 16%) as a dark red oil.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.00-7.54 (m, 4H).  $^{13}\text{C NMR}$  (63 MHz,  $\text{CDCl}_3$ )  $\delta$  146.5, 137.6, 128.8, 127.5, 125.8, 123.0, 119.7.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -58.22, -58.30. **LRMS** (APCI +ve) 548.80 (33%), 546.80 (88%), 544.81 (100%), 542.81 (39%), 502.85 (60%), 500.86 (93%), 498.86 (47%), 456.91 (20%). **HRMS** (APCI +ve) 454.9101 (calculated 454.9106 for  $\text{C}_{14}\text{H}_4\text{N}_2\text{Cl}_4\text{F}_6$ ).

#### 6.1.27 4-Bromo-2,6-difluoroaniline (136)

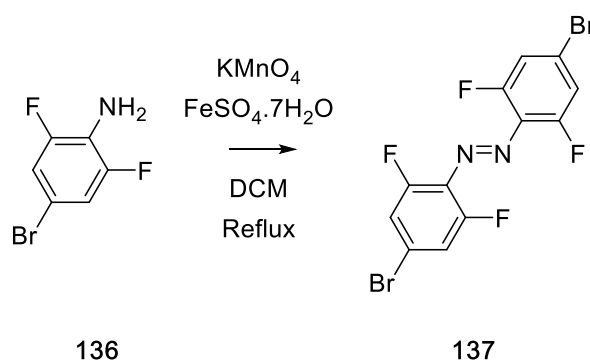


2,6-Difluoroaniline (5.00 g, 38.7 mM, 1 eq.) was added to a solution of *N*-bromosuccinimide (6.90 g 38.9 mM, 1 eq.) in acetonitrile (100 mL) and stirred overnight. Water (30 mL) was added and the mixture extracted with hexane. The

organic fraction was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **136** (8.00 g, 99%) as a pale pink solid.

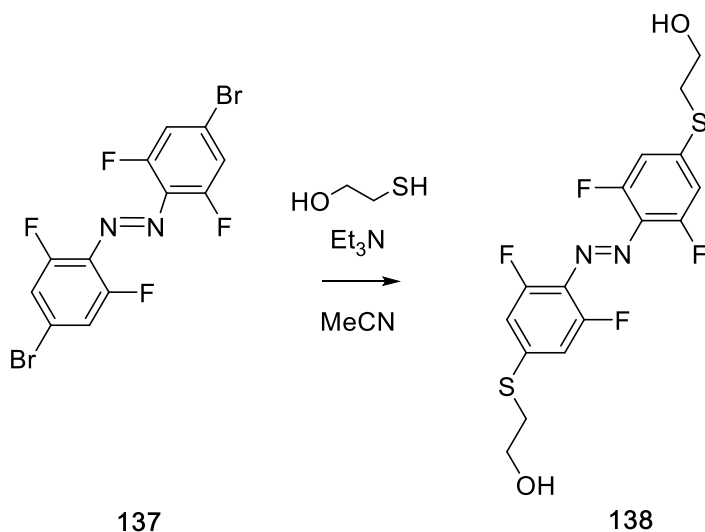
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.15-6.92 (m, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  151.7 (dd,  $J = 8, 8$  Hz), 123.5, 114.8 (dt,  $J = 13, 3$ Hz), 107.2 (d,  $J = 11$  Hz).  $^{19}\text{F NMR}$  (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -118, -119. **LRMS** (ESI -ve) 208.95 (26%), 207.94 (93%), 205.94 (100%,  $[\text{M}-1]^-$ ), 165.05 (70%), 118.90 (18%), 116.92 (24%). **HRMS** (ESI -ve) 205.9426 (calculated 205.9417 for  $\text{C}_6\text{H}_4\text{BrF}_2\text{N}$ ). **MP** 136-138 °C.

#### 6.1.28 1,2-Bis(4-bromo-2,6-difluorophenyl)diazene (**137**)



Potassium permanganate (8.00 g) and iron (II) sulfate heptahydrate (8.00 g) were ground together and suspended in dichloromethane (100 mL) to which was added 4-bromo-2,6-difluoroaniline (3.84 g, 18.5 mM, 1eq.). The suspension was refluxed for 16 hours then filtered through celite and the solvent removed under reduced pressure. The residue was purified over silica gel eluting with 10% ethyl acetate in hexane to yield **137** (0.6 g, 15.7 %) as a red solid as a mixture of isomers.

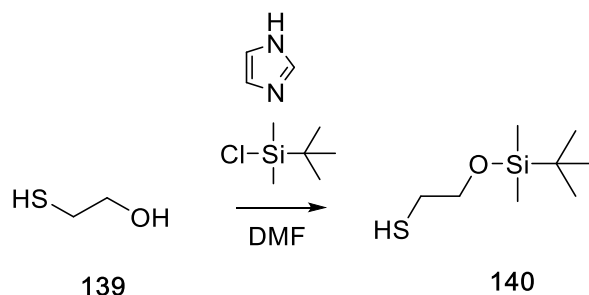
$^1\text{H NMR}$  (300 MHz  $\text{DMSO-d}_6$ )  $\delta$  7.80 (d,  $J = 3.0$  Hz, 2H), 8.00 (d,  $J = 3.0$  Hz, 2H)  $^{13}\text{C NMR}$  (75 MHz  $\text{CDCl}_3$ )  $\delta$  161.4, 160.1, 154.3, 154.3, 153.6, 152.1, 150.4, 130.4, 130.3, 130.3, 125.4, 123.3, 122.48, 120.1, 118.0, 117.9, 117.7, 117.6, 117.3, 117.2, 117.1, 117.0 (observed signals form complex multiplets from  $^{19}\text{F}$  and a mixture of isomers).  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -118.61. **LRMS** (ES -ve) 411.97 (58%,  $\text{M}^+$ ), 410.97 (30%), 412.97 (30%), 218.99 (100%), 220.99 (90%), 190.97 (90%), 192.97 (89%), 126.05 (45%), 112.04 (90%), 69.01 (48%). **HRMS** (ESI +ve) 410.8750 (calculated 410.8756 for  $\text{C}_{12}\text{H}_5\text{Br}_2\text{F}_4\text{N}_2$ ). **MP** 164-165 °C.

6.1.29 (*E*)-2,2'-((diazene-1,2-diylbis(3,5-difluoro-4,1-phenylene))bis(sulfaneydiyl))bis(ethan-1-ol) (**138**)

4,4'-Dibromo-2,2',6,6'-tetrafluoroazobenzene (100 mg, 0.24 mM, 1 eq.) was dissolved in acetonitrile (10 mL) to which was added 2-mercaptoethanol (0.2 mL, 2.8 mM, 12 eq.) and triethylamine (0.1 mL, 0.7 mM, 3 eq.) and the resulting solution was stirred overnight. The solvent was then removed under reduced pressure and the residue partitioned between water and ethyl acetate. The organic layer was washed with brine, dried over magnesium sulfate and filtered. The solvent was removed under reduced pressure and the residue was recrystallised from hot 20% methanol in ethyl acetate to yield **138** (58 mg, 59%) as a mixture of isomers.

$^1\text{H NMR}$  (300 MHz  $\text{CDCl}_3$ )  $\delta$ , 7.62-7.54 (m, 2H), 7.44-7.42 (m, 2H), 5.07-4.97 (m, 2H), 2.57-2.66 (m, 4H), 3.00-3.20 (m, 4H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  157.9, 157.3, 156.2, 155.5, 145.3, 145.2, 139.5, 137.2, 136.7, 135.2, 124.5, 124.3, 124.1, 124.1, 226.3, 116.2, 60.0, 59.8, 59.7, 59.6, 59.5, 49.1, 36.1, 35.6, 35.5, 35.4, 35.2, 35.1 (observed signals form complex multiplets).  $^{19}\text{F NMR}$  (376 MHz, DMSO)  $\delta$  -118.88, -119.78. **LRMS** No ions. **MP** 164-165 °C.

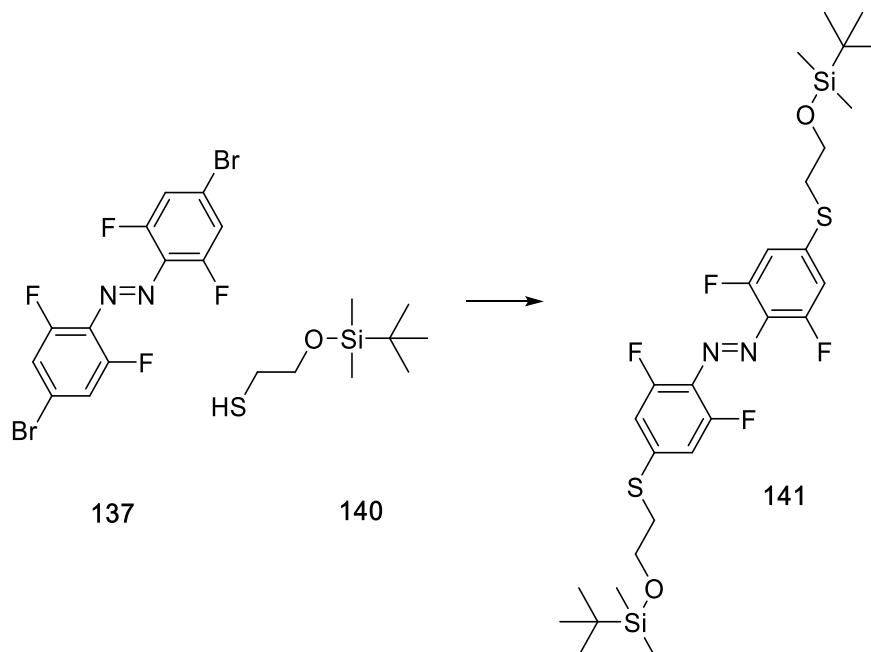
## 6.1.30 2-((Tert-butyldimethylsilyl)oxy)ethane-1-thiol (140)



Imidazole (0.58 g, 8.5 mM, 1.25 eq.) was dissolved in dimethylformamide (2 mL, dried over molecular sieves) under argon to which was added 2-mercaptoethanol (0.56 g, 7.1 mM, 1.05 eq.). The solution was cooled to 0 °C then chloro-*tert*-butyldimethylsilyl (1.03 g, 6.8 mM, 1 eq.) was added in further dimethylformamide (2 mL). This solution was then warmed to room temperature, stirred overnight then diluted with water and hexane. The organic layer was washed with sodium hydrogen carbonate, dried over magnesium sulfate then passed through a bed of silica gel and the solvent removed under reduced pressure to yield **140** (214 mg, 16 %) as a colourless oil.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 3.85 (t, *J* = 6.8 Hz, 2H), 2.81 (t, *J* = 6.8 Hz, 2H), 0.96 (s, 9H), 0.15 (s, 6H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ, 61.9, 41.4, 25.9, 18.4, -2.3. LRMS (ESI +ve) 193.05 (13%), 191.13 (28%), 189.12 (30%), 183.03 (62%), 165.05 (30%), 163.05 (31%), 153.01 (56%), 149.08 (31%), 147.07 (100%). HRMS (EI +ve) 192.0923 (calculated 192.1004 for C<sub>8</sub>H<sub>20</sub>OSiS).

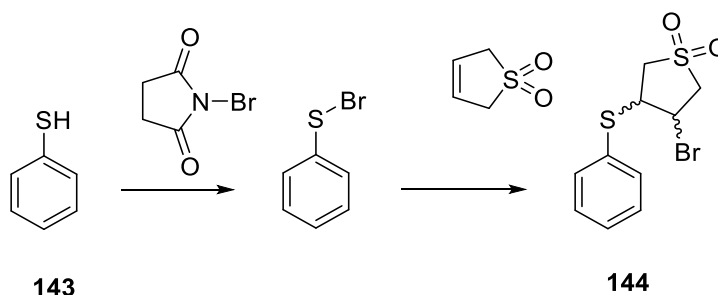
### 6.1.31 (1,2-Bis(4-((tert-butyldimethylsilyl)oxy)ethyl)thio)-2,6-difluorophenyl)diazene (141)



4,4'-Dibromo-2,2',6,6'-tetrafluoroazobenzene (50 mg 0.12 mM, 1 eq.) was dissolved in acetonitrile (10 mL) to which was added 2-((tert-butyldimethylsilyl)oxy)ethane-1-thiol and (50 mg, 0.24 mM, 2 eq.) sodium hydride (60% suspension in mineral oil, 50 mg, 1.25 mM, 10 eq.). The resulting solution was stirred overnight, then the solvent was removed under reduced pressure and the residue partitioned between water and ethyl acetate. The organic layer was washed with brine, dried over magnesium sulfate and filtered. The solvent was removed under reduced pressure and the residue was purified over silica gel eluting with 5% ethyl acetate in hexane to yield **141** (27 mg, 35.1 %) as a red oil.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.46-7.14 (m, 4H), 3.82-3.78 (m, 4H), 3.24-3.04 (m, 3H), 2.75 (t, *J* = 6.8 Hz, 1H), 0.82 (s, 18H), 0.0 (s, 12H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 137.5, 124.26, 115.87 (d, *J* = 12.0 Hz) 116.03, 115.71, 61.52, 41.40, 35.60, 25.90, 18.36, -5.32. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -117.56. LRMS (ESI +ve) 635.19 (14%), 634.19 (17%), 633.18 (38%), 559.16 (20%), 455.32 (22%), 454.31 (100%). HRMS (EI +ve) 633.2203 (calculated 633.2203 for C<sub>28</sub>H<sub>41</sub>F<sub>4</sub>N<sub>2</sub>O<sub>2</sub>S<sub>2</sub>Si<sub>2</sub>).

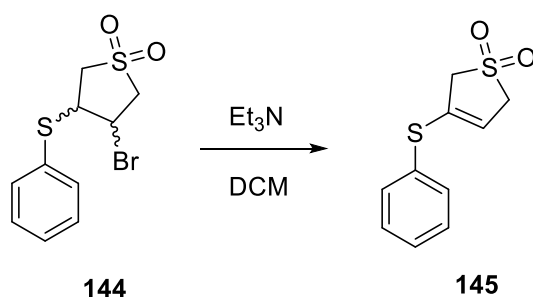
## 6.1.32 3-Bromo-4-(phenylthio)tetrahydrothiophene 1,1-dioxide (144)



Thiophenol (1.00 g, 9.07 mM, 1 eq.) was added dropwise to a stirred solution of *N*-bromosuccinimide (1.65 g, 9.28 mM, 1.2 eq.) in dichloromethane (10 mL) and the reaction was stirred for 30 minutes. Butadiene sulfone (1.1 g, 9.07 mM, 1 eq.) was added and the resulting mixture was stirred for 16 hours at room temperature. The solvent was then removed under reduced pressure and the residue was purified over silica gel eluting with dichloromethane to yield **144** (1.98 g, 71%) as a white crystalline solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.61-7.35 (m, 5H), 4.35 (q,  $J = 6.9$  Hz, 1H), 3.98 (dd,  $J = 14.2, 7.1$  Hz, 1H), 4.07 (dd,  $J = 14.2, 7.1$  Hz, 1H), 3.73 (dd,  $J = 13.8, 7.3$  Hz, 1H), 3.52 (dd,  $J = 14.0, 6.6$  Hz, 1H), 3.17 (dd,  $J = 13.8, 7.2$  Hz, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  134.4, 129.9, 129.8, 129.7, 59.4, 55.9, 51.6, 43.5. **LRMS** (EI +ve) 307.94 (16%), 305.94 (15%), 263.99 (13%), 163.06 (20%), 129.06 (5%), 109.01 (14%), 85.95 (64%), 84.95 (100%), 68.99 (25%). **HRMS** (EI +ve) 305.9384 (calculated 305.9384 for  $\text{C}_{10}\text{H}_{11}\text{O}_2\text{S}_2\text{Br}$ ).

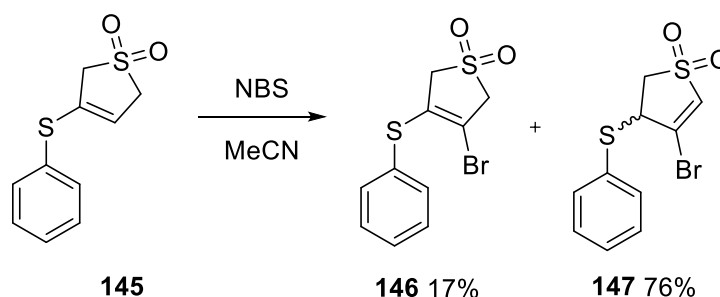
## 6.1.33 3-(Phenylthio)-2,5-dihydrothiophene 1,1-dioxide (145)



To a solution of 3-bromo-4-(phenylthio)tetrahydrothiophene-1,1-dioxide (1.00 g, 3.25 mM, 1 eq.) in dichloromethane (50 mL) was added triethylamine (2.00 mL, 14.35 mM, 4.4 eq.). The resulting solution was stirred overnight. The solution was then washed with water and the organic fraction dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **145** (0.79 g, 96%) as a white solid.

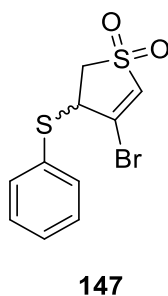
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.44-7.26 (m, 5H), 5.70 (m, 1H), 3.80 (m, 2H), 3.66 (m, 2H).  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  133.8, 133.1, 130.2, 129.8, 129.6, 120.0, 58.3, 58.0. **LRMS** (EI +ve) 226.01 (21%), 162.05 (78%), 161.04 (49%), 147.03 (56%), 134.02 (23%), 129.06 (86%), 128.06 (24%), 110.02 (38%), 109.02 (24%), 91.05 (35%), 70.99 (21%), 68.99 (100%). **HRMS** (EI +ve) 226.0127 (calculated 226.0122 for  $\text{C}_{10}\text{H}_{11}\text{O}_2\text{S}_2$ ).

#### 6.1.34 4-Bromo-3-(phenylthio)-2,3-dihydrothiophene 1,1-dioxide (146) and 3-bromo-4-(phenylthio)-2,5-dihydrothiophene 1,1-dioxide (147)



To a solution of 3-(phenylthio)-2,5-dihydrothiophene-1,1-dioxide (500 mg, 2.21 mM, 1.0 eq.) in acetonitrile (10 mL) was added *N*-bromosuccinimide (450 mg, 2.54 mM, 1.15 eq.). The resulting solution was stirred for 12 hours at 40 °C. The solvent was then removed under reduced pressure, the residue was redissolved in chloroform, filtered and the solvent again removed under reduced pressure. The residue was purified over silica gel eluting with 20% ethyl acetate in hexane to yield compound **146** (116 mg, 17%) as a colourless oil and compound **147** (510 mg, 76%) as white crystals.

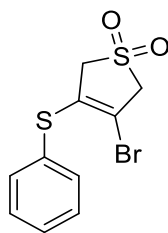
#### 6.1.35 4-Bromo-3-(phenylthio)-2,3-dihydrothiophene 1,1-dioxide (147)



$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.61-7.19 (m, 5H), 5.78 (s, 1H), 5.10-5.00 (m, 1H), 3.91 (dd,  $J = 14.4, 7.6$  Hz, 1H), 3.66 (dd,  $J = 14.4, 3.2$  Hz, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  157.5, 135.2, 131.0, 130.5, 127.7, 122.7, 59.8, 39.3. **LRMS** (EI +ve) 305.9 (68%), 303.9 (51%), 192.0 (20%), 160.0 (100%), 159.1 (42%), 159.0 (40%), 147.0 (40%). **HRMS** (EI +ve) 305.9384 (calculated 305.9384 for  $\text{C}_{10}\text{H}_{11}\text{O}_2\text{S}_2\text{Br}$ ).



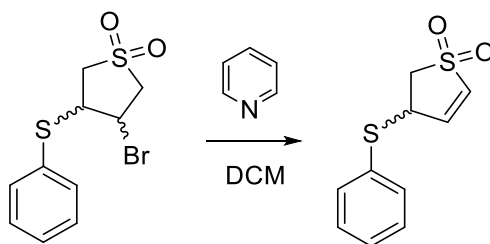
## 6.1.36 3-Bromo-4-(phenylthio)-2,5-dihydrothiophene 1,1-dioxide (146)



146

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.35-7.51 (m, 5H), 6.82-6.59 (CH, m, 2H), 4.48 (CH, dd,  $J$  = 4.9, 2.9 Hz, 1H), 3.65 (CH<sub>a</sub>H<sub>b</sub>, dd,  $J$  = 14.0, 8.1 Hz, 1H), 3.24 (CH<sub>a</sub>H<sub>b</sub>, dd,  $J$  = 14.0, 4.7 Hz, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  139.8, 134.1, 132.9, 130.5, 129.7, 129.4, 54.30, 44.9. **LRMS** (EI +ve), 226.01 (39%), 186.05 (23%), 185.05 (21%), 162.03 (100%), 161.03 (90%), 147.02 (88%), 129.06 (91%), 128.06 (54%), 110.01 (70%), 109.01 (53%) 85.01 (47%). **HRMS** (EI +ve) 303.9224 (calculated 303.9224 for  $\text{C}_{10}\text{H}_9\text{O}_2\text{S}_2\text{Br}$ ).

## 6.1.37 3-(Phenylthio)-2,3-dihydrothiophene 1,1-dioxide (142)



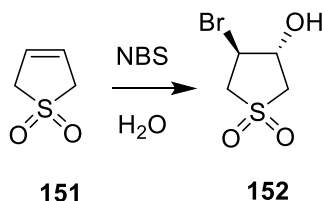
144

142

To a solution of 3-bromo-4-(phenylthio)tetrahydrothiophene-1,1-dioxide (500 mg, 1.63 mM, 1 eq.) in dichloromethane (50 mL) was added pyridine (0.33 mL, 4.07 mM, 2.5 eq.) and the solution was refluxed for three days. The solution was then washed with water and the organic fraction dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **142** (278 mg, 65%) as a white solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.33-7.47 (m, 5H), 6.80-6.56 (m, 2H), 4.45 (dd,  $J$  = 4.9, 2.9 Hz, 1H), 3.63 (dd,  $J$  = 14.0, 8.1 Hz, 1H), 3.21 (dd,  $J$  = 14.0, 4.9 Hz, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  139.8, 134.1, 132.9, 130.5, 129.7, 129.4, 54.3, 44.9. **LRMS** (EI +ve) 226.01 (17%), 162.05 (100%), 161.04 (50%), 147.03 (52%), 129.07 (58%), 128.06 (27%), 110.02 (47%), 109.01 (20%), 79.99 (47%), 68.99 (31%), 63.96 (28%). **HRMS** (EI +ve) 226.0128 (calculated 226.0122 for  $\text{C}_{10}\text{H}_{10}\text{O}_2\text{S}_2$ ).

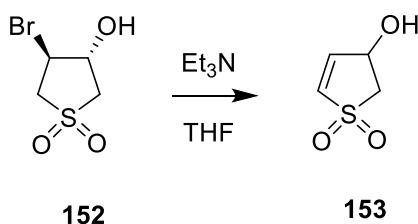
## 6.1.38 3-Bromo-4-hydroxytetrahydrothiophene 1,1-dioxide (152)



Butadiene sulfone (5.00 g, 42.3 mM, 1 eq.) was dissolved in water (100 mL) and *N*-bromosuccinimide (7.55 g, 42.3 mM, 1 eq.) added. The resulting solution was stirred for 3 days at room temperature. The solid was then filtered off and washed with cold sodium carbonate then dried under filtration to yield **152** (5.10 g, 56%) as a white solid.

<sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>COCD<sub>3</sub>) δ 4.86-4.71 (m, 1H), 4.64 (dd, *J* = 11.6, 6.0 Hz, 1H), 3.90 (dd, *J* = 14.2, 6.7 Hz, 1H), 3.63 (dd, *J* = 13.7, 6.1 Hz, 1H), 3.44 (dd, *J* = 14.2, 5.5 Hz, 1H), 3.14 (dd, *J* = 13.7, 4.4 Hz, 1H). <sup>13</sup>C NMR (75 MHz, CD<sub>3</sub>COCD<sub>3</sub>) δ, 74.2, 58.1, 56.9, 47.1. HRMS (ESI +ve) 212.9228 (calculated 212.9221 for C<sub>4</sub>H<sub>6</sub>O<sub>3</sub>SBr).

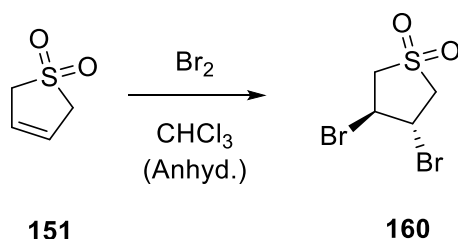
## 6.1.39 3-Hydroxy-2,3-dihydrothiophene 1,1-dioxide (153)



To a solution of 3-bromo-4-hydroxytetrahydrothiophene-1,1-dioxide (0.34 g, 1.50 mM, 1 eq.) in tetrahydrofuran was added triethylamine (0.44 mL, 3.2 mM, 2.1 eq.). The resulting solution was stirred for 3 hours then filtered through a thin pad of silica gel and the solvent removed under reduced pressure to yield **153** (0.22 g, 99%) as an orange oil.

<sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>OD) δ 6.85 (m, 2H), 5.09 (dddd, *J* = 7.5, 3.9, 2.4, 1.3 Hz, 1H), 3.67 (dd, *J* = 13.8, 7.5 Hz, 1H), 3.04 (dd, *J* = 13.8, 3.9 Hz, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ, 140.3, 132.2, 67.2, 55.7. LRMS (ESI +ve) 115.99 (13%), 99.02 (100%), 90.98 (38%), 86.99 (27%), 71.01 (33%), 70.03 (99%), 69.03(54%), 68.02 (49%), 64.96 (26%), 63.96 (38%), 62.99(24%), 55.01 (94%). HRMS (ESI +ve) 115.9936 (calculated 115.9932 for C<sub>4</sub>H<sub>4</sub>O<sub>3</sub>S).

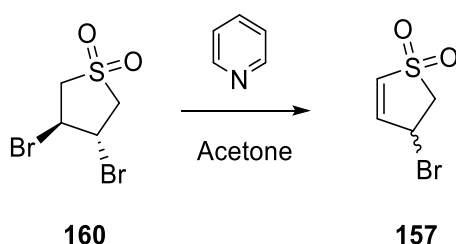
## 6.1.40 3,4-Dibromotetrahydrothiophene 1,1-dioxide (160)



Butadiene sulfone (5.00 g, 42 mM, 1 eq.) was dissolved in dry chloroform (8 mL), stirred under argon and heated to reflux. Bromine (6.76 g, 42 mM, 1 eq.) in further dry chloroform (6 mL) was added dropwise and the reaction was stirred for 2 hours. The reaction was then cooled and the solvent removed under reduced pressure. The residue was washed repeatedly with chloroform until the product was colourless. The residue was recrystallised from dichloromethane to give **160** (10.41 g, 95%) as a white crystalline solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  4.88-4.73 (m, 2H), 4.16-3.93 (m, 2H), 3.62-3.49 (m, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$ , 58.3, 45.6. **LRMS** (EI +ve) 279.83 (12%), 277.84 (52%), 275.84 (13%), 214.87 (10%), 212.87 (72%), 210.87 (16%), 134.93 (100%), 120.93 (100%), 105.93 (100%), 104.93 (100%), 89.01 (22%), 83.95 (44%), 81.92 (52%), 80.92 (67%), 78.92 (42%), 54.04 (78%), 53.03 (100%), 51.02 (44%). **IR** (solid) 700.2, 817.9, 833.3, 871.9, 904.7, 954.8, 1097.5, 1113.0, 1140.0, 1174.7, 1205.6, 1234.5, 1261.5, 1286.6, 1311.7, 1404.2, 2935.8, 2955.1, 2984.0, 3020.7. **HRMS** (EI +ve) 275.8462 (calculated 275.8455 for  $\text{C}_4\text{H}_6\text{O}_2\text{SBr}_2$ ).

## 6.1.41 3-Bromo-2,3-dihydrothiophene 1,1-dioxide (157)

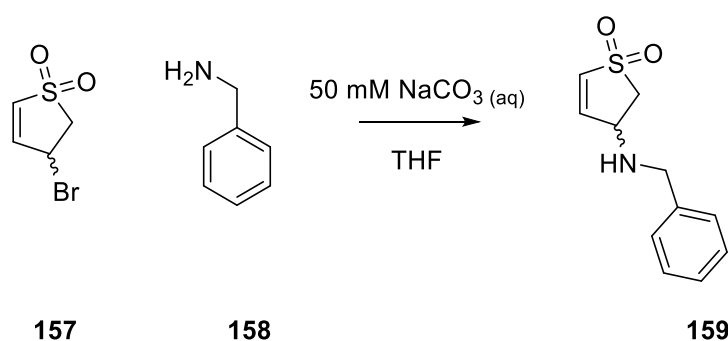


3,4-Dibromotetrahydrothiophene-1,1-dioxide (2.00 g, 7.20 mM, 1 eq.) was dissolved in acetone (14 mL) to which was added pyridine (1.14 g, 14.4 mM, 2 eq.). The reaction was then stirred for 14 hours at room temperature. The solvent was then removed under reduced pressure the any oil was triturated with ether (30 mL). The oil was then

trituated with hot toluene (3 x 50 mL) the combined extracts were dried over magnesium sulfate filtered and dried to give **157** (1.17 g, 83%) as a light orange oil.

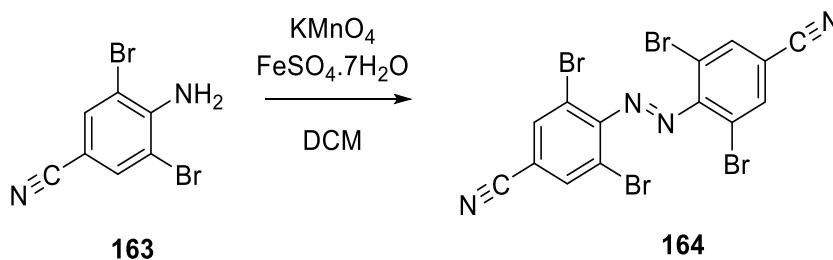
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  6.81 (ddd,  $J = 7.9, 6.6, 2.2$  Hz, 2H), 5.15 (dtd,  $J = 7.8, 3.2, 1.3$  Hz, 1H), 3.84 (dd,  $J = 14.6, 7.7$  Hz, 1H), 3.64-3.51 (m, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  139.5, 133.0, 56.8, 37.5. **LRMS** 197.92 (9%), 195.92 (8%), 133.96 (49%), 131.96 (49%), 103.9 (22%), 89.00 (24%), 78.04 (45%), 68.02 (73%), 63.96 (58%), 53.03 (38%). **HRMS** (EI +ve) 195.9193 (calculated 195.9194 for  $\text{C}_4\text{H}_5\text{O}_2\text{SBr}$ ).

#### 6.1.42 3-(Benzylamino)-2,3-dihydrothiophene 1,1-dioxide (**159**)



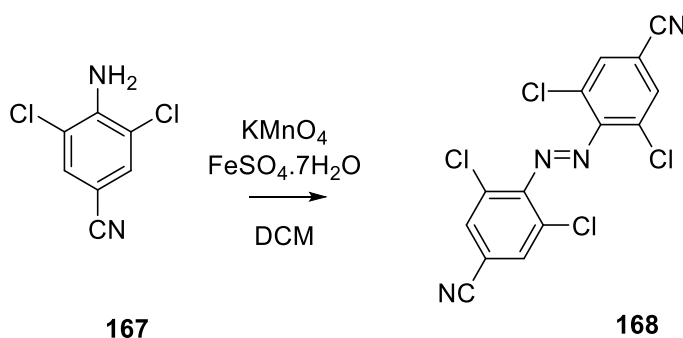
3-Bromo-2,3-dihydrothiophene-1,1-dioxide (0.100 g, 0.50 mM, 1 eq.) and benzylamine (0.054 g, 0.50 mM, 1 eq.) were dissolved in a solution of sodium carbonate in deuterated water (50 mM, 1.80 mL) and tetrahydrofuran (0.20 mL). The reaction was stirred at 50 °C for 2 hours then extracted with dichloromethane. The organic fraction was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **159** (0.119 g, 77%) as a white crystalline solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.36-7.32 (m, 5H), 6.74 (dd,  $J = 6.7, 2.8$  Hz, 1H), 6.66 (dd,  $J = 6.7, 1.8$  Hz, 1H), 4.24 (ddd,  $J = 5.8, 4.3, 2.7$  Hz, 1H), 3.86 (d,  $J = 7.8$  Hz, 2H), 3.47 (dt,  $J = 13.2, 6.6$  Hz, 1H), 3.11 (dd,  $J = 13.6, 4.0$  Hz, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  141.4, 138.7, 132.7, 128.8, 128.7, 128.2, 127.7, 127.4, 56.3, 54.7, 51.4. **LRMS** (AP EI +ve) 224.07 (22%), 132.08 (100%), 124.00 (25%), 122.03 (64%). **HRMS** (ES -ve) 224.0742 (calculated 224.0745 for  $\text{C}_{11}\text{H}_{14}\text{NO}_2\text{S}$ ).

6.1.43 (*E*)-4,4'-(Diazene-1,2-diyl)bis(3,5-dibromobenzonitrile) (**164**)

Potassium permanganate (4.00 g) and iron (II) sulfate heptahydrate (4.00 g) were ground together to a fine powder, then suspended in dichloromethane (50 mL) and dibromo-2,6-ethyl-(4-aminobenzoates (2.0g, 7.25 mM, 1 eq added. The suspension was refluxed overnight then cooled, filtered through celite and the solvent removed under reduced pressure. The resulting solid was recrystallised from ethyl acetate to yield **164** (0.38 g, 19%) as a dark red solid.

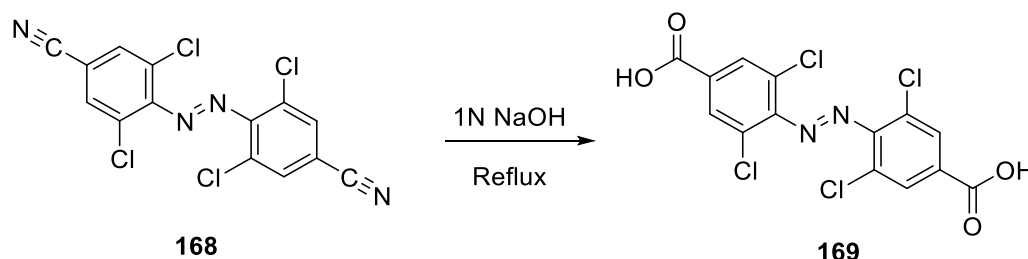
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.00 (s, 4H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 151.6, 136.5, 116.4, 115.0. LRMS (ESI -ve) 549.70 (20%), 547.71 (32%), 545.71 (21%), 250.15 (38%), 220.18 (27%), 219.16 (100%), 80.91 (28%), 79.91 (27%). HRMS (APCI +ve) 544.7226 (calculated 544.77242 for C<sub>14</sub>H<sub>5</sub>N<sub>4</sub>). IR (solid) cm<sup>-1</sup> 3063, 2230, 1530, 1422, 1371, 1364, 1209, 1196, 880, 758, 611.

6.1.44 (*E*)-4,4'-(Diazene-1,2-diyl)bis(3,5-dichlorobenzonitrile) (**168**)

To a suspension of freshly ground potassium permanganate (4.0 g) and iron (II) sulfate heptahydrate (4.0 g) in dichloromethane (25 mL) was added 2,6-dichloro-4-nitrilaniline (2.0 g, 10.07 mM, 1 eq.) and the resulting suspension was then refluxed for 18 hours. The reaction was cooled, filtered through celite and the solvent removed under reduced pressure. The residue was recrystallised from ethyl acetate to yield **168** (256 mg, 14%) as a red solid.

$^1\text{H NMR}$  (250 MHz,  $\text{CDCl}_3$ )  $\delta$  7.72 (s, 4H),  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  132.8, 132.6, 128.3, 126.9, 114.3. **LRMS** (ES  $-ve$ ) 371.91 (15%), 369.92 (37%), 367.92 (33%), 188.96 (12%), 186.96 (72%), 184.97 (100%), 148.99 (22%). 96.96 (42%). **HRMS** (APCI  $+ve$ ) 630.7077 (calculated 630.7085 for  $\text{C}_{14}\text{H}_5\text{N}_2\text{Br}_4\text{F}_6$ ). **IR** (solid)  $\text{cm}^{-1}$  3950, 3376, 3057, 2234, 1616, 1541, 1381, 1204, 1198, 828, 814, 797, 745, 621.

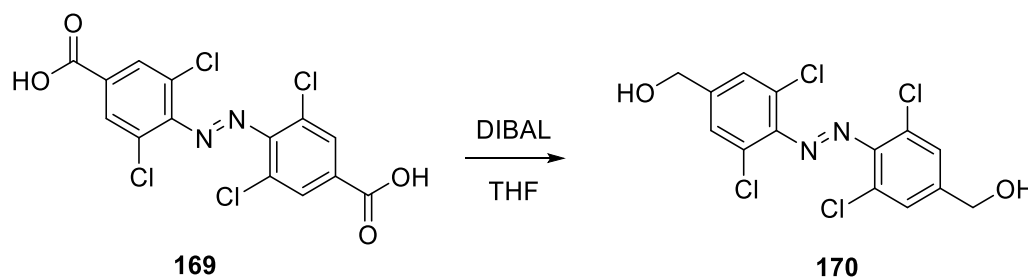
#### 6.1.45 (*E*)-4,4'-(Diazene-1,2-diyl)bis(3,5-dichlorobenzoic acid) (169)



(*E*)-4,4'-(Diazene-1,2-diyl)-bis(-3,5-dichlorobenzonitrile) (55 mg, 0.15 mM, 1 eq.) was suspended in sodium hydroxide solution (1 M, 10 mL) and refluxed for 16 hours. After cooling to room temperature the solution was acidified with hydrochloric acid (1 M) and the precipitate was filtered off. This residue was dissolved in ethyl acetate and this solution was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to give **169** (53 mg, 88%) as a white solid.

$^1\text{H NMR}$  (300 MHz,  $\text{D}_2\text{O}$ )  $\delta$  7.83 (s, 4H),  $^{13}\text{C NMR}$  (75 MHz,  $\text{D}_2\text{O}$ )  $\delta$  171.5, 162.5, 157.6, 147.7, 138.9, 129.8, 126.2. **LRMS** (ESI  $+ve$ ) 408.89 (54%), 406.90 (100%), 404.90 (82%), 325.18 (27%), 293.19 (27%), 311.17 (27%), 293.18 (23%), 190.95 (26%), 188.95 (38%), 112.98 (28%), 98.94 (31%), 96.95 (27%). **HRMS** (EI  $-ve$ ) 404.8995 (calculated 404.9003 for  $\text{C}_{14}\text{H}_5\text{Cl}_4\text{N}_2\text{O}_4$ ). **IR** (solid)  $\text{cm}^{-1}$  1690, 1624, 1589, 1582, 1553, 1537, 1441, 1431, 1385, 1292, 1287, 1273, 818, 768, 789.

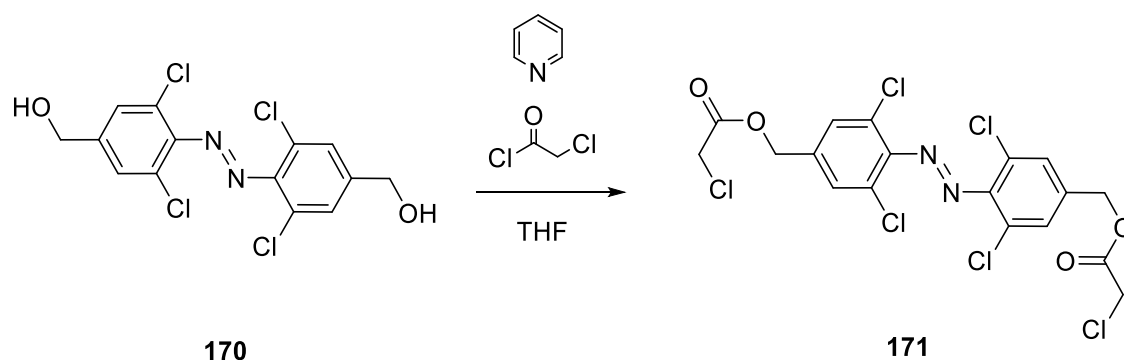
#### 6.1.46 (*E*)-(Diazene-1,2-diyl)bis(3,5-dichloro-4,1-phenylene)dimethanol (170)



To a solution of (*E*)-4,4'-(diazene-1,2-diyl)-bis-(3,5-dichlorobenzoic acid) (53 mg, 0.13 mM, 1 eq.) in tetrahydrofuran (5 mL) was added diisobutylaluminium hydride solution (0.100 mL, 1.00 mM, 7.7 eq.) and the reaction stirred overnight. Saturated sodium potassium tartrate solution was added, the phases were separated and the organic layer was dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The residue was recrystallised from methanol to give **170** (41 mg, 82%) as a white solid.

<sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD) δ 7.56 (d, *J* = 0.6 Hz, 4H), 4.68 (s, 4H). <sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD) δ 145.9, 145.2, 127.0, 126.9, 62.0. LRMS (EI -ve) 380.94 (67%), 378.94 (100%), 376.94 (86% [M]<sup>-</sup>). 350.94 (29%), 349.94 (56%), 347.94 (46%). HRMS (EI -ve) 376.9413 (calculated 376.9418 for C<sub>14</sub>H<sub>9</sub>Cl<sub>4</sub>N<sub>2</sub>O<sub>2</sub>).

#### 6.1.47 (*E*)-(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))bis(methylene) bis(2-chloroacetate) (**171**)

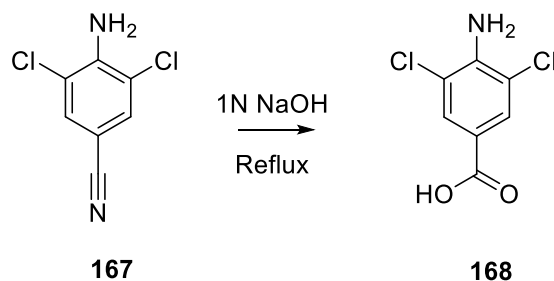


(*E*)-(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene)) (30 mg, 0.054 mM, 1 eq.) was dissolved in tetrahydrofuran and the flask flushed with argon. Chloroacetyl chloride (100 mL, excess), was added dropwise and the reaction was stirred for 20 minutes at room temperature. Pyridine (100 μL, excess) was added dropwise over 20 minutes and the solution stirred overnight. The solvent was removed under reduced pressure and the residue was partitioned between dichloromethane and sodium hydrogen carbonate. The organic layer was dried over sodium sulfate and filtered, then the solvent was removed under reduced pressure and the resulting solid was recrystallised from methanol to yield **171** (39 mg, 93%) as deep red crystals.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.71 (s, 4H), 5.21 (s, 4H), 4.16 (s, 4H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 167.0, 148.7, 137.9, 133.0, 116.2, 65.5, 40.7. HRMS (EI +ve) 529.8932 (calculated

529.8928 for  $C_{18}H_{12}N_2O_4Cl_6$ ). IR (solid)  $cm^{-1}$  3069, 2955, 2922, 2853, 1749, 1591, 1557, 1441, 1402, 1369650, 1310, 1258, 1198, 1186, 1159, 1020, 982, 928, 897, 885, 854, 800, 789, 777, 723, 685, 586, 557, 523, 444, 424, 419.

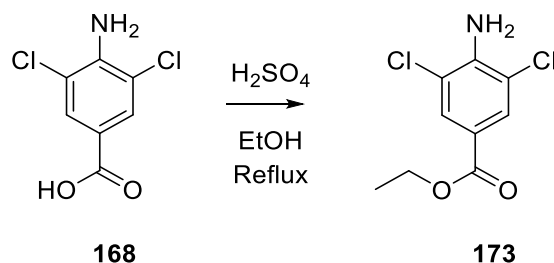
#### 6.1.48 4-Amino-3,5-dichlorobenzoic acid (168)



4-Amino-3,5-dichlorobenzoic acid (1.0 g, 5.40 mM, 1 eq.) was suspended in sodium hydroxide (1 M, 30 mL) and refluxed until thin layer chromatography showed an absence of starting material. The solution was cooled to room temperature and acidified with hydrochloric acid solution (1 M), causing precipitation of the product. The precipitate was filtered off, dissolved in ethyl acetate, dried over anhydrous magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **168** (1.00 g, 91 %) as a white solid.

<sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ 7.71 (s, 2H), 6.35 (s, 2H), <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>) δ 166.0, 145.6, 129.7, 118.9, 117.6. LRMS (EI +ve) 206.98 (92%), 204.00 (100% [M]<sup>+</sup>), 189.98 (91%), 187.97 (99%), 161.98 (20%), 159.99 (32%), 124.01 (28%), 119.01 (26%), 100.00 (27%). HRMS (EI +ve) 204.9695 (calculated 204.9697 for C<sub>7</sub>H<sub>5</sub>Cl<sub>2</sub>NO<sub>2</sub>). IR (solid)  $cm^{-1}$  3497, 3376, 1680, 1584, 1555, 1532, 1501, 1418, 1331, 1250, 1223, 1208, 1194, 1180, 1130, 1113, 910, 897, 785, 764, 723, 677, 652, 646.

#### 6.1.49 Ethyl-(4-amino-3,5-chloro)-benzoate (173)



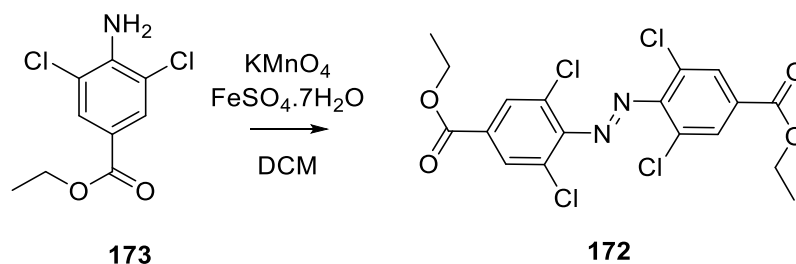
4-Amino-3,5-dichlorobenzoic acid (900 mg, 3.86 mM, 1 eq.) was dissolved in ethanol (10 mL) and sulfuric acid (2 mL) was added. This solution was refluxed for 14 hours then



cooled and neutralized with saturated sodium hydrogen carbonate solution and extracted twice with dichloromethane. The combined organic fractions were dried over magnesium sulfate, filtered, and then concentrated under reduced pressure to yield **173** (800 mg, 80%) as a low melting pale brown solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.87 (s, 2H), 4.87 (s, 2H), 4.32 (q,  $J = 7.1$  Hz, 2H), 1.37 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{DMSO-d}_6$ )  $\delta$  164.9, 143.9, 129.7, 120.2, 118.5, 61.1, 14.4. **LRMS** (EI -ve) 434.00 (23%), 432.00 (62%), 430.01 (68%), 283.27 (29%), 270.94 (36%), 265.94 (98%), 266.94 (100%  $[\text{M}+\text{Cl}]^-$ ), 255.24 (24%), 233.99 (53%  $[\text{M}]^-$ ), 232.00 (80%  $[\text{M}-\text{H}]^-$ ), 204.01 (37%), 112.98 (27%), 96.96 (27%), 80.91 (23%). **HRMS** (EI -ve) 233.0009 (calculated 233.0010 for  $\text{C}_9\text{H}_9\text{Cl}_2\text{NO}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3480, 3374, 2988, 1699, 1611, 1553, 1495, 1474, 1410, 1395, 1368, 1323, 1252, 1227, 1146, 1140, 1113, 1020, 924, 905, 895, 781, 756, 694.

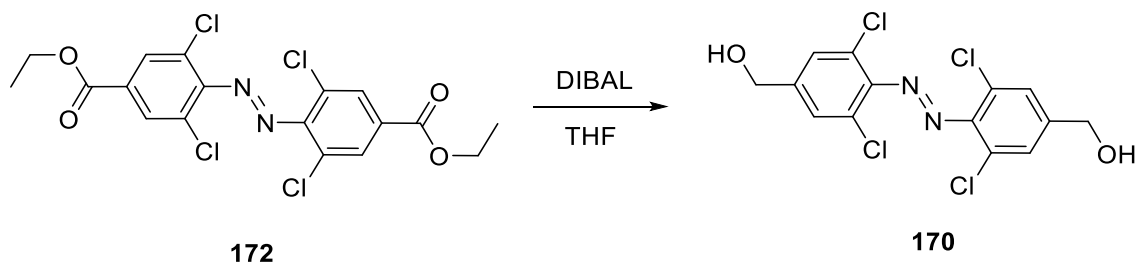
#### 6.1.50 Diethyl-4,4'-(diazene-1,2-diyl)-(E)-bis(3,5-dichlorobenzoate) (**172**)



Potassium permanganate (2.00 g) and iron (II) sulfate heptahydrate (2.00 g) were ground together to a fine power, then suspended in dichloromethane (25 mL) and ethyl(4-amino-3,5-dichloro)benzoate (0.70 g, 2.99 mM, 1 eq.) added. The suspension was refluxed overnight then cooled, filtered through celite and the solvent removed under reduced pressure. The resulting solid was recrystallised from ethyl acetate to yield **172** (86 mg, 13%) as a grey solid.

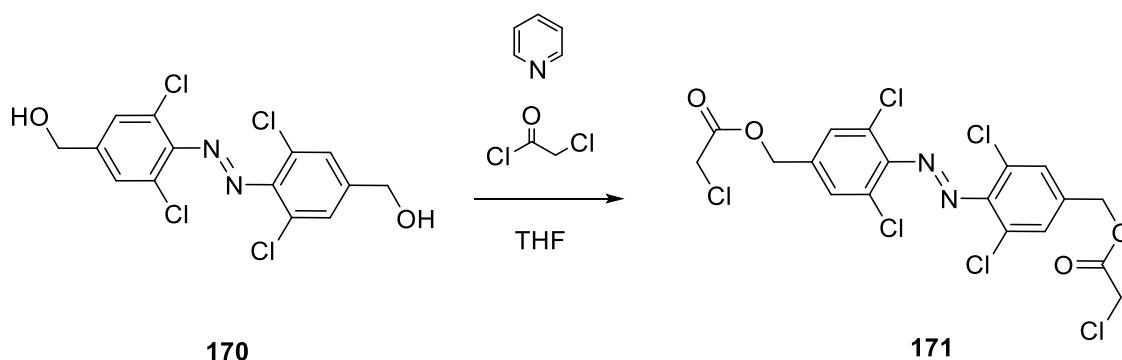
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.10 (s, 4H), 4.41 (q,  $J = 7.1$  Hz, 4H), 1.41 (t,  $J = 7.1$  Hz, 6H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  163.7, 150.3, 131.9, 130.5, 127.2, 62.1, 14.3. **LRMS** (EI -ve) 465.97 (50%), 463.97 (100%), 461.97 (75%  $[\text{M}]^-$ ), 313.08 (28%), 278.98 (58%), 276.98 (86%), 239.06 (38%), 172.83 (25%), 112.99 (23%), 96.96 (28%). **HRMS** (ES -ve) 461.9710 (calculated 461.9708 for  $\text{C}_{18}\text{H}_{14}\text{Cl}_4\text{N}_2\text{O}_4$ ). **IR** (solid)  $\text{cm}^{-1}$  1720, 1552, 1477, 1383, 1363, 1267, 1145, 1014, 929, 895, 864, 819, 814, 759.

## 6.1.51 (Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))dimethanol (170)



Diethyl-4,4'-[(diazene-1,2-diyl)bis(3,5-dichloro)]benzoate (140 mg, 0.302 mM, 1 eq.) was dissolved in dry tetrahydrofuran (30 mL) and the flask flushed with argon. Diisobutylaluminium hydride solution (1 M in toluene, 6 eq.) was slowly added and the reaction was stirred overnight. Saturated potassium sodium tartrate (20 mL) was added and the suspension was stirred vigorously for 1 hour. The organic layer was separated, dried over sodium sulfate, filtered and evaporated under reduced pressure. The solid residues were purified over silica gel eluting with 20-50% ethyl acetate/hexane to yield **170** (74 mg, 66%) as a pale orange solid.

$^1\text{H NMR}$  (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.56 (d,  $J = 0.6$  Hz, 4H), 4.68 (s, 4H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  145.9, 145.2, 127.0, 126.9, 62.0. **LRMS** (EI -ve) 380.94 (67%), 378.94 (100%), 376.94 (86%  $[\text{M}]^-$ ), 350.94 (29%), 349.94 (56%), 347.94 (46%). **HRMS** (EI -ve) 376.9413 (calculated 376.9418 for  $\text{C}_{14}\text{H}_9\text{Cl}_4\text{N}_2\text{O}_2$ ).

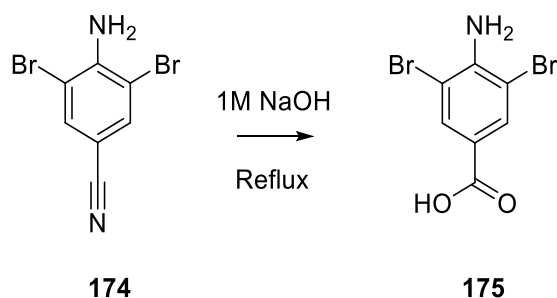
6.1.52 (*E*)-(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))-bis-(methylene)-bis-(2-chloroacetate) (171)

(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))dimethanol (30 mg, 0.054 mM, 1 eq.) was dissolved in tetrahydrofuran and the flask flushed with argon. Chloroacetyl chloride (100  $\mu\text{L}$ , excess) was added dropwise and the reaction was stirred for 20

minutes at room temperature. Pyridine (100  $\mu$ L, excess) was then added dropwise over 20 minutes and the solution stirred overnight. The solvent was removed under reduced pressure and the residue was partitioned between dichloromethane and sodium hydrogen carbonate. The organic layer was dried over sodium sulfate and filtered, then the solvent was removed under reduced pressure and the resulting solid was recrystallised from methanol to yield **171** (39 mg, 93%) as deep red crystals.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.71 (s, 4H), 5.21 (s, 4H), 4.16 (s, 4H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  167.0, 148.7, 137.9, 133.0, 116.2, 65.5, 40.7. **LRMS** (EI +ve) 501.98 (30%), 413.99 (27%), 264.00 (64%), 130.99 (100%). **HRMS** (EI +ve) 529.8932 (calculated 529.8928 for  $\text{C}_{18}\text{H}_{12}\text{N}_2\text{O}_4\text{Cl}_6$ ). **IR** (solid,  $\text{cm}^{-1}$ ) 3069, 2955, 2922, 2853, 1749, 1591, 1557, 1441, 1402, 1369, 1310, 1258, 1198, 1186, 1159, 1020, 982, 928, 897, 885, 854, 800, 789, 777, 723, 685, 650, 586, 557, 523, 444, 424, 419.

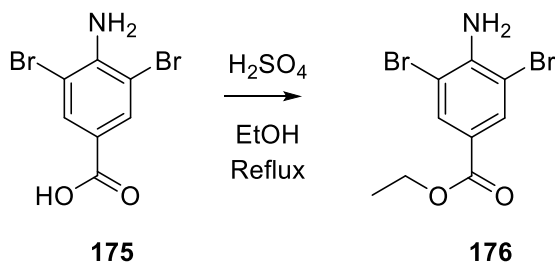
#### 6.1.53 4-Amino-3,5-dibromobenzoic acid (**175**)



4-Amino-3,5-dibromobenzonitrile (3.00 g, 10.8 mM, 1 eq.) was suspended in sodium hydroxide (1 M, 60 mL) and refluxed until thin layer chromatography showed an absence of starting material. The solution was cooled to room temperature and acidified with hydrochloric acid solution (1 M), causing precipitation of the product. The precipitate was filtered off, dissolved in ethyl acetate, dried over anhydrous magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **175** (3.20 g, 99%) as a white solid.

$^1\text{H NMR}$  (300 MHz,  $\text{DMSO-d}_6$ )  $\delta$  12.82 (s, 1H), 7.91 (s, 2H), 6.12 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{DMSO-d}_6$ )  $\delta$  165.7, 147.2, 133.6, 120.5, 106.8. **LRMS** (EI +ve) 296.89 (48%), 294.89 (100%), 292.89 (50%), 277.89 (18%), 277.89 (57%), 275.90 (30%), 219.0 (60%), 131.00 (37%). **HRMS** (EI +ve) 292.8687 (calculated 292.8687 for  $\text{C}_7\text{H}_5\text{Br}_2\text{NO}_2$ ) **IR** (solid)  $\text{cm}^{-1}$  3485, 3381, 2641, 2575, 2509, 1792, 1676, 1595, 1574, 1533, 1491, 1414, 1331, 1271, 1138, 1061, 1038, 928, 901, 766, 727, 677, 586, 546, 511.

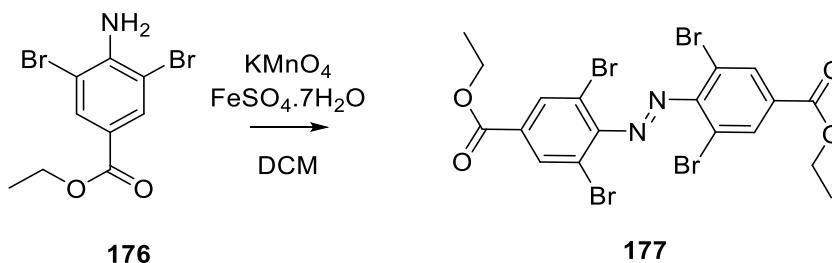
## 6.1.54 Ethyl 4-amino-3,5-dibromobenzoate (176)



4-Amino-3,5-dibromobenzoic acid (3.00 g, 9.31 mM, 1 eq.) was dissolved in ethanol (10 mL) and sulfuric acid (2 mL) this was refluxed for 14 hours. The solution was then neutralized with saturated sodium hydrogen carbonate solution and extracted twice with dichloromethane. The combined organic fractions were dried over magnesium sulfate, filtered, and then concentrated under reduced pressure to yield **176** (2.80 g, 85%) as a pale brown solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.07 (d,  $J = 3.2$  Hz, 2H), 4.98 (s, 2H), 4.32 (q,  $J = 7.1$  Hz, 2H), 1.37 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  164.6, 145.7, 133.4, 121.3, 107.4, 61.1, 14.4. **LRMS** (EI +ve) 324.93 (10%), 322.93 (19%), 320.93 (10%), 294.90 (36%), 284.31 (79%), 277.89 (50%,  $[\text{M}]^+$ ), 275.89 (24%), 256.27 (87%), 241.24 (25%), 219.01 (37%), 213.21 (42%), 207.06 (65%), 185.17 (31%), 171.16 (25%), 157.14 (20%), 129.10 (57%), 112.10 (22%), 111.11 (29%), 98.08 (100%), 87.11 (29%), 85.11 (29%), 73.04 (63%), 69.08 (52%), 57.08 (42%). **HRMS** (EI +ve) 320.8994 (calculated 320.9000 for  $\text{C}_9\text{H}_9\text{Br}_2\text{NO}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3431, 3321, 2990, 2920, 1796, 1713, 1605, 1543, 1481, 1466, 1449, 1391, 1364, 1298, 1254, 1138, 1113, 1059, 1018, 901, 862, 758, 729, 714, 671, 586, 554, 490, 467, 415.

## 6.1.55 Diethyl 4,4'-(diazene-1,2-diyl)bis(3,5-dibromo)benzoate (177)

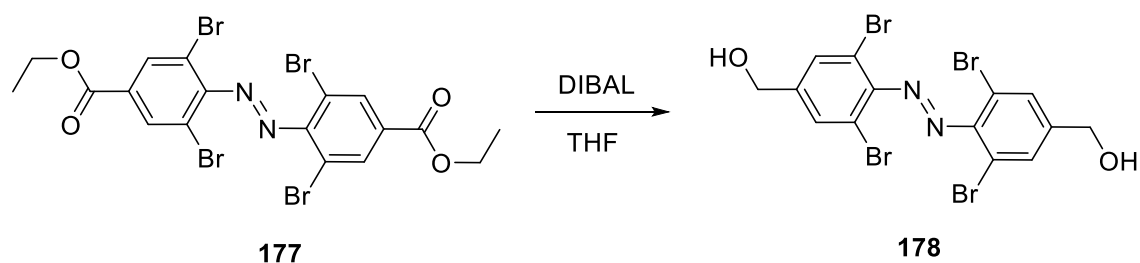


Potassium permanganate (10 g) and iron (II) sulfate heptahydrate (10 g) were ground together to a fine power, then suspended in dichloromethane (250 mL) and ethyl 4-amino-3,5-dibromobenzoate (2.5 g, 7.74 mM, 1 eq.) added. The suspension was refluxed overnight then cooled, filtered through celite and the solvent removed under reduced

pressure. The residue was recrystallised from ethyl acetate to yield **177** (86 mg, 13%) as a grey solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.27 (s, 4H), 4.35 (q,  $J = 7.1$  Hz, 4H), 1.42-1.25 (m, 6H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  163.5, 151.6, 134.4, 132.4, 115.6, 62.1, 14.3. **LRMS** (EI +ve) 644.77 (65%), 642.77 (100%), 640.77 (69%), 581.76 (28%), 579.76 (45%), 577.76 (29%), 504.86 (36%), 502.86 (40%), 487.84 (76%), 485.84 (80%), 483.84 (84%). **HRMS** (EI +ve) 638.7756 (calculated 638.7760 for  $\text{C}_{18}\text{H}_{14}\text{Br}_4\text{O}_4\text{H}$ ). **IR** (solid)  $\text{cm}^{-1}$  3321, 3090, 2990, 1715, 1605, 1545, 1470, 1433, 1368, 1248, 1130, 1111, 1013, 899, 862, 762, 739, 654, 469, 417.

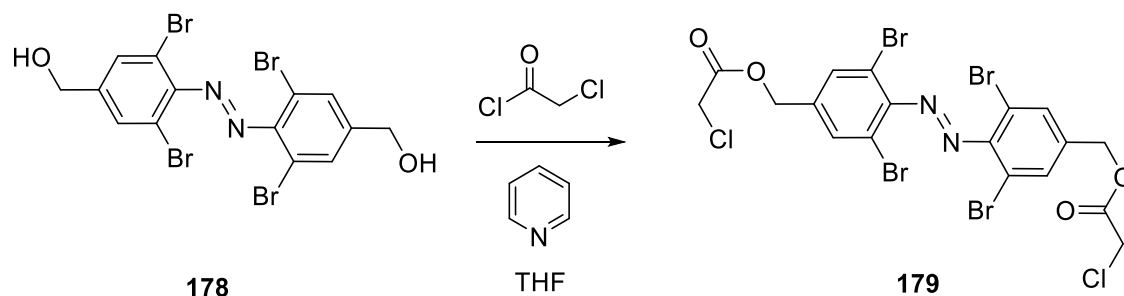
#### 6.1.56 (*E*)-(Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene))dimethanol (**178**)



Diethyl-4,4'-(diazene-1,2-diyl)bis(3,5-dibromobenzoate) (**177**) (120 mg, 0.187 mM, 1 eq.), was dissolved in dry tetrahydrofuran (6 mL) and the flask flushed with argon. Diisobutylaluminium hydride solution (1 M in toluene, 6 eq.) was slowly added and the reaction was stirred overnight. Saturated potassium sodium tartrate (12 mL) was added and the suspension was stirred vigorously for 1 hour. The organic layer was separated, dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The residue was purified over silica gel eluting with 20-50% ethyl acetate in hexane to yield **178** (104 mg, 63%) as a pale orange solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.77 (s, 4H), 4.67 (s, 4H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  147.3, 145.6, 131.0, 115.5, 61.8. **LRMS** (EI +ve) 559.80 (23%), 557.81 (38%), 55.80 (33%), 294.91 (41%), 292.90 (100%), 290.91 (55%), 277.90 (36%), 264.90 (40%), 246.88 (67%), 219.01 (29%), 155.97 (24%), 77.05 (21%), 69.00 (32%). **HRMS** (EI +ve) 532.7498 (calculated 533.7476 for  $\text{C}_{14}\text{H}_{10}\text{Br}_4\text{N}_2\text{O}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  2916, 2851, 2490, 2236, 2070, 1746, 1585, 1541, 1435, 1391, 1258, 1194, 1111, 1026, 976, 922, 899, 847, 799, 733, 492.

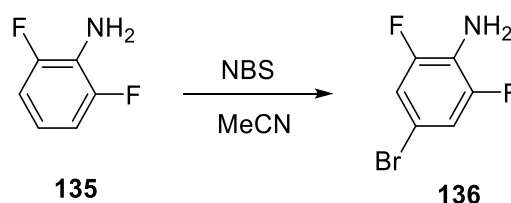
### 6.1.57 (*E*)-(Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene))bis(methylene) bis(2-chloroacetate) (**179**)



(*E*)-(diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene))dimethanol (30 mg, 0.054 mM, 1 eq.) was dissolved in tetrahydrofuran and the flask flushed with argon. Chloroacetyl chloride (100  $\mu\text{L}$ , excess), was added dropwise and the reaction was stirred for 20 minutes at room temperature. Pyridine (100  $\mu\text{L}$ , excess) was then added dropwise over 20 minutes and the solution stirred overnight. The solvent was then removed under reduced pressure and the residue was partitioned between dichloromethane and sodium hydrogen carbonate. The organic layer was dried over sodium sulfate and filtered, then the solvent was removed under reduced pressure and the resulting solid was recrystallised from methanol to yield **179** (38 mg, 99%) as deep red crystals.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.71 (s, 2H), 5.21 (s, 2H), 4.16 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  167.0, 148.7, 137.9, 133.0, 116.2, 65.5, 40.7. **HRMS** (NSI +ve) 706.6980 (calculated 706.6978 for  $\text{C}_{18}\text{H}_{12}\text{N}_2\text{O}_4\text{Br}_4\text{Cl}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3065, 2920, 2851, 1765, 1749, 1593, 1547, 1445, 1402, 1371, 1356, 1304, 1175, 1018, 970, 949, 924, 910, 864, 795, 770, 743, 723, 689, 567, 511, 500, 420, 413, 403.

### 6.1.58 4-Bromo-2,6-difluoroaniline (**136**)

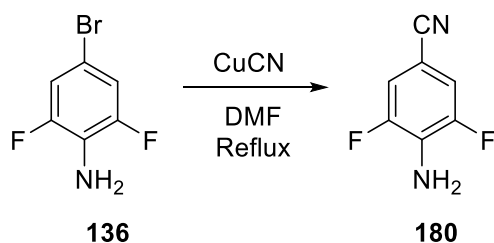


4-Bromo-2,6-difluoroaniline was synthesised according to a published procedure.<sup>61</sup> 2,6-Difluoroaniline (5.0 g, 38.7 mM, 1 eq.) was added to a solution of *N*-bromosuccinimide (6.9 g, 38.9 mM, 1 eq.) in acetonitrile (100 mL) and the resulting solution was stirred for 24 hours. Water (30 mL) was added and the mixture was extracted with hexane (3

x 30 mL). The organic layers were combined, dried over magnesium sulfate and filtered before the solvent was removed under reduced pressure to yield **136** (8.0 g, 100%) as a pink solid.

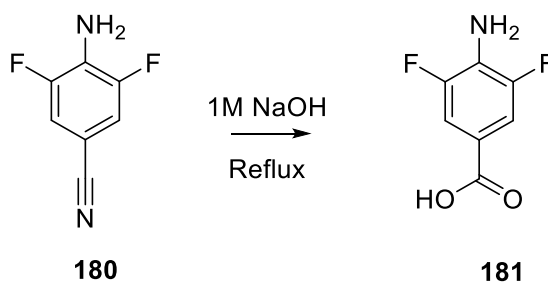
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.15-6.92 (m, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  151.8 (Dd,  $J = 9$  Hz,  $J = 9$  Hz), 123.5, 114.8 (dt,  $J = 13$ , 3 Hz), 107.2 (d,  $J = 12$  Hz).  $^{19}\text{F NMR}$  (282.24 MHz,  $\text{CDCl}_3$ )  $\delta$  -118, -119. **LRMS** (ESI -ve) 207.94 [ $\text{M}$ ] $^-$  (93%), 205.94 [ $\text{M}$ ] $^-$  (100%), 165.05 (70%), 118.9 (18%), 116.92 (24%). **HRMS** (ESI -ve) 205.9426 (Calculated 205.9417 for  $\text{C}_6\text{H}_4\text{BrF}_2\text{N}$ ). **IR** (solid)  $\text{cm}^{-1}$  3422, 3331, 1643, 1605, 1580, 1499, 1427, 1298, 1152, 966, 866, 839, 762, 716, 563, 509. **MP** 136-138  $^\circ\text{C}$ .

#### 6.1.59 4-Amino-3,5-difluorobenzonitrile (**180X**)



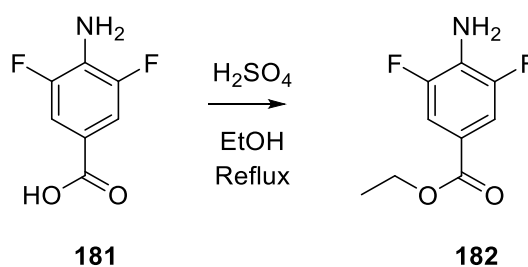
4-Bromo-2,6-difluoroaniline (2.00 g, 9.62 mM, 1 eq.) and copper cyanide (2.69 g, 28.8 mM, 3 eq.) were dissolved in dimethylformamide (20 mL) and refluxed for 14 hours following a published procedure.<sup>61</sup> The solution was cooled to room temperature and ammonia solution (12%, 100 mL) was added. This mixture was extracted with ethylacetate (5 x 100 mL) and the organic layers were combined, dried over magnesium sulfate filtered and the solvent removed under reduced pressure. The residue was purified over silica gel eluting with 33% hexane in dichloromethane to give **180** (0.50 g, 34%) as a pale yellow solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.06 (s, 2H), 4.28 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  150.5 (dd,  $J = 9$  Hz,  $J = 9$  Hz), 129.8 (t,  $J = 16$  Hz), 118.0 (t,  $J = 3$  Hz), 115.9-115.1 (m), 98.0 (t,  $J = 11$  Hz).  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -130.66. **LRMS** (EI -ve),  $m/z$  153.02 (100% [ $\text{M}$ ] $^-$ ), 133.01 (83%), 113 (66%). **HRMS** (EI -ve) 154.0345 (Calculated 154.0343 for  $\text{C}_7\text{H}_4\text{F}_2\text{N}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3484, 3484, 3358, 3351, 2228, 1638, 1574, 1530, 1445, 1348, 1335, 1277, 1146, 955, 866.

6.1.60 4-Amino-3,5-difluorobenzoic acid (**181**)

4-Amino-3,5-difluorobenzonitrile (256 mg, 1.66 mM, 1 eq.) was suspended in sodium hydroxide solution (1 M, 10 mL) and refluxed until thin layer chromatography showed an absence of starting material. The solution was cooled to room temperature and acidified with hydrochloric acid solution (1 M), causing precipitation of the product. The precipitate was filtered off and dissolved in ethyl acetate, dried over anhydrous magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **181** (268 mg, 93%) as an off-white solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.33 (dd,  $J = 7.2, 2.4$  Hz, 2H), 5.32 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  167.5 (t,  $J = 3$  Hz), 150.4 (d,  $J = 8$  Hz,  $J = 8$  Hz), 130.4 (t,  $J = 16$  Hz), 116.6 (t,  $J = 8$  Hz), 112.8–111.8 (m).  $^{19}\text{F NMR}$  (376 MHz, DMSO)  $\delta$  -131.61. **LRMS** (EI +ve) 173.03 (100%  $[\text{M}]^+$ ), 156.03 (95%), 128.03 (37%). **HRMS** (EI +ve) 173.0287 (calculated 173.0288 for  $\text{C}_7\text{H}_5\text{F}_2\text{NO}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  1680, 1636, 1580, 1499, 1423, 1333, 1298, 1238, 1148, 959, 864, 833, 758, 719, 563.

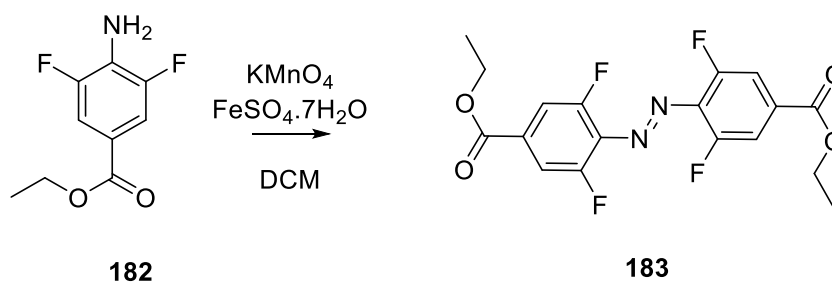
6.1.61 Ethyl(4-amino-3,5-fluoro)benzoate (**182**)

4-Amino-3,5-difluorobenzoic acid (516 mg, 9.31 mM) was dissolved in ethanol (100 mL) with sulfuric acid (2 mL) and the solution was refluxed for 14 hours. The solution was then neutralized with saturated sodium hydrogen carbonate solution and extracted twice with dichloromethane. The combined organic fractions were dried over magnesium sulfate, filtered, and then concentrated under reduced pressure to yield **182** (451 mg, 75 %) as a dark brown solid.



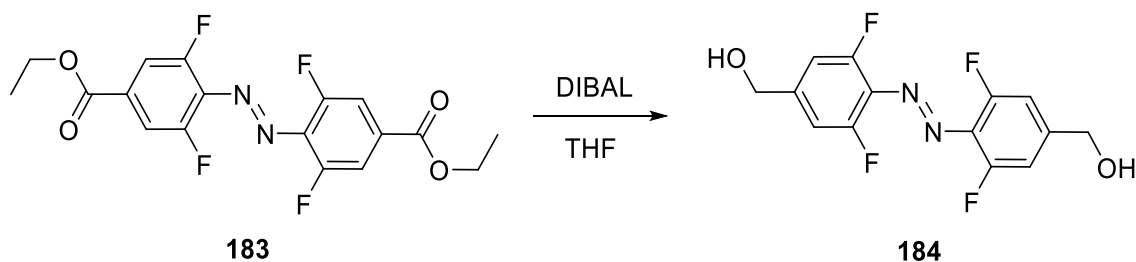
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.54 (dd,  $J = 7.2, 2.1$  Hz, 2H), 4.34 (q,  $J = 7.1$  Hz, 2H), 3.95 (s, 2H), 1.38 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  165.2, 151.2 (d,  $J = 1$  Hz,  $J = 1$  Hz), 128.8 (t,  $J = 2$  Hz), 118.5 (t,  $J = 2$  Hz), 112.6 (m), 61.1, 14.3.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -132.93. **LRMS** (EI +ve) 201.06 (87%), 173.03 (62%), 156.01 (100%), 149.03 (25%), 128.03 (48%). **HRMS** (EI +ve) 201.0596 (Calculated 201.0601 for  $\text{C}_9\text{H}_9\text{F}_2\text{NO}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3372, 1697, 1630, 1585, 1528, 1477, 1443, 1396, 1371, 1333, 1269, 1223, 1128, 1092, 1024, 939, 881, 758, 745, 719, 544, 519, 419.

#### 6.1.62 Diethyl 4,4'-(diazene-1,2-diyl)bis(3,5-difluorobenzoate) (**183**)



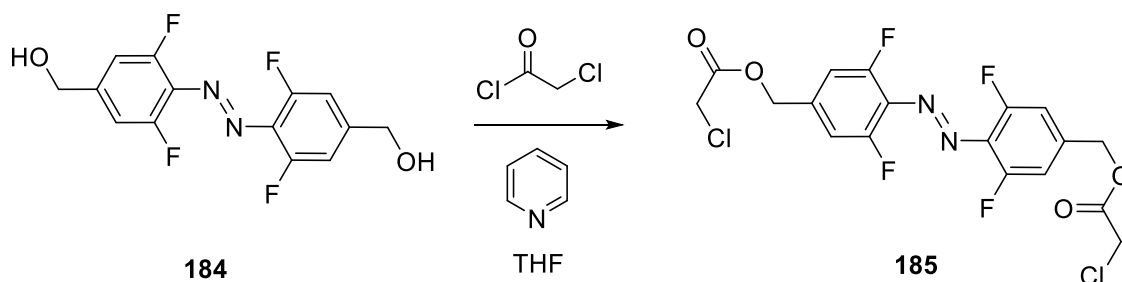
Potassium permanganate (2.00 g) and iron (II) sulfate heptahydrate (2.00 g) were ground together to a fine powder, then suspended in dichloromethane (100 mL) and 4-amino-3,5-difluorobenzoate (400 mg, 1.99 mM, 1 eq.). The suspension was refluxed overnight then cooled, filtered through celite and the solvent removed under reduced pressure. The residue was recrystallised from ethyl acetate to yield **183** (55 mg, 14%) as a pale pink solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.76 (dd,  $J = 10.5, 1.5$  Hz, 4H), 4.44 (q,  $J = 7.1$  Hz, 4H), 1.44 (t,  $J = 7.1$  Hz, 6H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  163.7, 156.7, 153.3, (d,  $J = 128$  Hz), 133 (m), 114.0 (d,  $J = 2$  Hz) 113.0 (d,  $J = 2$  Hz), 62.2, 14.2.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -119.49. **LRMS** (EI +ve) 398.09 (90%  $[\text{M}]^+$ ), 382.1 (25%), 353.06 (23%), 213.03 (100%), 201.06 (41%), 185.04 (37%), 173.03 (29%), 156.02 (94%), 128.03 (24%), 101.02 (47%). **HRMS** (EI +ve) 398.0895 (calculated 398.0890 for  $\text{C}_{18}\text{H}_{14}\text{F}_4\text{N}_2\text{O}_4$ ). **IR** (solid)  $\text{cm}^{-1}$  3372, 3250, 3090, 2961, 2936, 1767, 1722, 1678, 1626, 1576, 1528, 1470, 1433, 1398, 1371, 1331, 1231, 1194, 1088, 1049, 1018, 885, 862, 797, 770, 748, 544, 511, 500.

6.1.63 (*E*)-(Diazene-1,2-diylbis(3,5-difluoro-4,1-phenylene))dimethanol (**183**)

Diethyl 4,4'-(diazene-1,2-diyl)-bis-(3,5-difluorobenzoate) (50 mg, 0.126 mM, 1 eq.) was dissolved in dry tetrahydrofuran (5 mL) and the flask flushed with argon. Diisobutylaluminium hydride solution (0.76 mL, 1 M in toluene, 6 eq.) was slowly added and the reaction was stirred overnight. Saturated potassium sodium tartrate (10 mL) was added and the suspension was stirred vigorously for 1 hour. The organic layer was separated, dried over sodium sulfate, filtered and evaporated under reduced pressure. The residues was purified over silica gel eluting with 20-50% ethyl acetate in hexane to yield **184** (35 mg, 89%) as a pale solid.

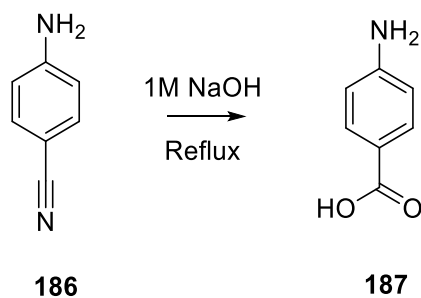
<sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>OD) δ 7.06 (d, *J* = 10.2 Hz, 3H *E*-isomer), 6.86 (d, *J* = 8.8 Hz, 1H *Z*-isomer), 4.56 (d, *J* = 6.3 Hz, 1H *E*-isomer), 4.45 (s, 1H *Z*-isomer), <sup>13</sup>C NMR (75 MHz, CD<sub>3</sub>OD) δ 157.1 (d, *J* = 1 Hz), 153.8, (d, *J* = 1 Hz), 148.2 (m), 147.9, 129.9, 109.8 (d, *J* = 1 Hz) 109.5 (d, *J* = 1 Hz), 62.3 (mixtures of isomers). <sup>19</sup>F NMR (376 MHz, MeOD) δ -122.99. LRMS (EI +ve) 315.07 (45% [M]<sup>+</sup>), 314.07 (70% [M-H]<sup>+</sup>), 213.05 (57%), 201.6 (37%), 173.04 (33%), 172.05 (57%), 171.03 (96%), 157.03 (39%), 156.03 (78%), 143.049 (37%), 142.04 (23%), 125.02 (100%). HRMS (EI +ve) 315.0744 (calculated 315.0757 for C<sub>14</sub>H<sub>10</sub>F<sub>4</sub>N<sub>2</sub>O<sub>2</sub>). IR (solid) cm<sup>-1</sup> 2860, 2610, 2475, 1701, 1609, 1580, 1551, 1499, 1423, 1422, 1317, 1306, 1244, 1211, 1180, 1109, 1078, 1016, 806, 739, 679, 629, 606, 527, 474.

6.1.64 (*E*)-(Diazene-1,2-diylbis(3,5-difluoro-4,1-phenylene))bis(methylene)-bis-(2-chloroacetate) (**185**)

(*E*)-(Diazene-1,2-diylbis-(3,5-difluoro-4,1-phenylene))dimethanol (23 mg, 0.073 mM, 1 eq.) was dissolved in tetrahydrofuran and the flask flushed with argon. Chloroacetyl chloride (100  $\mu$ L, excess), was added dropwise and the reaction was stirred for 20 minutes at room temperature. Pyridine (100  $\mu$ L, excess) was then added dropwise over 20 minutes and the solution stirred overnight. The solvent was then removed under reduced pressure and the residue was partitioned between dichloromethane and sodium hydrogen carbonate. The organic layer was dried over sodium sulfate and filtered, then the solvent was removed under reduced pressure and the resulting solid was recrystallised from methanol to yield **185** (30 mg, 87%) as deep red crystals.

$^1\text{H NMR}$  (600 MHz,  $\text{CDCl}_3$ )  $\delta$ , 7.08 (d,  $J = 9.1$  Hz, 3H *E*-isomer), 6.89 (d,  $J = 7.7$  Hz, 1H *Z*-isomer), 5.23 (s, 3H), 5.14 (s, 1H), 4.16 (s, 3H), 4.12 (s, 1H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  166.9, 157.3, 153.8, 139.8, 139.67, 112.1, 111.7, 65.9, 65.6 (mixture of isomers).  $^{19}\text{F NMR}$  (565 MHz,  $\text{CDCl}_3$ )  $\delta$  -118.11, -118.12, -119.55, -119.57 (possible multiplets of isomers). **LRMS** (EI +ve) 468.01 (23%), 466.01 (48%  $[\text{M}]^+$ ), 249.01 (21%), 247.01 (100%), 125.02 (78%). **HRMS** (EI +ve) 466.0113 (calculated 466.0110 for  $\text{C}_{18}\text{H}_{12}\text{N}_2\text{O}_4\text{Cl}_2\text{F}_4$ ) **IR** (solid)  $\text{cm}^{-1}$  2961, 2920, 2849, 1759, 1630, 1580, 1441, 1410, 1373, 1314, 1186, 1171, 1076, 1053, 1020, 980, 943, 928, 864, 851, 799, 785, 750, 737, 712, 602, 586, 567, 529, 446.

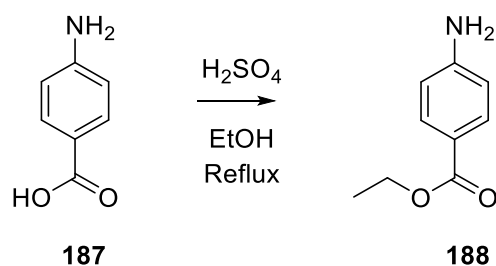
#### 6.1.65 4-Amino-3-benzoic acid (**187**)



4-Amino-3-benzonitrile (3.00 g, 25.0 mM, 1 eq.) was suspended in sodium hydroxide (1 M, 150 mL) and refluxed until thin layer chromatography showed an absence of starting material. The solution was cooled to room temperature and acidified with hydrochloric acid solution (1 M), causing precipitation of the product. The precipitate was filtered off and dissolved in ethyl acetate, dried over anhydrous magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **187** (1.89 g, 55%) as an off-white solid.

$^1\text{H NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  10.10 (s, 1H), 7.91 (d,  $J = 8.2$  Hz, 2H), 7.22 (d,  $J = 8.3$  Hz, 2H).  $^{13}\text{C NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  167.3, 142.0, 131.4, 126.3, 120.5. **LRMS** (EI +ve) 137.04 (100%  $[\text{M}]^+$ ), 120 (100%). **HRMS** (EI +ve) 214.9584 (Calculated 214.9582 for  $\text{C}_7\text{H}_6\text{BrNO}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3426, 3325, 3090, 2980, 1715, 1670, 1605, 1541, 1481, 1470, 1431, 1391, 1364, 1300, 1250, 1155, 1130, 1109, 1057, 1018, 899, 864, 824, 758, 731, 681, 631, 559, 532, 503, 490, 473, 451.

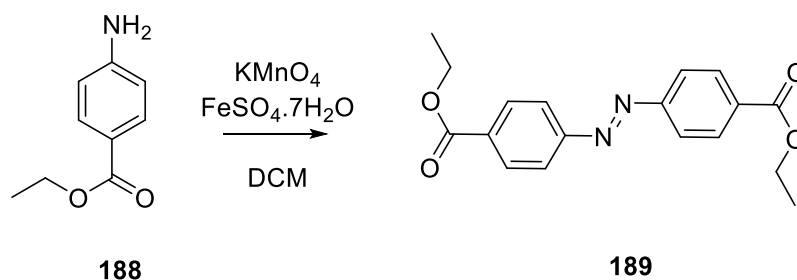
#### 6.1.66 Ethyl (4-amino-benzoate) (188)



4-Aminobenzoic acid (1.00 g, 6.06 mM, 1 eq.) was dissolved in ethanol (60 mL) and sulfuric acid (2 mL) and refluxed for 14 hours. The solution was then neutralized with saturated sodium hydrogen carbonate solution and extracted twice with dichloromethane. Combined organic fractions were dried over magnesium sulfate, filtered, and then concentrated under reduced pressure to yield **188** (1.20 g, 100%) as a pale purple solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.88 (dd,  $J = 6.8, 1.9$  Hz, 2H), 6.66 (dd,  $J = 6.8, 1.9$  Hz, 2H), 4.34 (q,  $J = 7.1$  Hz, 2H), 4.07 (s, 2H), 1.38 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  166.7, 150.7, 131.6, 120.1, 113.8, 60.4, 14.5. **LRMS** (EI +ve) 165.07 (100%  $\text{M}^+$ ), 137.04 (75%), 121.05 (21%), 120.03 (100%), 92.05 (82%), 84.94 (90%), 82.94 (100%), 65.04 (49%). **HRMS** (EI -ve) 165.0794 (calculated 165.0790 for  $\text{C}_9\text{H}_7\text{NO}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3375, 3310, 2980, 2926, 2851, 1956, 1711, 1603, 1520, 1464, 1449, 1406, 1385, 1366, 1261, 1213, 1171, 1094, 1009, 868, 858, 766, 696, 542.

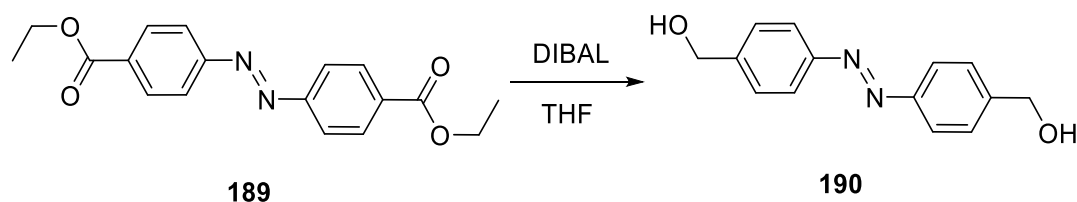
#### 6.1.67 Diethyl-4,4'-(diazene-1,2-diyl)benzoate (189)



Potassium permanganate (4.00 g) and iron (II) sulfate heptahydrate (4.00 g) were ground together to a fine power and suspended in dichloromethane (250 mL). 4-Aminobenzoate (800 mg, 4.84 mM, 1 eq.) was added and the suspension was refluxed overnight. The reaction was then cooled, filtered through celite and the solvent removed under reduced pressure. The resulting solid was recrystallised from ethyl acetate to yield **189** (172 mg, 22%) as a pale orange solid.

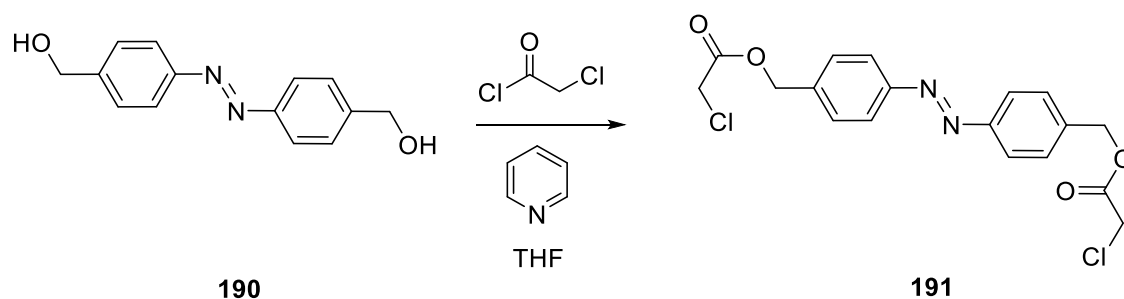
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.36-8.12 (m, 4H), 8.12-7.89 (m, 4H), 4.67-3.65 (m, 4H), 1.45 (dd,  $J = 9.3, 4.6$  Hz, 6H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  166.0, 154.9, 132.76, 130.7, 122.9, 61.4, 14.4. **LRMS** (EI +ve) 326.13 (93%  $\text{M}^+$ ), 281.09 (20%), 177.07 (20%), 149.06 (100%). **HRMS** (EI +ve) 326.1272 (calculated 326.1267 for  $\text{C}_{18}\text{H}_{18}\text{N}_2\text{O}_4$ ). **IR** (solid)  $\text{cm}^{-1}$  3375, 3310, 2980, 2926, 2851, 1956, 1711, 1603, 1520, 1464, 1449, 1406, 1385, 1366, 1261, 1213, 1171, 1094, 1009, 868, 858, 766, 696, 542.

#### 6.1.68 (*E*)-(Diazene-1,2-diylbis(4,1-phenylene))dimethanol (**190**)



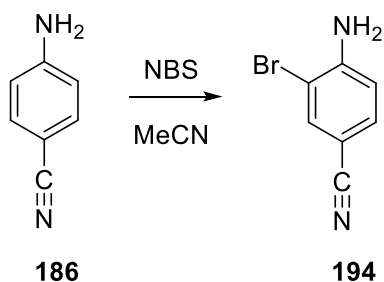
Diethyl-4,4'-(diazene-1,2-diylbenzoate) (**189**) (172 mg, 0.527 mM, 1 eq.) was dissolved in dry tetrahydrofuran (30 mL) and the flask flushed with argon. Diisobutylaluminium hydride solution (1 M in toluene, 6 eq.) was slowly added and the reaction was stirred overnight. Saturated potassium sodium tartrate (30 mL) was added and the suspension was stirred vigorously for 1 hour. The organic layer was separated, dried over sodium sulfate, filtered and evaporated under reduced pressure. The solid residues were purified over silica gel eluting with 20-50% ethyl acetate/hexane to yield **190** (78 mg, 61%) as a pale solid.

$^1\text{H NMR}$  (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.91 (d,  $J = 8.3$  Hz, 4H), 7.56 (d,  $J = 8.3$  Hz, 4H), 4.60 (s, 4H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  151.9, 144.9, 127.1, 126.9, 122.4, 120.2, 63.3. **LRMS** (EI +ve) 242.10 (100%  $[\text{M}]^+$ ), 135.05 (33%), 107.05 (67%), 89.04 (36%). **HRMS** (EI +ve) 242.1050 (calculated 242.1055 for  $\text{C}_{14}\text{H}_{14}\text{N}_2\text{O}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3275, 3175, 2916, 2851, 1931, 1601, 1495, 1456, 1414, 1375, 1342, 1304, 1288, 1202, 1153, 1105, 1020, 1007, 851, 829, 719, 696, 637, 569, 519.

6.1.69 *(E)*-(Diazene-1,2-diylbis(4,1-phenylene))bis(methylene)-bis(2-chloroacetate) (**191**)

*(E)*-(Diazene-1,2-diylbis(4,1-phenylene))dimethanol (50 mg, 0.207 mM, 1 eq.) was dissolved in tetrahydrofuran and the flask flushed with argon. Chloroacetyl chloride (100  $\mu$ L, excess), was added dropwise and the reaction was stirred for 20 minutes at room temperature then pyridine (100  $\mu$ L, excess) was then added dropwise over a further 20 minutes and the solution stirred overnight. The solvent was then removed under reduced pressure and the residue was partitioned between dichloromethane and sodium hydrogen carbonate. The organic layer was dried over sodium sulfate, filtered, and the solvent was removed under reduced pressure. The resulting solid was recrystallised from methanol to yield **191** (70 mg, 86%) as deep red crystals.

$^1\text{H NMR}$  (600 MHz,  $\text{CDCl}_3$ )  $\delta$ , 7.95 (d,  $J = 8.4$  Hz, 4H), 7.55 (d,  $J = 8.5$  Hz, 4H), 5.32 (s, 4H), 4.16 (s, 4H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  167.2 152.5, 137.9, 129.1, 123.2, 67.3, 40.9. **LRMS** (EI +ve) 396.05 (37%), 394.05 (59%  $[\text{M}]^+$ ), 211.03 (43%), 90.05 (26%), 89.04 (60%), 83.09 (29%). **HRMS** (EI +ve) 242.1050 (calculated 242.1055 for  $\text{C}_{14}\text{H}_{14}\text{N}_2\text{O}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  2949, 2922, 1744, 1603, 1499, 1449, 1412, 1373, 1315, 1256, 1209, 1200, 1190, 1159, 1107, 999, 980, 966, 955, 928, 885, 876, 856, 835, 777, 718, 677, 640, 590, 559, 529, 513.

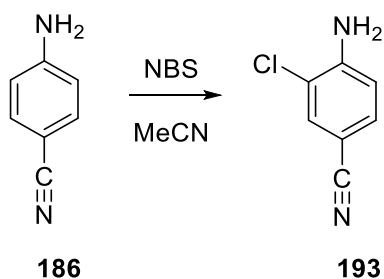
6.1.70 2-Bromo-4-aminobenzonitrile (**194**)

4-Aminobenzonitrile (1.00 g, 8.46 mM, 1 eq.) was dissolved in acetonitrile (30 mL), *N*-bromosuccinimide (1.51 g, 8.46 mM, 1 eq.) was added and the resulting solution was

stirred overnight. The solvent was removed under reduced pressure and the residue was dissolved in dichloromethane and washed with potassium hydroxide. The organic layer was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to give **194** (1.65 g, 99%) as an off-white solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.67 (d,  $J = 1.3$  Hz, 1H), 7.43-7.25 (m, 1H), 6.76 (d,  $J = 8.4$  Hz, 1H), 4.72 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  148.3, 136.4, 132.5, 118.8, 114.8, 107.8, 100.9. **LRMS** (ES -ve) 196.94 (71%  $[\text{M}]^-$ ), 194.94 (72%  $[\text{M}]^-$ ), 115.02 (100%), 80.91 (43%), 78.91 (44%). **HRMS** (EI +ve) 194.9559 (calculated 194.9958 for  $\text{C}_7\text{H}_4\text{N}_2\text{Br}$ ). **IR** (solid)  $\text{cm}^{-1}$  3451, 3333, 2228, 1761, 1703, 1609, 1528, 1481, 1402, 1304, 1202, 1165, 1053, 891, 872, 814, 739, 716, 640, 590, 474, 449, 426, 409, 403.

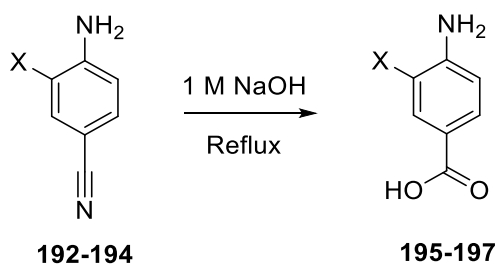
#### 6.1.71 4-Amino-3-chlorobenzonitrile (**193**)



4-Aminobenzonitrile (3.00 g, 25.0 mM, 1 eq.) was dissolved in acetonitrile (30 mL), *N*-chlorosuccinimide (3.39 g, 25.0 mM, 1 eq.) was added and the solution was stirred overnight. The solvent was removed under reduced pressure and the residue dissolved in dichloromethane and washed with potassium hydroxide. The organic layer was dried over magnesium sulfate filtered and concentrated. The solid residue was purified by silica gel eluting with 30% ethyl acetate in hexane to yield **193** (2.84 g, 74%) as an off-white solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.52 (d,  $J = 1.8$  Hz, 1H), 7.33 (dd,  $J = 8.4, 1.8$  Hz, 1H), 6.75 (d,  $J = 8.4$  Hz), 4.61 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  147.0, 133.3, 132.0, 118.9, 118.5, 115.1, 100.8. **LRMS** (EI +ve) 154.01 (47%), 152.01 (100%  $[\text{M}]^+$ ), 142.16 (24%), 125.00 (18%). **HRMS** (EI +ve) 152.0138 (calculated 152.0141 for  $\text{C}_7\text{H}_5\text{N}_2\text{Cl}$ ). **IR** (solid)  $\text{cm}^{-1}$  3474, 3350, 3208, 2220, 1902, 1769, 1630, 1597, 1506, 1416, 1335, 1317, 1200, 1161, 1078, 1043, 881, 812, 712, 675, 583, 523, 517, 430.

## 6.2 General method for the synthesis of 4-amino-3-halobenzoic acids



Variouly halogenated substituted 4-aminobenzonitriles were suspended in sodium hydroxide (1 M, 3-6 mL/mM) and refluxed until thin layer chromatography showed an absence of starting material. The solution was cooled to room temperature and acidified with hydrochloric acid solution (1 M), causing precipitation of the product. The precipitate was filtered off then dissolved in ethyl acetate, dried over anhydrous magnesium sulfate, filtered and the solvent removed under reduced pressure.

### 4-Amino-3-bromobenzonitrile (197)

Reacting **194** (1.30 g, 6.60 mM, 1 eq.) yielded **197** (1.32 g, 93%) as an off-white solid.

$^1\text{H NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  12.41 (s, 1H), 7.94-7.75 (m, 1H), 7.63 (dd,  $J$  = 8.5, 1.2 Hz, 1H), 6.79 (d,  $J$  = 8.5 Hz, 1H), 6.13 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  166.8, 150.4, 134.4, 130.5, 119.2, 114.5, 106.3. **LRMS** (ESI -ve) 428.86 (20%), 426.86 (37%), 424.86 (20%), 215.94 (31%), 213.94 (29%), 96.95 (32%), 80.91 (100%). **HRMS** (EI -ve) 214.9584 (calculated 214.9582 for  $\text{C}_7\text{H}_6\text{BrNO}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3426, 3325, 3090, 2980, 1715, 1670, 1605, 1541, 1481, 1470, 1431, 1391, 1364, 1300, 1250, 1155, 1130, 1109, 1057, 1018, 899, 864, 824, 758, 731, 681, 631, 559, 532, 503, 490, 473, 451.

### 4-Amino-3-chlorobenzoic acid (196)

Reacting **193** (2.5 g, 16.40 mM, 1 eq.) yielded **196** (2.32 g, 83 %) as a white solid.

$^1\text{H NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  12.41 (s, 1H), 7.71 (d,  $J$  = 1.9 Hz, 1H), 7.59 (dd,  $J$  = 8.5, 1.9 Hz, 1H), 6.79 (d,  $J$  = 8.5 Hz, 1H), 6.18 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  167.0, 149.4, 131.1, 130.0, 118.7, 116.4, 114.6. **LRMS** (EI +ve) 173.00 (23%), 171.01 (97%  $[\text{M}]^+$ ), 156.00 (36%), 154.01 (100%). **HRMS** (EI +ve) 171.0093 (calculated 171.0087 for  $\text{C}_7\text{H}_6\text{NO}_2\text{Cl}$ ). **IR** (solid)  $\text{cm}^{-1}$  3499, 3399, 1665, 1614, 1557, 1558, 1512, 1435, 1406 1331, 1283, 1256, 1157, 1067, 1040, 918, 880, 829, 822, 766, 704, 631, 559, 540.



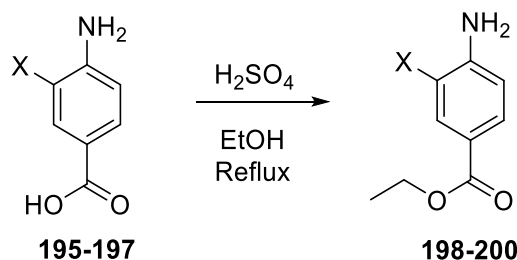
#### 4-Amino-3-fluorobenzoic acid (195)

Reacting **192** (2.0 g, 14.7 mM, 1 eq.) yielded **195** (2.20 g, 96 %) as a white solid.

$^1\text{H NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  12.37 (s, 1H), 7.56-7.41 (m, 2H), 6.77 (t,  $J = 8.6$  Hz, 1H), 5.99 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  167.3 (d,  $J = 1$  Hz), 149.7, (d,  $J = 116$  Hz), 141.9, (d,  $J = 2$  Hz) 127.4, (d,  $J = 1$  Hz), 117.8, (d,  $J = 7.5$  Hz) 116.3 (d,  $J = 22.5$  Hz) 115.2 (d,  $J = 7.5$  Hz).

$^{19}\text{F NMR}$  (376 MHz, DMSO- $d_6$ )  $\delta$  -135.81. **LRMS** (EI +ve) 155.04 (100%  $[M]^+$ ), 137.04 (97%), 110.04 (30%), 88.05 (27%). **HRMS** (EI +ve) 155.0383 (calculated 155.0383 for  $\text{C}_7\text{H}_6\text{NO}_2\text{F}$ ). **IR** (solid)  $\text{cm}^{-1}$  3514, 3414, 1672, 1634, 1574, 1530, 1456, 1412, 1298, 1254, 1198, 1152, 1088, 1043, 930, 887, 827, 760, 635, 559, 548, 434, 430.

### 6.3 General method for the synthesis of ethyl(4-amino-3 halobenzoates)



Variously substituted 4-aminobenzoic acids (1 eq.) were dissolved in ethanol (10 mL/mM) with sulfuric acid (2 mL) and refluxed for 14 hours. The solution was then neutralized with saturated sodium hydrogen carbonate solution and extracted twice with dichloromethane. Combined organic fractions were dried over magnesium sulfate, filtered, and the solvent removed under reduced pressure.

#### 6.3.1 (4-Amino-3-bromo)benzoate (200)

Reacting **197** (1.00 g,) yielded **200** (1.07 g, 95%) as a pale brown solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.13 (d,  $J = 1.9$  Hz, 1H), 7.81 (dd,  $J = 8.4, 1.9$  Hz, 1H), 6.75 (d,  $J = 8.4$  Hz, 1H), 4.54 (s, 2H), 4.34 (q,  $J = 7.1$  Hz, 2H), 1.38 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  165.6, 148.1, 134.5, 130.3, 121.1, 114.3, 107.9, 60.7, 14.4. **LRMS** (EI +ve) 244.98 (76%), 242.9831 (77%), 218.99 (23%), 216.96 (54%), 214.96 (56%), 199.94 (100%), 197.94 (98%), 171.96 (26%), 169.96 (26%), 90.03 (26%). **HRMS** (EI +ve) 242.9892 (calculated 242.9895 for  $\text{C}_9\text{H}_{10}\text{BrNO}_2$ ). **IR** (solid)  $\text{cm}^{-1}$  3470, 3360, 2976, 2905, 1688, 1626,

1593, 1510, 1481, 1416, 1364, 1337, 1275, 1231, 1153, 1123, 1103, 1018, 893, 822, 797, 758, 702, 679, 629, 552, 527, 434, 415.

### 6.3.2 (4-Amino-3-chloro)benzoate (199)

Reacting **196** (1.00 g) reacted to yield **199** (1.06 g, 91%) as a pale brown solid.

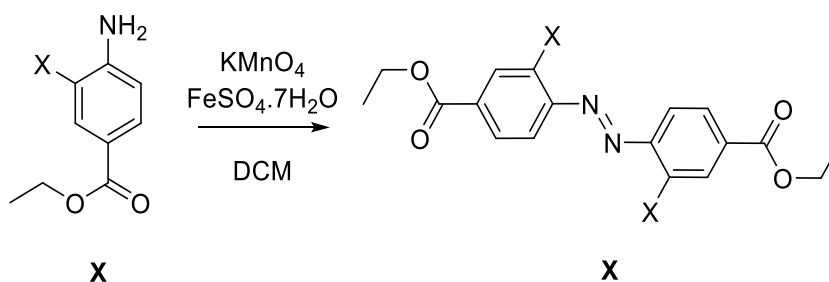
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.98 (d, *J* = 1.9 Hz, 1H), 7.79 (dd, *J* = 8.4, 1.9 Hz, 1H), 6.76 (d, *J* = 8.4 Hz, 1H), 4.79-4.35 (m, 2H), 4.32 (d, *J* = 7.1 Hz, 2H), 1.39 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 165.8, 147.0, 131.2, 129.6, 120.8, 118.2, 114.4, 60.7, 14.4. LRMS (EI +ve) 201.04 (26%), 199.04 (87% [M]<sup>+</sup>), 171.01 (77%), 156.00 (70%), 153.99 (100%). HRMS (EI +ve) 199.0400 (calculated 199.0401 for C<sub>9</sub>H<sub>10</sub>NO<sub>2</sub>Cl<sub>1</sub>). IR (solid) cm<sup>-1</sup> 3487, 3356, 3208, 2980, 2901, 1688, 1624, 1593, 1508, 1474, 1423, 1391, 1366, 1329, 1281, 1242, 1159, 1134, 1113, 1040, 1022, 899, 870, 833, 760, 706, 629, 552, 534, 440, 419.

### 6.3.3 Ethyl(4-amino-3,5-fluoro)benzoate (198)

Reacting **195** (2.00 g, 13 mM, 1 eq.) yielded **198** (2.17 g, 92 %) as a dark brown solid.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.64 (m, 2H), 6.74 (m, 1H), 4.31 (q, *J* = 7.2 Hz, 2H), 4.18 (s, 2H), 1.35 (t, *J* = 7.2 Hz, 3H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ, 166.0, 151.9, 139.4 (d, *J* = 1 Hz), 126.8, 120.3, (d, *J* = 1 Hz), 116.5 (d, *J* = 3 Hz), 115.3 (d, *J* = 1 Hz), 60.7, 14.4. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -136.10. LRMS (EI +ve) 183.07 (76% [M]<sup>+</sup>), 155.04 (37%), 138.02 (100%), 110.04 (28%). HRMS (EI +ve) 183.0694 (calculated 183.0696 for C<sub>9</sub>H<sub>10</sub>NO<sub>2</sub>F). IR (solid) cm<sup>-1</sup> 3402, 3331, 3215, 2988, 1688, 1634, 1611, 1585, 1520, 1445, 1364, 1315, 1285, 1246, 1200, 1155, 1115, 1094, 1076, 1024, 941, 887, 826, 760, 667, 631, 575, 563, 536, 480, 469.

## 6.4 General method for the synthesis of haloazobenzene ethyl esters



Potassium permanganate (1.2 g/mM) and iron (II) sulfate heptahydrate (1.2 g/mM) were ground together to a fine powder, then suspended in dichloromethane and various halo-substituted ethyl(4-aminobenzoates) were added. The suspension was refluxed

overnight then cooled, filtered through celite and the solvent removed under reduced pressure. The resulting solid was recrystallised from ethyl acetate.

#### 6.4.1 Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-bromobenzoate) (203)

Reacting **200** yielded **103** (81 mg, 10%) as a pale orange solid.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.47 (s, 1H), 8.10 (d, J = 8.3 Hz, 1H), 7.80 (d, J = 8.3 Hz, 1H), 4.45 (d, J = 7.1 Hz, 2H), 1.58 (t, J = 7.1 Hz, 3H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 164.7, 151.9, 135.2, 134.2, 129.4, 126.1, 118.3, 61.8, 14.3. LRMS (EI +ve) 485.95 (33%), 483.95 (72%), 481.95 (38%), 256.98 (61%), 254.98 (62%), 228.97 (93%), 226.97 (100%), 200.94 (56%), 198.94 (59%), 182.93 (24%), 149.06 (42%), 143.95 (25%), 142.95 (27%), 103.02 (27%). HRMS (EI +ve) 481.9486 (calculated 481.9477 for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub>Br<sub>2</sub>). IR (solid, cm<sup>-1</sup>) 3416, 3335, 3225, 2990, 2951, 2930, 2891, 1913, 1676, 1628, 1595, 1570, 1512, 1476, 1441, 1385, 1364, 1308, 1273, 1169, 1126, 1111, 1076, 1020, 880, 847, 766, 696, 638, 500.

#### 6.4.2 Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-chlorobenzoate) (202)

Reacting **199** yielded **202** (122 mg, 7%) of as a pale orange solid.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.26 (d, J = 1.6 Hz, 1H), 8.02 (dd, J = 8.4, 1.6 Hz, 1H), 7.79 (d, J = 8.4 Hz, 1H), 4.43 (q, J = 7.1 Hz, 2H), 1.43 (t, J = 7.1 Hz, 3H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 164.8, 151.04, 136.0, 134.1, 132.1, 128.6, 118.0, 61.8, 14.3. LRMS (EI +ve) 396.04 (25%), 394.05 (394.05 [M]<sup>+</sup>), 211.02 (35%), 183.02 (37%), 175.99 (44%), 168.99 (52%), 149.99 (51%), 145.00 (23%), 114.00 (100%). HRMS (EI +ve) 394.0467 (calculated 394.0487 for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub>Cl<sub>2</sub>). IR (solid) cm<sup>-1</sup> 3487, 3358, 2974, 1817, 1722, 1628, 1595, 1562, 1510, 1477, 1450, 1389, 1314, 1267, 1240, 1219, 1138, 1109, 1057, 1022, 905, 866, 847, 770, 721, 708, 633, 596, 503, 478.

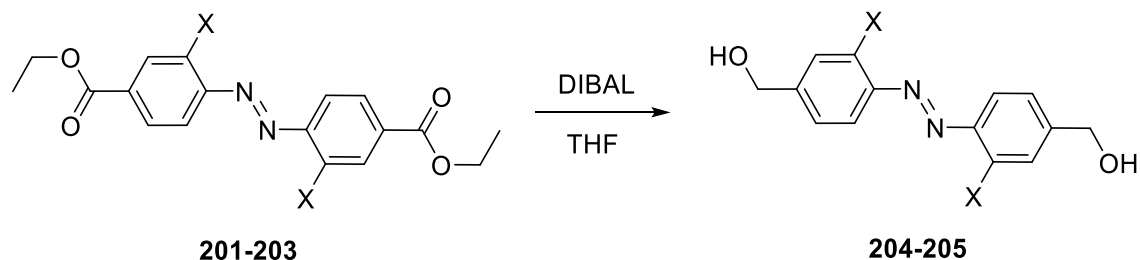
#### 6.4.3 Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-fluorobenzoate) (201)

Reacting **198** yielded **201** (191 mg, 9.7 %) of as a pale orange solid as a mixture of rotamers and isomers.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 9.12 (d, J = 1.3 Hz, 0.18H), 9.01 (d, J = 1.3 Hz, 0.30H), 8.85 (t, J = 1.4 Hz, 0.50H), 8.53-8.38 (m, 0.5H), 8.32 (dd, J = 9.1, 3.0 Hz, 0.5H), 8.10 (dd, J = 10.3, 1.6 Hz, 0.18H), 8.01-7.76 (m, 0.39H), 7.26 (s, 0.39H), 4.46 (dq, J = 25.3, 7.1 Hz, 4H), 1.54-1.25 (m, 6H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 164.9, 161.7, 158.3, 143.1, 135.1, 135.0, 133.0, 130.2, 125.7, 118.7, 118.4, 117.8, 61.8, 14.3. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -121.76, -122.49. LRMS (EI +ve) 362.11 (90%), 342.10 (27%), 324.11 (23%), 297.07 (26%), 195.05 (100%),

167.05 (97%), 139.02 (25%), 122.02 (20%), 94.02 (24%), 83.03 (28%). **HRMS** (EI +ve) 362.1076 (calculated 362.1078 for C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub>F<sub>2</sub>).

## 6.5 General method for the reduction of haloazobenzene diethyl esters to diols



Azobenzene ethyl esters were dissolved in dry tetrahydrofuran (30 mL/mM) and the flask flushed with argon. Diisobutylaluminium hydride solution (1 M in toluene, 6 eq.) was slowly added and the reaction was stirred overnight. Saturated potassium sodium tartrate (60 mL/mM) was added and the suspension was stirred vigorously for 1 hour. The organic layer was separated, dried over sodium sulfate, filtered and evaporated under reduced pressure. The solid residue was purified over silica gel eluting with 20-50% ethyl acetate in hexane.

### 6.5.1 (E)-(diazene-1,2-diylbis(3-bromo-4,1-phenylene))dimethanol (206)

Reacting **203** yielded **206** (62.7 mg, 94%) as a pale white solid.

<sup>1</sup>H NMR (300 MHz, DMSO-d<sub>6</sub>) δ 7.88-7.80 (m, 2H), 7.66 (d, *J* = 8.3 Hz, 2H), 7.49 (dd, *J* = 8.3, 1.7 Hz, 2H), 4.61 (d, *J* = 5.5 Hz, 4H) (rotamers or isomers). <sup>13</sup>C NMR (151 MHz, DMSO-d<sub>6</sub>) δ 150.6, 149.3, 148.0, 145.0, 131.6, 131.1, 126.9, 126.7, 125.9, 118.1, 117.7, 116.4, 62.2, 62. (isomers or rotamers). **LRMS** (EI +ve) 401.92 (24%), 399.92 (69%), 397.92 (37% M<sup>+</sup>), 263.99 (34%), 218.99 (72%), 214.96 (88%), 212.96 (100%), 186.96 (41%), 168.95 (45%), 166.95 (81%), 130.99 (48%), 78.05 (32%). **HRMS** (EI +ve) 397.9270 (calculated 397.9265 for C<sub>14</sub>H<sub>12</sub>N<sub>2</sub>O<sub>2</sub>Br<sub>2</sub>). **IR** (solid) cm<sup>-1</sup> 2916, 2841, 2415, 2330, 1593, 1558, 1443, 1402, 1379, 1350, 1306, 1294, 1223, 1130, 1096, 1038, 935, 889, 880, 814, 721, 667, 588, 554, 469, 459, 432, 411.

### 6.5.2 (E)-(diazene-1,2-diylbis(3-chloro-4,1-phenylene))dimethanol (205)

Reacting **202** yielded **205** (50 mg, 18%) as a pale white solid.

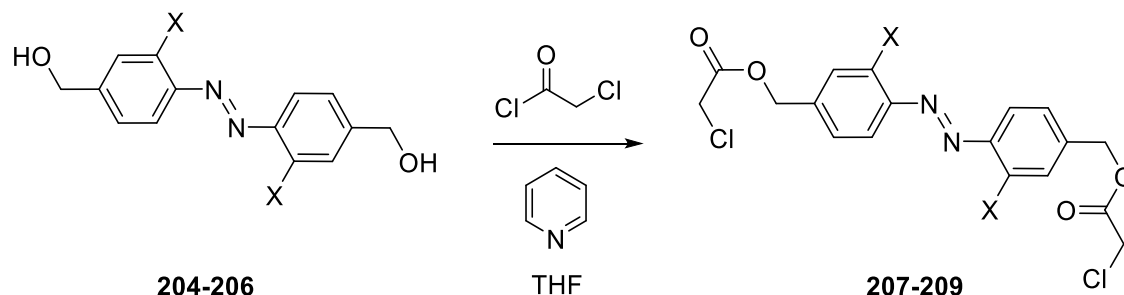
<sup>1</sup>H NMR (600 MHz, DMSO-d<sub>6</sub>) δ 7.72-7.64 (m, 4H), 7.51-7.42 (m, 3H), 7.14 (d, *J* = 8.1 Hz, 1H), 6.47 (d, *J* = 8.1 Hz, 1H), 5.52 (t, *J* = 5.8 Hz, 2H), 5.35 (t, *J* = 5.8 Hz, 1H), 4.61 (d, *J* = 5.7 Hz, 4H), 4.44 (d, *J* = 5.7 Hz, 2H) (rotamers or isomers). <sup>13</sup>C NMR (151 MHz, DMSO-d<sub>6</sub>) δ 149.4, 149.3, 147.1, 145.0, 135.0, 128.5, 128.1, 126.2, 126.0, 125.9, 118.4, 117.9, 62.3, 62.0 (isomers or rotamers). LRMS (EI +ve) 310.03 (5% [M]<sup>+</sup>), 263.99 (100%), 118.99 (29%), 99.99 (34%), HRMS (EI +ve) 310.0277 (calculated 310.0276 for C<sub>14</sub>H<sub>12</sub>N<sub>2</sub>O<sub>2</sub>Cl<sub>2</sub>).

### 6.5.3 (E)-(Diazene-1,2-diylbis(3-fluoro-4,1-phenylene))dimethanol (204)

Reacting **201** yielded **204** (47 mg, 30%) as a pale white solid

<sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>OD) δ 7.73 (t, *J* = 8.0 Hz, 1H), 7.29 (ddd, *J* = 10.0, 8.0, 1.2 Hz, 2H), 4.68 (s, 2H). <sup>13</sup>C NMR (75 MHz, CD<sub>3</sub>OD) δ 160.3 (d, *J* = 127 Hz), 148.5 (d, *J* = 7.5 Hz), 139.4, 122.1, (d, *J* = 7.5 Hz), 117.2, 114.3 (d, *J* = 22 Hz), 62.7. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -123.66. LRMS (ESI +ve) (378.09 (22%), 316.08 (21%), 315.07 (100%) [M+H]<sup>+</sup>. HRMS (EI +ve) 279.0947 (calculated 279.0945 for C<sub>14</sub>H<sub>13</sub>N<sub>2</sub>O<sub>2</sub>F<sub>2</sub>).

### 6.6 General method for the chloroacetylation of azobenzene diols



Azobenzene diols **204-206** were dissolved in tetrahydrofuran and the flask flushed with argon. Chloroacetyl chloride (100  $\mu$ L, excess), was added dropwise and the reaction was stirred for 20 minutes at room temperature then pyridine (100  $\mu$ L, excess) was then added dropwise over a further 20 minutes and the solution stirred overnight. The solvent was then removed under reduced pressure and the residue was partitioned between dichloromethane and sodium hydrogen carbonate. The organic layer was dried over sodium sulfate and filtered, then the solvent was removed under reduced pressure and the resulting solid was recrystallised from methanol.

### 6.6.1 (E)-(Diazene-1,2-diylbis(3-bromo-4,1-phenylene))bis(methylene)-bis-(2-chloroacetate) (209)

Reacting **206** yielded **209** (70 mg, 99%) as deep red crystals as a mixture of isomers and rotamers.

<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 7.77 (dd, *J* = 9.8, 5.0 Hz, 12H), 7.62 (d, *J* = 1.5 Hz, 2H), 7.40 (dd, *J* = 8.3, 1.8 Hz, 6H), 7.08 (dd, *J* = 8.1, 1.7 Hz, 2H), 6.35 (d, *J* = 8.1 Hz, 2H), 5.26 (s, 12H), 5.11 (s, 4H), 4.14 (d, *J* = 10.8 Hz, 12H), 4.09 (s, 4H) (rotamers or isomers). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 167.0, 152.2, 149.4, 139.8, 136.1, 133.3, 132.7, 127.8, 126.6, 118.6, 117.8, 117.0, 66.3, 66.3, 66.3, 66.0, 40.7, 29.7 (rotamers or isomers). LRMS (EI +ve) 555.86 (18%), 553.86 (52%), 551 (57%), 549.87 (23% [M]<sup>+</sup>), 292.93 (21%), 290.94 (88%), 288.94 (70%), 262.93 (44%), 260.93 (34%), 169.95 (21%), 168.95 (100%), 188.95 (95%), 89.04 (24%), 84.94 (37%), 82.95 (53%). HRMS (EI +ve) 549.8671 (calculated 549.8697 for C<sub>18</sub>H<sub>14</sub>N<sub>2</sub>O<sub>4</sub>Br<sub>2</sub>Cl<sub>2</sub>). IR (solid) cm<sup>-1</sup> 2961, 2918, 2851, 1738, 1597, 1562, 1547, 1477, 1460, 1412, 1377, 1314, 1258, 1229, 1192, 1167, 1130, 1043, 964, 922, 887, 839, 779, 745, 733, 692, 667, 613, 575, 517.

### 6.6.2 (E)-(Diazene-1,2-diylbis(3-chloro-4,1-phenylene))bis(methylene)-bis-(2-chloroacetate) (208)

Reacting **205** yielded **208** (51 mg, 66%) as deep red crystals.

<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 7.78 (d, *J* = 8.2 Hz, 2H), 7.59 (s, 2H), 7.35 (d, *J* = 8.2 Hz, 2H), 5.26 (s, 4H), 4.15 (s, 4H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 167.0, 148.6, 139.6, 136.2, 130.3, 127.1, 118.4, 66.4, 40.7. LRMS (EI +ve) 465.96 (24%), 463.97 (48%), 461.97 (37% [M]<sup>+</sup>), 246.98 (50%), 244.99 (76%), 216.98 (31%), 125.00 (40%), 123.00 (100%), 89.04 (39%). HRMS (EI +ve) 461.9714 (calculated 461.708 for C<sub>18</sub>H<sub>14</sub>N<sub>2</sub>O<sub>4</sub>Cl<sub>4</sub>). IR (solid) cm<sup>-1</sup> 3013, 2963, 1744, 1651, 1601, 1562, 1531, 1481, 1462, 1412, 1400, 1312, 1258, 1180, 1130, 1057, 989, 966, 928, 899, 881, 837, 793, 779, 739, 704, 677, 619, 584, 569, 521, 484, 440, 405.

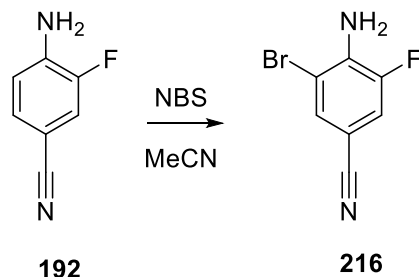
### 6.6.3 (E)-(Diazene-1,2-diylbis(3-fluoro-4,1-phenylene))bis(methylene) bis(2-fluoroacetate) (207)

Reacting **204** yielded **207** (30 mg, 65%) as deep red crystals.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.86-7.74 (m, 2H), 7.36-7.17 (m, 4H), 5.27 (s, 4H), 4.15 (d, *J* = 3.3 Hz, 4H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 167.0, 161.1, 159.4, 140.6, 123.8, 121.8, 118.3, 116.7, 66.4, 40.7. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -123.66. δ LRMS (ES+ve) 445.13 (30%),

433.04 (64%), 393.04 (42%), 383.02 (45%), 371.11 (37%), 337.06 (60%). **HRMS** (EI +ve) 431.0398 (calculated 431.0377 for C<sub>18</sub>H<sub>15</sub>N<sub>2</sub>O<sub>4</sub>F<sub>2</sub>Cl<sub>2</sub>).

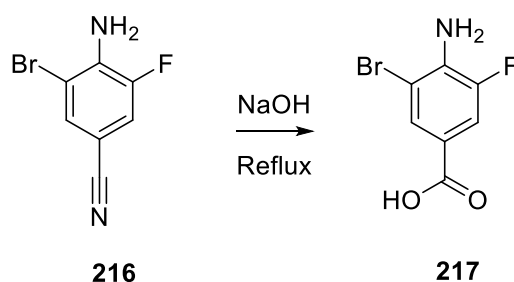
#### 6.6.4 4-Amino-3-bromo-5-fluorobenzonitrile (216)



4-Amino-3-fluorobenzonitrile (5.00 g 36.7 mM, 1 eq.), was dissolved in acetonitrile (100 mL), *N*-bromosuccinimide (9.81 g, 55.1 mM, 1.5 eq.) was added and the solution was stirred overnight. The solvent was removed under reduced pressure and the residue dissolved in dichloromethane and washed with potassium hydroxide. The organic layer was dried over magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **216** (7.84 g, 99%) an off-white solid.

<sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 7.55 (t, *J* = 1.4 Hz, 1H), 7.30-7.23 (m, 1H), 4.71 (s, 2H). <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 150.2, 148.6, 138.5 (d, *J* = 1 Hz), 132.2 (d, *J* = 1 Hz), 117.7, (d, *J* = 1 Hz), 108.4 (d, *J* = 1 Hz), 100.3 (d, *J* = 1 Hz). <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -129.08. **LRMS** (EI +ve) 215.95 (98%), 213.95 (100% [M]<sup>+</sup>). **HRMS** (ESI -ve) 213.9547 (Calculated 213.9542 for C<sub>7</sub>H<sub>4</sub>N<sub>2</sub>FBr). **IR** (solid) cm<sup>-1</sup> 3464, 3339, 2228, 1628, 1557, 1504, 1429, 1317, 1248, 1126, 1076, 961, 868, 762, 719, 604, 536, 459, 420, 407.

#### 6.6.5 4-Amino-3-bromo-5-fluorobenzoic acid (217)

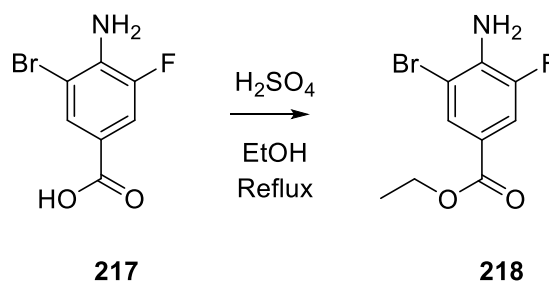


4-Amino-3-bromo-5-fluorobenzonitrile (7.00 g, 33 mM, 1 eq.) was suspended in sodium hydroxide (1 M, 80 mL) and refluxed until thin layer chromatography showed an absence of starting material. The solution was cooled to room temperature and acidified with hydrochloric acid solution (1 M), causing precipitation of the product.

The precipitate was filtered off then dissolved in ethyl acetate, dried over anhydrous magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **217** (6.73 g, 88%) as a white solid.

$^1\text{H NMR}$  (250 MHz, DMSO- $d_6$ )  $\delta$  7.80-7.71 (m, 1H), 7.52 (dd,  $J = 11.6, 1.8$  Hz, 1H), 6.21 (s, 2H).  $^{13}\text{C NMR}$  (63 MHz, DMSO- $d_6$ )  $\delta$  165.6, 148.95 (d,  $J = 145$  Hz), 139.4 (d,  $J = 2$  Hz), 129.6, (d) 117.9 (d,  $J = 1$  Hz), 115.0, (d,  $J = 2$  Hz), 106.4 (d,  $J = 1$  Hz).  $^{19}\text{F NMR}$  (376 MHz, DMSO)  $\delta$  -129.14. **LRMS** (EI +ve) 234.95 (98%), 232.95 (100%  $[\text{M}]^+$ ), 217.94 (71%), 215.95 (72%), 213.99 (30%), 187.95 (23%), 168.99 (26%). **HRMS** (ESI -ve) 232.9489 (Calculated 232.9488 for  $\text{C}_7\text{H}_5\text{NO}_2\text{FBr}$ ). **IR** (solid)  $\text{cm}^{-1}$  3503, 3395, 1686, 1603, 1560, 1514, 1439, 1416, 1348, 1323, 1267, 1211, 1094, 949, 893, 835, 764, 683, 606, 550, 461, 455, 428, 420, 411, 403.

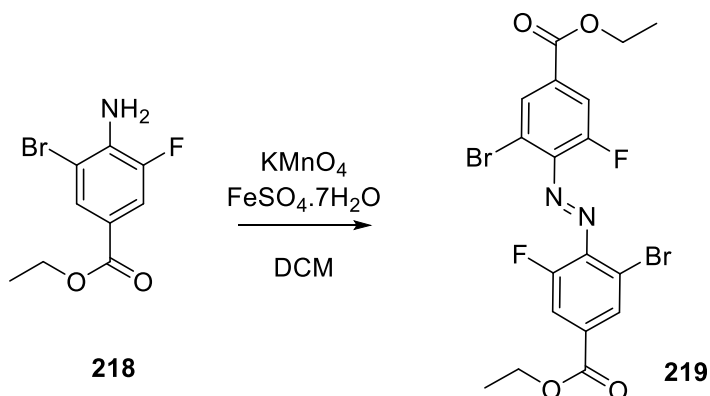
#### 6.6.6 Ethyl 4-amino-3-bromo-5-fluorobenzoate (**218**)



4-Amino-3-bromo-5-fluorobenzoic acid (6.5 g 27.9 mM, 1 eq.) was dissolved in ethanol (100 mL) and sulfuric acid (5 mL) and the resulting solution was refluxed for 14 hours. The solution was then neutralized with saturated sodium hydrogen carbonate solution and extracted twice with dichloromethane. The combined organic fractions were dried over magnesium sulfate, filtered, and then concentrated under reduced pressure to yield **218** (6.29 g, 86%) as an off white solid.

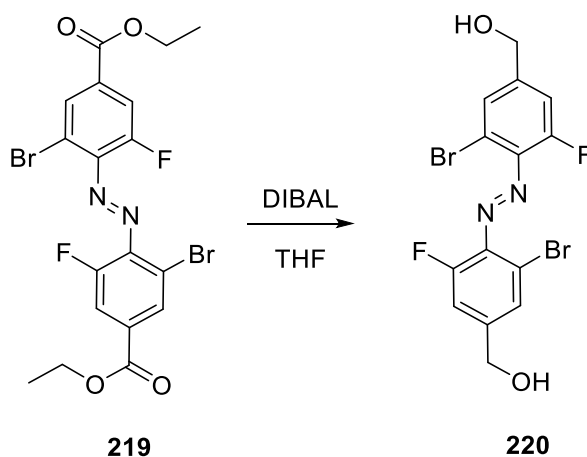
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.96 (t,  $J = 1.6$  Hz, 1H), 7.65 (dd,  $J = 11.2, 1.6$  Hz, 1H), 4.77-4.35 (m, 2H), 4.32 (d,  $J = 7.1$  Hz, 2H), 1.39 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  164.8 (d,  $J = 1$  Hz), 149.3, (d,  $J = 310$  Hz), 140.1 (d,  $J = 12$  Hz), 129.8, 117.3, (d,  $J = 1$  Hz) 115.3 (d,  $J = 2$  Hz), 107.4 (d,  $J = 75$  Hz), 61.1, 14.4.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -130.92. **LRMS** (ESI -ve) 261.96 (100%), 259.96 (95%), 233.95 (43%), 231.95 (34%), 126.98 (30%), 124.96 (40%). **HRMS** (EI +ve) 260.9799 (Calculated 260.9801 for  $\text{C}_9\text{H}_9\text{NO}_2\text{FBr}$ ).



6.6.7 Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-bromo-5-fluorobenzoate) (**219**)

Potassium permanganate (12.0 g) and iron (II) sulfate heptahydrate (12.0 g) were ground together to a fine powder, suspended in dichloromethane (250 mL) and ethyl 4-amino-3-bromo-5-fluorobenzoate (6.00 g, 23 mM, 1 eq.) was added. The suspension was refluxed overnight then cooled, filtered through celite and the solvent removed under reduced pressure. The resulting solid was recrystallised from ethyl acetate to yield **219** (0.40 g, 6.7%) as a pale orange solid.

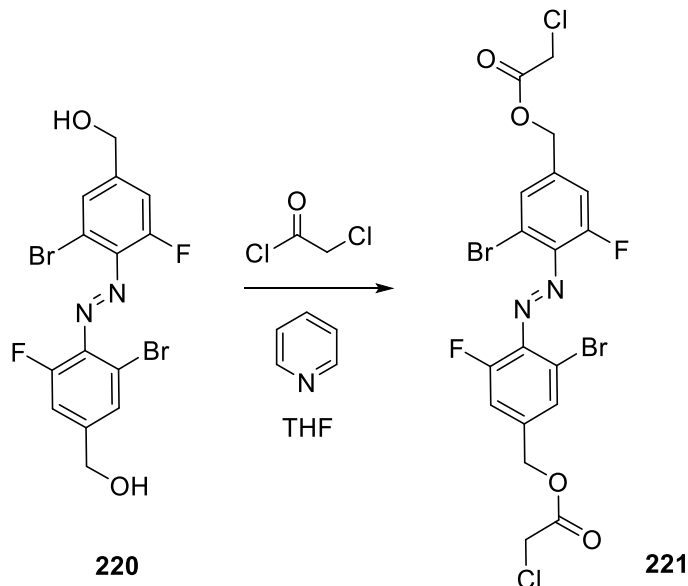
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.23 (t, *J* = 1.6 Hz, 1H), 7.86 (dd, *J* = 10.6, 1.6 Hz, 1H), 4.43 (q, *J* = 7.1 Hz, 2H), 1.43 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 163.7 (d, *J* = 2 Hz), 151.4, (d, *J* = 260 Hz), 142.5, (d, *J* = 38 Hz), 133.3, 130.3 (d, *J* = 1 Hz), 121.3 (d), 117.9 (d, *J* = 2 Hz), 62.2, 14.3. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -122.06, -130.93. LRMS (EI +ve) 522.07 (17%), 520.07 (33%), 518.07 (20% [M]<sup>+</sup>), 281.13 (46%), 273.04 (48%), 216.99 (27%), 207.09 (100%). HRMS (ESI -ve) 518.9370 (Calculated 518.9367 for C<sub>18</sub>H<sub>15</sub>N<sub>2</sub>O<sub>2</sub>F<sub>2</sub>Br<sub>2</sub>). IR (solid) cm<sup>-1</sup> 3067, 2986, 1711, 1566, 1481, 1449, 1412, 1393, 1368, 1292, 1225, 1175, 1119, 1096, 1020, 974, 910, 895, 876, 856, 806, 766, 748, 691, 544, 529, 424.

6.6.8 (E)-(Diazene-1,2-diylbis(3-bromo-5-fluoro-4,1-phenylene))dimethanol (**220**)

Diethyl 4,4'-(diazene-1,2-diyl)-(E)-bis(3-bromo-5-fluorobenzoate) (300 mg, 0.577 mM, 1 eq.) was dissolved in dry tetrahydrofuran (30 mL) and the flask flushed with argon. Diisobutylaluminium hydride solution (1 M in toluene, 3.34 mL, 3.46 mM, 6 eq.) was slowly added and the reaction was stirred overnight. Saturated potassium sodium tartrate (30 mL) was added and the suspension was stirred vigorously for 1 hour. The organic layer was separated, dried over sodium sulfate, filtered and evaporated under reduced pressure. The solid residue was purified over silica gel eluting with 20-50% ethyl acetate in hexane to yield **220** (170 mg, 68%) as a off white solid.

<sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>OD) δ 7.66 (s, 2H), 7.32 (d, *J* = 10.9 Hz, 2H), 4.70 (s, 4H). <sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD) δ 147.9, 144.9, 143.8, 137.7, 131.2 (d, *J* = 15 Hz), 116.2, 63.2, (d, *J* = 15 Hz). <sup>19</sup>F NMR (376 MHz, MeOD) δ -125.08. LRMS (ESI -ve) 437.90 (33%), 435.91 (57%), 433.91 [M]<sup>-</sup> (33%), 363.97 (20%), 232.96 (90%), 230.96 (100%), 217.95 (58%), 216.96 (50%), 215.95 (58%), 202.94 (27%), 186.94 (86%), 184.96 (79%), 96.04 (39%), 93.00 (43%). HRMS (ESI -ve) 433.9075 (Calculated 433.9077 for C<sub>14</sub>H<sub>10</sub>N<sub>2</sub>O<sub>2</sub>Br<sub>2</sub>F<sub>2</sub>).

#### 6.6.9 (E)-(Diazene-1,2-diylbis(3-bromo-5-fluoro-4,1-phenylene))bis(methylene)-bis(2-chloroacetate) (**221**)

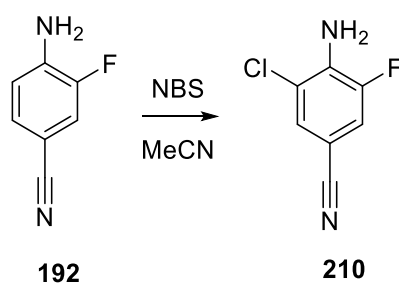


(E)-(Diazene-1,2-diylbis(3-bromo-5-fluoro-4,1-phenylene))dimethanol (17 mg, 0.039 mM, 1 eq.) was dissolved in tetrahydrofuran and the flask flushed with argon. Chloroacetyl chloride (100  $\mu$ L, excess), was added dropwise and the reaction was stirred for 20 minutes at room temperature. Pyridine (100  $\mu$ L, excess) was then added dropwise over 20 minutes and the solution stirred overnight. The solvent was then removed

under reduced pressure and the residue was partitioned between dichloromethane and sodium hydrogen carbonate. The organic layer was dried over sodium sulfate and filtered, then the solvent was removed under reduced pressure and the resulting solid was recrystallised from methanol to yield **221** (9.2 mg, 40%) as deep red crystals.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.59 (s, 2H), 7.31-7.18 (m, 2H), 5.25 (s, 4H), 4.17 (s, 4H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{CDCl}_3$ )  $\delta$  167.0, 152.7, 151.0, 139.0, 128.5, 122.4, 116.2 (d,  $J = 1$  Hz), 68.0, 65.7, 40.7.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -115.22, -121.82. **LRMS** (EI +ve) 612.84 (31%), 610.84 (33%), 590.86 (26%), 588.86 (30%,  $[\text{M}]^+$ ), 454.29 (40%), 414.28 (27%), 413.27 (100%). **HRMS** (ESI -ve) 586.8604 (Calculated 586.8587 for  $\text{C}_{18}\text{H}_{13}\text{N}_2\text{O}_4\text{Br}_2\text{F}_2\text{Cl}_2$ ).

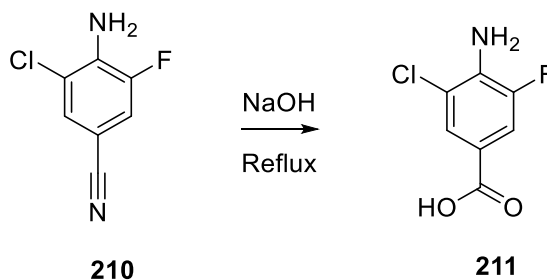
#### 6.6.10 4-Amino-3-chloro-5-fluorobenzonitrile (**210**)



4-Amino-3-fluorobenzonitrile (5.00 g 36.7 mM, 1 eq.) was dissolved in acetonitrile (100 mL) and *N*-chlorosuccinimide (7.35 g, 55.1 mM, 1.5 eq.) was added. The solution was stirred overnight. The solvent was removed under reduced pressure and the residue dissolved in dichloromethane and washed with potassium hydroxide. The organic layer was dried over magnesium sulfate, filtered and concentrated to yield **210** (5.58 g, 89%) as an off-white solid.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.40 (t,  $J = 1.6$  Hz, 1H), 7.23 (dd,  $J = 10.2, 1.6$  Hz, 1H), 4.65 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  151.4, 148.2, 137.3 (d,  $J = 15$  Hz), 129.3 (d,  $J = 1$  Hz), 119.6 (d,  $J = 1$  Hz), 117.5, (d,  $J = 22$  Hz), 99.5 (d,  $J = 2$  Hz).  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -129.99. **LRMS** (ESI -ve) 172.00 (30%), 170.00 (100%  $[\text{M}]^+$ ). **HRMS** (EI +ve) 170.0049 (Calculated 170.0047 for  $\text{C}_7\text{H}_4\text{N}_2\text{FCl}$ ). **IR** (solid)  $\text{cm}^{-1}$  3476, 3339, 3208, 3084, 2232, 1630, 1557, 1510, 1433, 1323, 1250, 1227, 1177, 1126, 1086, 962, 883, 866, 766, 721, 683, 631, 604, 548, 538, 511, 469, 444, 428, 411, 403.

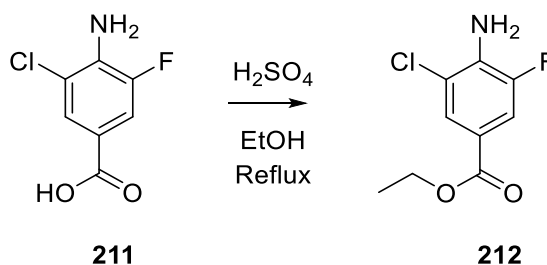
## 6.6.11 4-Amino-3-chloro-5-fluorobenzoic acid (211)



4-Amino-3-bromo-5-fluorobenzonitrile (5.00 g, 29 mM, 1 eq.) was suspended in sodium hydroxide (1 M, 80 mL) and refluxed until thin layer chromatography showed no starting material remained. The solution was cooled to room temperature and acidified with hydrochloric acid solution (1 M), causing precipitation of the product. The precipitate was filtered off then dissolved in ethyl acetate, dried over anhydrous magnesium sulfate, filtered and the solvent removed under reduced pressure to yield **211** (4.01 g, 72%) as a white solid.

<sup>1</sup>H NMR (300 MHz, DMSO-d<sub>6</sub>) δ 12.79 (s, 1H), 7.65-7.57 (m, 1H), 7.48 (dd, *J* = 11.5, 1.8 Hz, 1H), 6.28 (s, 2H). <sup>13</sup>C NMR (75 MHz, DMSO-d<sub>6</sub>) δ 166.3, 151.3, 148.1, 138.9 (d, *J* = 2 Hz), 127.0, 117.5, (d, *J* = 7.5 Hz), 114.9 (d, *J* = 22 Hz). <sup>19</sup>F NMR (376 MHz, DMSO) δ -130.19. **LRMS** (EI +ve) 190.99 (20%), 189.00 (80%), 175.99 (22%), 173.99 (32%), 172.00 (100%), 144.00 (45%), 108.03 (28%). **HRMS** (ESI -ve) 188.9993 (Calculated 188.9993 for C<sub>7</sub>H<sub>5</sub>NO<sub>2</sub>FCl). **IR** (solid) cm<sup>-1</sup> 3509, 3399, 1686, 1605, 1570, 1518, 1441, 1414, 1354, 1327, 1269, 1213, 1096, 951, 916, 893, 862, 764, 725, 696, 611, 550, 473, 467, 457, 438, 420, 413, 409, 401.

## 6.6.12 Ethyl 4-amino-3-chloro-5-fluorobenzoate (212)

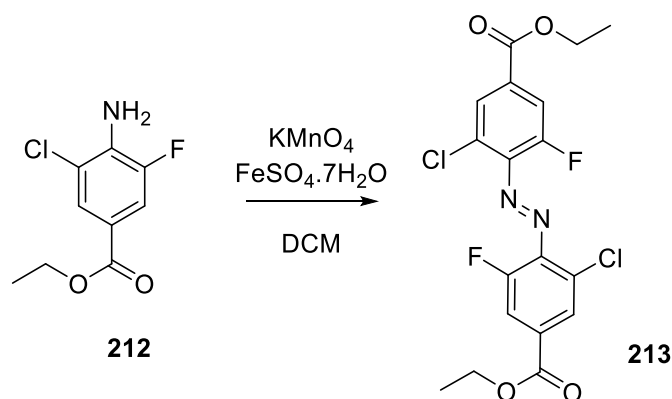


4-Amino-3-chloro-5-fluorobenzoic acid (2.5 g, 13 mM, 1 eq.) was dissolved in ethanol (100 mL) and sulfuric acid (5 mL) and the resulting solution was refluxed for 14 hours. The solution was then neutralized with saturated sodium hydrogen carbonate solution and extracted twice with dichloromethane. Combined organic fractions were dried

over magnesium sulfate, filtered, and then concentrated under reduced pressure to yield **212** 2.37 g (83%) as an off white solid.

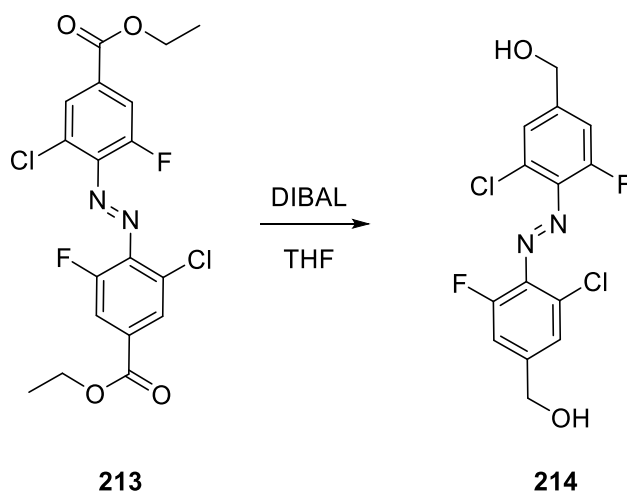
$^1\text{H NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  7.66-7.59 (m, 1H), 7.50 (dd,  $J = 11.5, 1.8$  Hz, 1H), 6.36 (s, 2H), 4.24 (q,  $J = 7.1$  Hz, 2H), 1.29 (dd,  $J = 8.3, 5.9$  Hz, 3H).  $^{13}\text{C NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  164.7 (d,  $J = 1$  Hz), 148.1, 139.0 (d,  $J = 2$  Hz), 126.8, 117.6 (d,  $J = 2$  Hz), 116.5 (d,  $J = 1$  Hz), 114.9 (d,  $J = 2$  Hz), 114.6, 61.0, 14.6.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -122.84. **LRMS** (ESI -ve) 217.03 (50%  $[\text{M}]^+$ ), 188.99 (38%), 173.99 (36%), 171.99 (100%), 144.00 (22%), 130.99 (30%), 108.02 (20%). **HRMS** (EI +ve) 217.0304 (Calculated 217.0306 for  $\text{C}_9\text{H}_9\text{NO}_2\text{FCl}$ ). **IR** (solid  $\text{cm}^{-1}$ ) 3493, 3374, 2980, 1697, 1616, 1595, 1570, 1514, 1477, 1431, 1396, 1369, 1346, 1323, 1261, 1206, 1117, 1096, 1067, 1028, 959, 932, 883, 853, 762, 712, 617, 594, 561, 534, 494, 467, 459, 451, 446, 420, 411.

#### 6.6.13 Diethyl 4,4'-(diazene-1,2-diyl)(E)-bis(3-chloro-5-fluorobenzoate) (**213**)



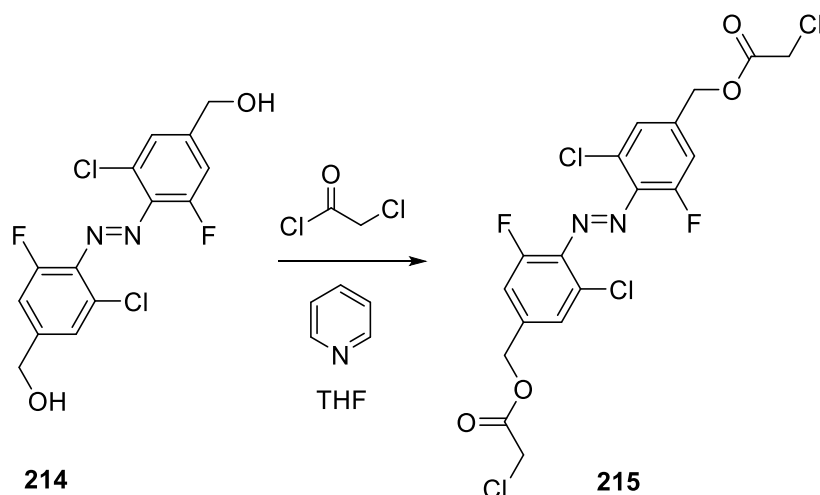
Potassium permanganate (4 g) and iron (II) sulfate heptahydrate (4 g) were ground together to a fine powder, then suspended in dichloromethane (250 mL) and ethyl 4-amino-3-chloro-5-fluorobenzoate (2.00 g, 9.2 mM, 1 eq.) was added. The suspension was refluxed overnight then cooled, filtered through celite and the solvent removed under reduced pressure. The resulting solid was recrystallised from ethyl acetate to yield **213** (99 mg, 5%) as a fine red crystals.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.05 (t,  $J = 1.5$  Hz, 2H), 7.82 (dd,  $J = 10.5, 1.5$  Hz, 2H), 4.43 (q,  $J = 7.1$  Hz, 4H), 1.43 (t,  $J = 7.1$  Hz, 6H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{CDCl}_3$ )  $\delta$  163.8, 151.9, (d,  $J = 127$  Hz), 142.1, (d,  $J = 7.5$  Hz), 133.0, (d,  $J = 7.5$  Hz), 131.8, 127.4 (d,  $J = 15$  Hz), 117.0 (d,  $J = 15$  Hz), 62.1, 14.2.  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -130.58, -131.96. **LRMS** (EI +ve) 432.02 (34%), 430.03 (50%  $[\text{M}]^+$ ), 231.02 (36%), 229.02 (100%), 201.01 (23%), 172.98 (48%), 116.99 (26%). **HRMS** (ES +ve) 431.0383 (Calculated 431.0377 for  $\text{C}_{18}\text{H}_{18}\text{N}_2\text{O}_4\text{F}_2\text{Cl}_2$ ).

6.6.14 (*E*)-(Diazene-1,2-diylbis(3-chloro-5-fluoro-4,1-phenylene))dimethanol (214)

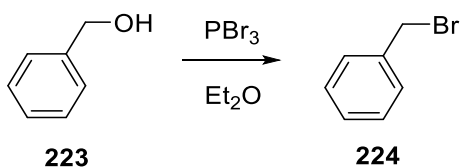
Diethyl 4,4'-(diazene-1,2-diyl)-(*E*)-bis(3-chloro-5-fluorobenzoate) (39 mg 0.07 mM) was dissolved in dry tetrahydrofuran (30 mL) and the flask flushed with argon. Diisobutylaluminium hydride solution (0.7 mL, 0.7 mM 1 M in toluene) was slowly added and the reaction was stirred overnight. Potassium sodium tartrate (5 mL) was added and the suspension was stirred vigorously for 1 hour. The organic layer was separated, dried over sodium sulfate, filtered and evaporated under reduced pressure. The solid residue was purified over silica gel eluting with 20-50% ethyl acetate/hexane to yield **214** (21 mg, 68%) as a pale solid.

<sup>1</sup>H NMR (600 MHz, MeOD)  $\delta$  7.71 (d,  $J = 5.6$  Hz, 3H), 7.50 (s, 1H), *trans* 4.75 (s, 1H), *cis* 4.68 (s, 1H). <sup>13</sup>C NMR (151 MHz, MeOD)  $\delta$  147.08, 126.57 (t,  $J = 1$  Hz), 121.15, 113.97, 113.95, 113.83, 113.82 (t,  $J = 1$  Hz), 62.07. <sup>19</sup>F NMR (376 MHz, MeOD)  $\delta$  -117.22, -124.29. LRMS (EI +ve) 348.00 (38%), 346.01 (58% [M]<sup>+</sup>), 189.00 (34%), 187.01 (100%), 142.99 (26%), 140.99 (77%), 95.03 (20%). HRMS (EI +ve) 346.0077 (calculated 346.0087 C<sub>14</sub>H<sub>10</sub>Cl<sub>2</sub>F<sub>2</sub>N<sub>2</sub>O<sub>2</sub>).

6.6.15 (*E*)-(Diazene-1,2-diylbis(4,1-phenylene))bis(methylene) bis(2-chloroacetate) (**215**)

(*E*)-(Diazene-1,2-diylbis(3-chloro-5-fluoro-4,1-phenylene))dimethanol (38 mg 0.11 mM, 1 eq.) was dissolved in tetrahydrofuran and the flask flushed with argon. Chloroacetyl chloride (100  $\mu$ L, excess), was added dropwise and the reaction was stirred for 20 minutes at room temperature. Pyridine (100  $\mu$ L, excess) was then added dropwise over 20 minutes and the solution stirred overnight. The solvent was then removed under reduced pressure and the residue was partitioned between dichloromethane and sodium hydrogen carbonate. The organic layer was dried over sodium sulfate and filtered, then the solvent was removed under reduced pressure and the resulting solid was recrystallised from methanol to yielded **215** (42 mg, 77%) as deep red crystals.

**$^1\text{H}$  NMR** (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.32 (s, 2H), 7.11 (dd,  $J = 10.8, 1.6$  Hz, 2H), 5.16 (s, 4H), 4.10 (s, 4H).  **$^{13}\text{C}$  NMR** (151 MHz,  $\text{CDCl}_3$ )  $\delta$  167.1, 138.5, 132.7, 125.4, 115.8, 115.1, 110.5, 65.8, 40.4.  **$^{19}\text{F}$  NMR** (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -116.50, -122.27. **LRMS** (EI +ve) 501.97 (100%), 499.94 (28%), 497.95 (25%), 264.98 (54%), 262.98 (69%), 142.99 (20%), 140.99 (64%). **HRMS** (EI +ve) 497.9516 (calculated 497.9519 for  $\text{C}_{18}\text{H}_{12}\text{Cl}_4\text{F}_2\text{N}_2\text{O}_4$ ).

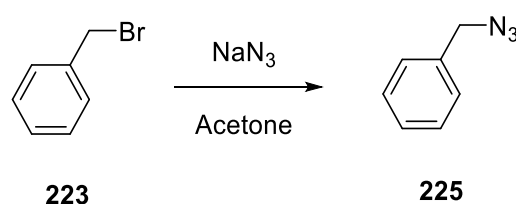
6.6.16 Benzyl bromide (**224**)

Benzyl alcohol (0.50 g, 4.62 mM, 1 eq.) was dissolved in diethyl ether (20 mL) to which was added phosphorus tribromide (0.625 g, 2.31 mM, 0.5 eq.) and the resulting solution was stirred for 3 hours at room temperature. The reaction mixture was washed with

brine, dried over magnesium sulfate, filtered and the solvent removed under reduced pressure. The residue was purified over silica gel eluting with dichloromethane to yield **224** (0.70 g, 89%) as a pale yellow oil.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.46-7.23 (m, 5H), 4.51 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  137.8, 129.1, 128.8, 128.4, 33.6. **LRMS** (EI +ve) 171.97 (16%), 169.97 (15%), 92.06 (21%), 91.05 (100%), 89.04 (13%), 65.04 (20%), 63.02 (12%). **HRMS** (EI +ve) 169.9729 (calculated Cal 169.9731  $\text{C}_7\text{H}_7\text{Br}$ ).

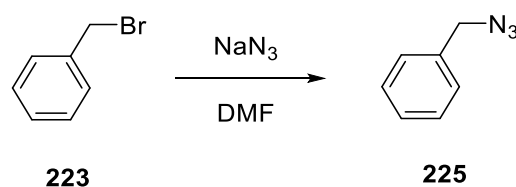
#### 6.6.17 Benzyl azide (**225**)



Sodium azide (0.862 g, 13.3 mM, 1 eq.) was dissolved in acetone (20 mL) to which benzyl bromide (0.96 g, 5.65 mM) was added dropwise over 20 minutes at 0 °C. The reaction was then warmed to room temperature and stirred overnight to which was added water (20 mL). The solvent was removed under reduced pressure extracted into dichloromethane to yield **225** 0.64 g (85%)

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.48-7.23 (m, 5H), 4.35 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  135.4, 128.9, 128.3, 128.3, 54.8. **LRMS** (EI +ve) 133.06 (28%), 105.04 (42%), 104.5 (92%), 91.05 (100%), 85.95 (29%), 83.95 (43%), 78.04 (33%), 77.04 (81%), 51.02 (37%). **HRMS** (EI +ve) 133.0635 (calculated Cal 133.0640  $\text{C}_7\text{H}_7\text{N}_3$ ).

#### 6.6.18 Benzyl azide (**225**)



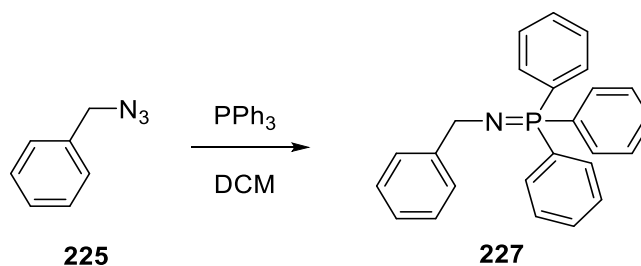
To a solution of sodium azide (1.44 g, 22 mM) in dimethylformamide (60 mL) was added benzyl bromide (0.96 g, 5.65 mM) and the resulting suspension was stirred under argon at 80 °C overnight. The reaction was diluted with ether and washed with water. The organic fraction was dried over magnesium sulfate, filtered and the solvent



removed under reduced pressure to yield to give the **225** (0.73 g, 99%) as a colourless oil.

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.48-7.23 (m, 5H), 4.35 (s, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  135.4, 128.9, 128.3, 128.3, 54.8. **LRMS** (EI +ve) 133.06 (28%), 105.04 (42%), 104.5 (92%), 91.05 (100%), 85.95 (29%), 83.95 (43%), 78.04 (33%), 77.04 (81%), 51.02 (37%). **HRMS** (EI +ve) 133.0635 (calculated 133.0640 for  $\text{C}_7\text{H}_7\text{N}_3$ ).

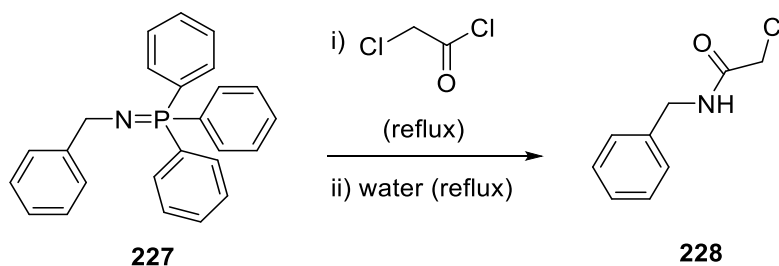
#### 6.6.19 *N*-Benzyl-1,1,1-triphenyl-*l*-5-phosphanimine (**227**)



To a solution of triphenyl phosphine (200 mg, 1.5 mM) in dichloromethane (10 mL) was added a solution of benzyl azide (394 mg, 1.5 mL) in further dichloromethane (10 mL) and the resulting mixture was stirred overnight. The solvent was then removed under reduced pressure and the residue purified over silica gel to yield **227** (430 mg, 78%).

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.77-7.05 (m, 20H), 4.46-4.24 (m, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  133.2, 132.2, 132.1, 128.6, 128.5.  $^{31}\text{P NMR}$  (202 MHz,  $\text{CDCl}_3$ )  $\delta$  29.20.

#### 6.6.20 *N*-Benzyl chloroacetamide (**228**)

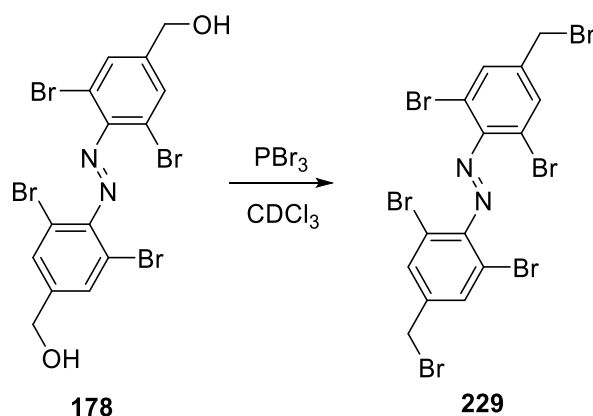


To a solution of *N*-benzyl-1,1,1-triphenyl-*l*-5-phosphanimine (100 mg, 0.27 mM) in toluene (10 mL) a solution of chloroacetyl chloride (30  $\mu\text{L}$  0.22 mM) in toluene (5 mL) was added dropwise and the resulting solution was refluxed for 3 hours. Water (2 mL) was added and the reaction was refluxed for a further 3 hours, then cooled to room temperature added to a solution of sodium hydroxide (10 mL, 1N). The mixture was extracted with ethyl acetate and the organic fraction was dried over sodium sulfate,

filtered and the solvent was removed under reduced pressure. The residue was purified over silica gel eluting with diethyl ether to yield **228** (40 mg, 80%),

$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.38-7.17 (m, 5H), 6.81 (s, 1H), 4.43 (d,  $J = 5.8$  Hz, 2H), 4.05 (d,  $J = 5.1$  Hz, 2H).  $^{13}\text{C NMR}$  (75 MHz,  $\text{CDCl}_3$ )  $\delta$  165.8, 137.3, 128.9, 127.9, 127.8, 43.9, 42.7. **LRMS** (EI +ve) 183.06 (50%), 185.04 (15%) 148.07 (100%), 149.08(10%), 107.05 (42%), 105.07 (23%), 91.05 (50%). **HRMS** (EI +ve) 183.0455 (Calculated 183.0451 for  $\text{C}_7\text{H}_4\text{N}_2\text{FBr}$ ).

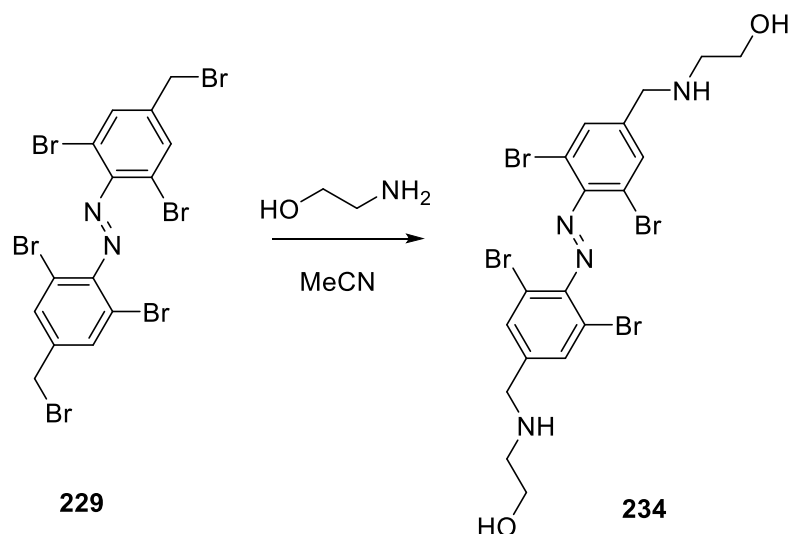
#### 6.6.21 (E)-1,2-bis(2,6-Dibromo-4-(bromomethyl)phenyl)diazene (**229**)



To a suspension of (*E*)-(diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene))dimethanol (100 mg, 0.18 mM, 1 eq.) in chloroform (30 mL) was added phosphorus tribromide (24.3 mg, 0.09 mM, 0.5 eq.) the resulting mixture was stirred overnight at 50 °C. The solvent is removed under reduced pressure and immediately purified by flash silica chromatography (5% methanol: dichloromethane) to yield **229** 82 mg (67%).

$^1\text{HNMR}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.25-7.17 (m, 4H), 5.28-5.20 (m, 4H).  $^{13}\text{C NMR}$  (151 MHz,  $\text{CDCl}_3$ )  $\delta$  133.8, 131.4, 123.2, 116.5, 63.5. **LRMS** (EI +ve) 685.58 (33%), 683.58 (46%), 681.58 (38%), 606.66 (25%), 604.66 (50%), 602.66 (54%), 600.66 (28%), 358.79 (35%), 356.79 (95%), 354.79 (100%), 330.78 (23%), 328.78 (66%), 326.97 (62%), 324.79 (27%), 249.86 (32%), 247.87 (61%), 245.87 (34%), 168.95 (77%), 166.95 (64%), 88.03 (32%). **HRMS** (EI +ve) 677.5794 (Calculated 677.5788 for  $\text{C}_{14}\text{H}_8\text{N}_2\text{Br}_6$ ). **IR** (solid)  $\text{cm}^{-1}$  2926, 2856, 2104, 1506, 1377, 1344, 989, 964, 868, 827, 744, 522, 436, 420.

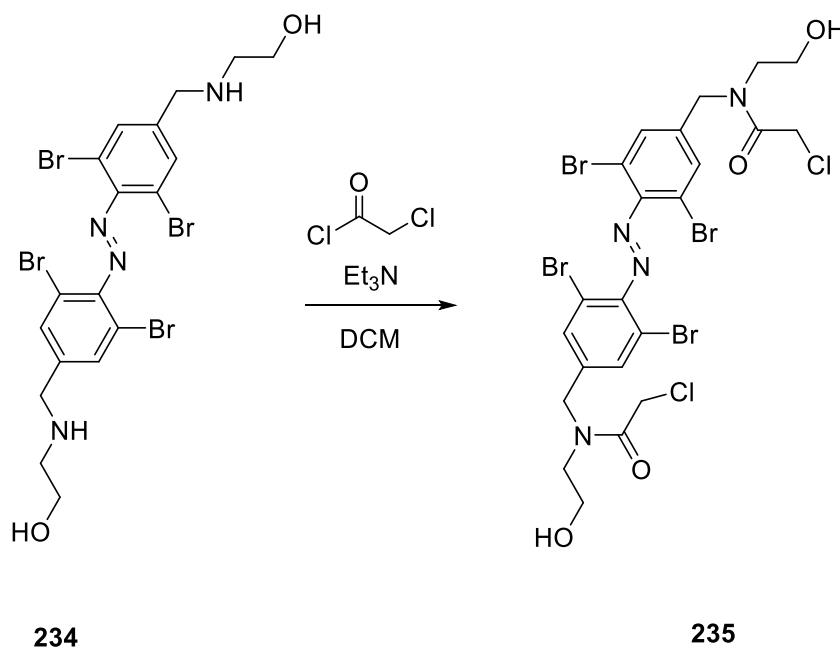
6.6.22 (*E*)-2,2'-(((diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene)) bis(methylene)) bis(azanediyl)) bis(ethan-1-ol) (**234**)



To a solution of (*E*)-2,2'-(((diazene-1,2-diyl-bis-(3,5-dibromo-4,1-phenylene)) bis(methylene))-bis(azanediyl))bis(ethan-1-ol) (20 mg, 0.0298 mM, 1 eq.) in acetonitrile (50 mL), was added ethanolamine (18 mg, 85 mM, excess.) and the resulting solution was stirred overnight. The solvent was then removed under reduced pressure and the oily residue was partitioned between water and hexane. The organic layer was washed with water, dried over magnesium sulfate, filtered and the solvent was removed under reduced pressure to yield **234** (14 mg, 74%) as a red solid.

<sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD) δ 7.82 (s, 4H), 3.87 (s, 4H), 3.71 (t, *J* = 5.5 Hz, 4H), 2.76 (t, *J* = 5.5 Hz, 4H). <sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD) δ 147.34, 143.5, 133.0, 115.5, 60.3, 51.3, 50.2. LRMS (EI +ve) 646.84 (54%), 644.64 (89%), 642.84 (63%), 342.17 (28%), 280.19 (66%), 115.09 (82%), 105.05 (100%). HRMS (EI +ve) 639.8322 (calculated 639.8320 for C<sub>18</sub>H<sub>20</sub>N<sub>4</sub>O<sub>2</sub>Br<sub>4</sub>).

6.6.23 (*E*)-*N,N'*-((Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene)) bis(methylene)) bis(2-chloro-*N*-(2-hydroxyethyl)acetamide) (**235**)



To a solution of (*E*)-2,2'-(((diazene-1,2-diylbis (3,5-dibromo-4,1-phenylene)) bis(methylene))-bis(azanediyl))bis(ethan-1-ol) (24 mg, 0.037 mM, 1 eq.) in dichloromethane (10 mL) was added chloroacetic acid (10.1 mg, 0.1 mM, 2.7 eq.) and the resulting solution was stirred for 15 minutes. Triethylamine (15 mg, 0.14 mM, 4 eq.) was added and the reaction was stirred for a further 3 hours. The solution was then washed with sodium bicarbonate and the organic fraction was dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The residue was purified over silica gel eluting with 10 % methanol in ethyl acetate to yield **235** (29 mg, 97%) as a dark red oil.

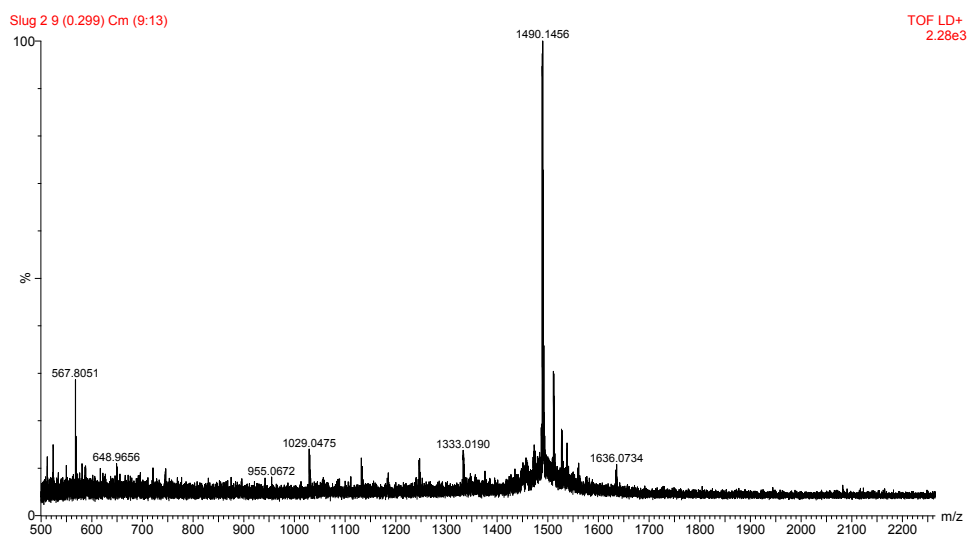
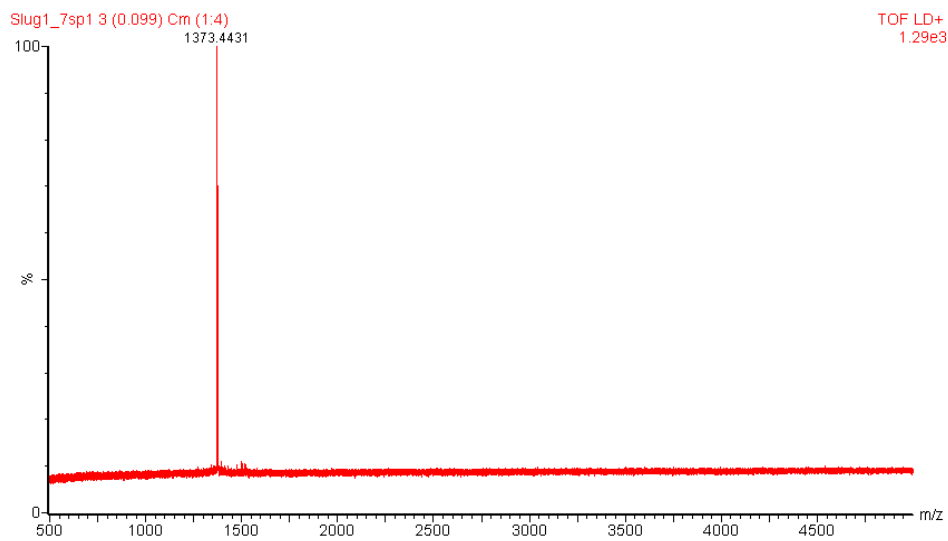
<sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>OD) δ 7.73 (s, 4H), 4.87 (s, 4H), 4.73 (s, 4H), 4.50 (dd, *J* = 5.2, 5.1 Hz, 4H), 3.76 (t, *J* = 5.2 Hz, 4H), 3.60 (t, *J* = 5.1 Hz, 2H). <sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD) δ 169.1, 147.6, 141.1, 132.2, 131.6, 115.6, 58.9, 50.0, 41.1, **LRMS** (ESI -ve) 834.74 (54%), 832.73 (80%), 830.73 (100%), 828.7492 (55%), 826.74 (12%), 454.24 (12%), 453.24 (53%). **HRMS** (ESI -ve) 826.7479 (calculated 826.7440 for C<sub>22</sub>H<sub>21</sub>Br<sub>4</sub>Cl<sub>2</sub>N<sub>4</sub>O<sub>4</sub>).

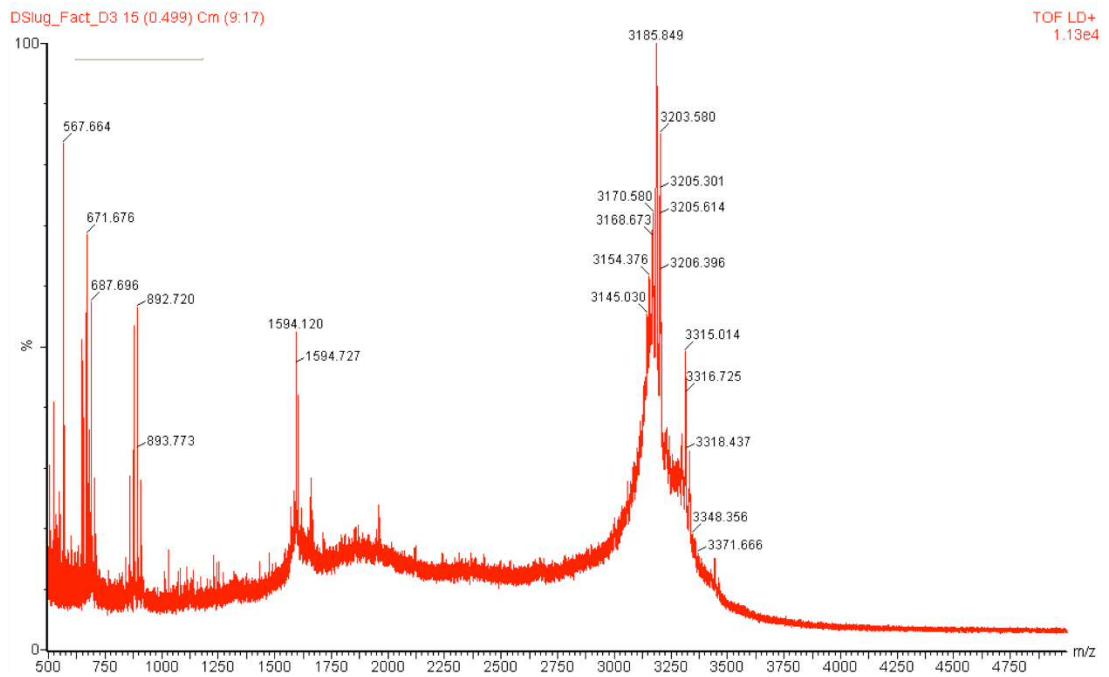
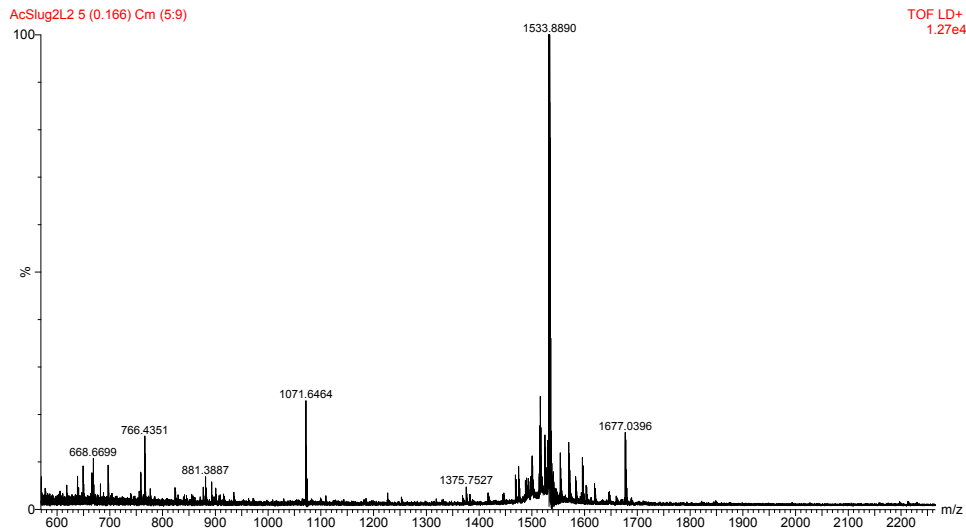
# Chapter 7

# Appendix

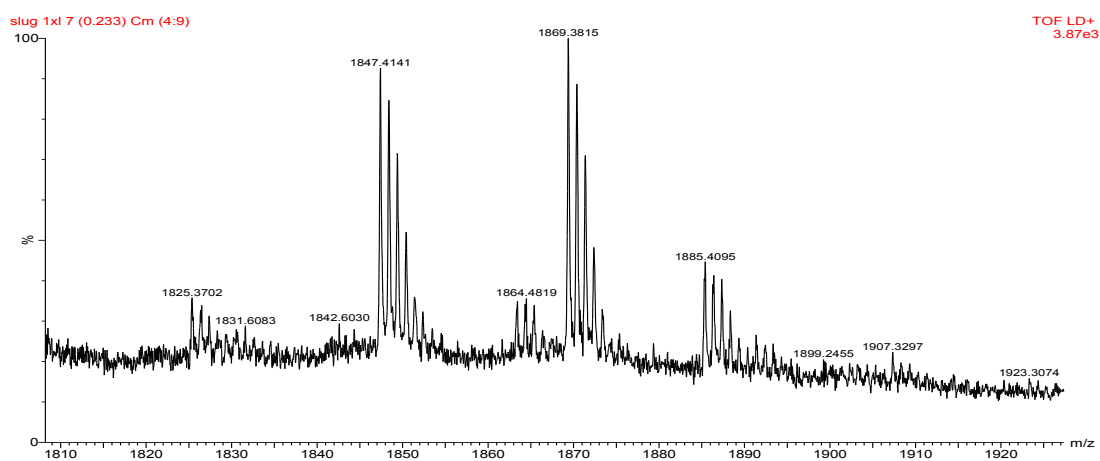
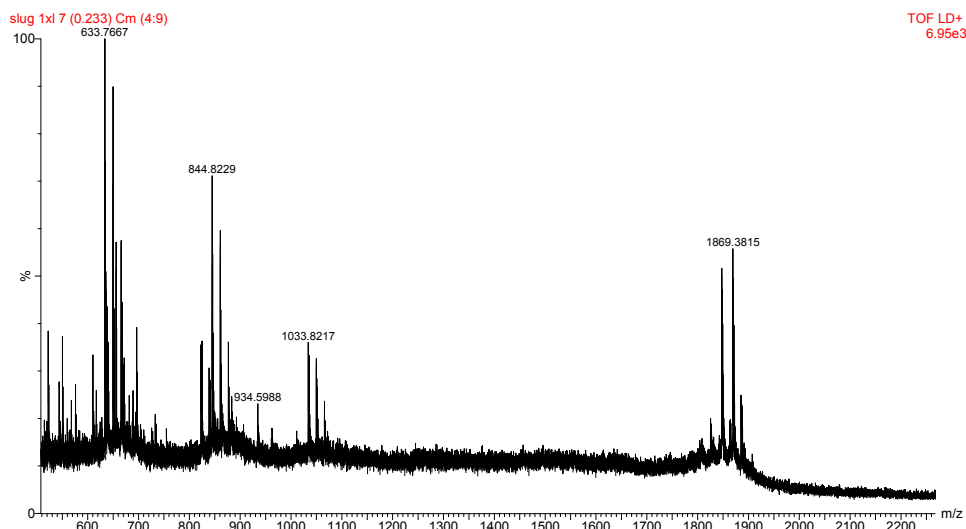
## 7 Appendix

### 7.1 Slug peptide masses

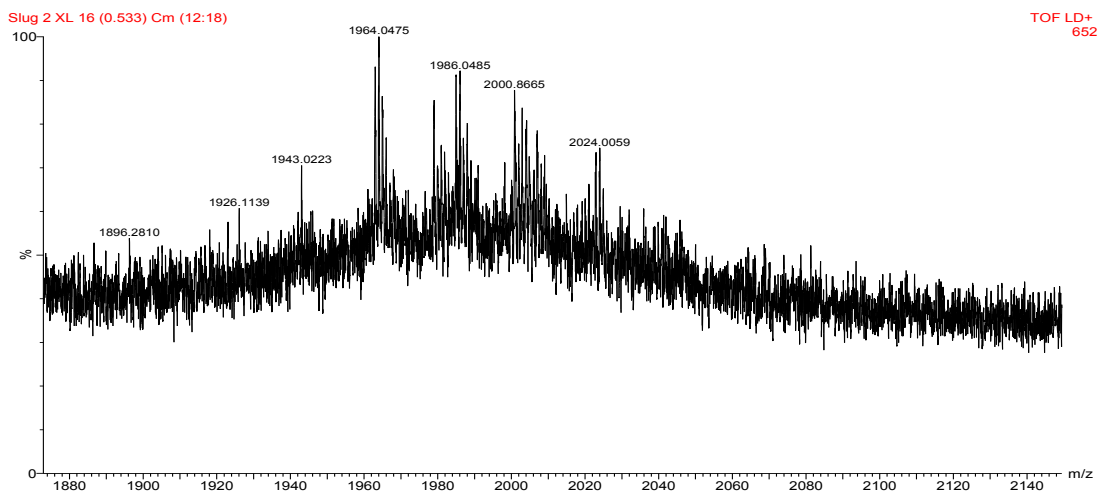
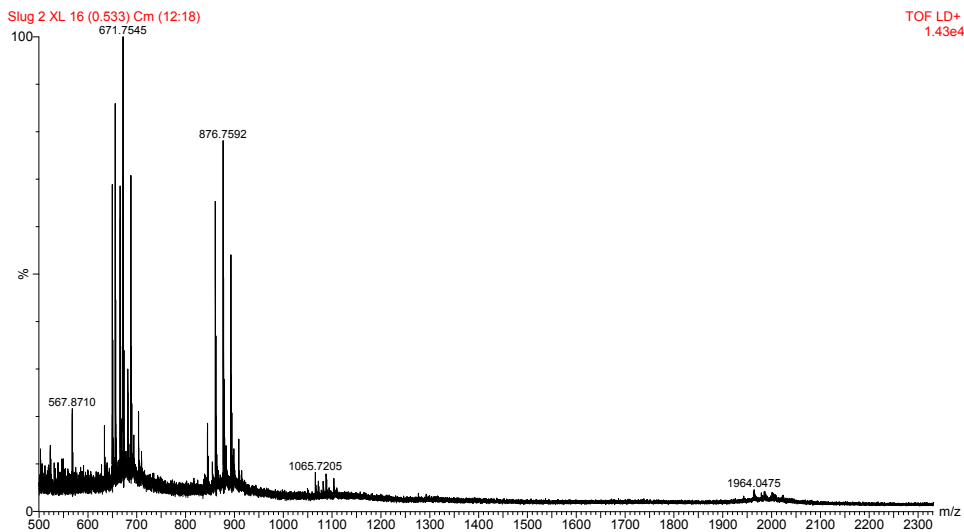


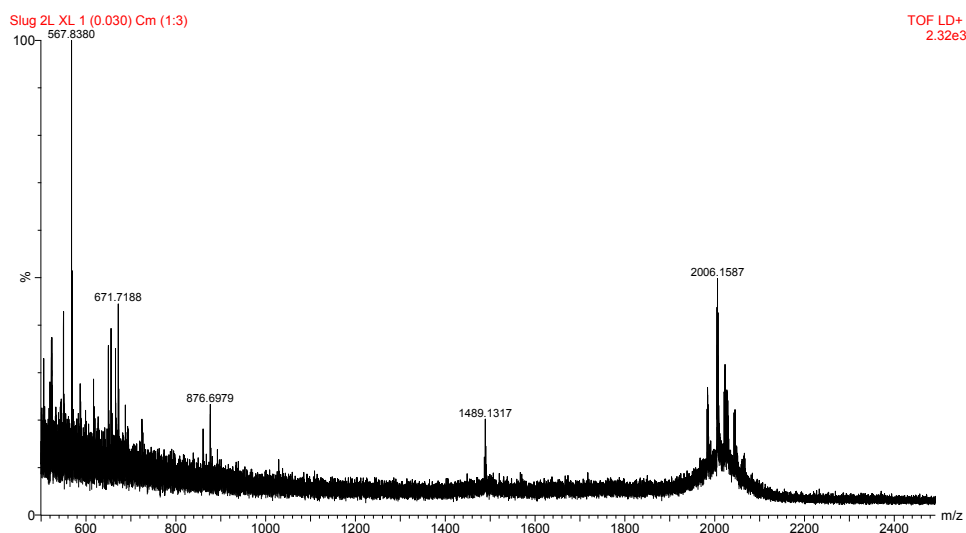
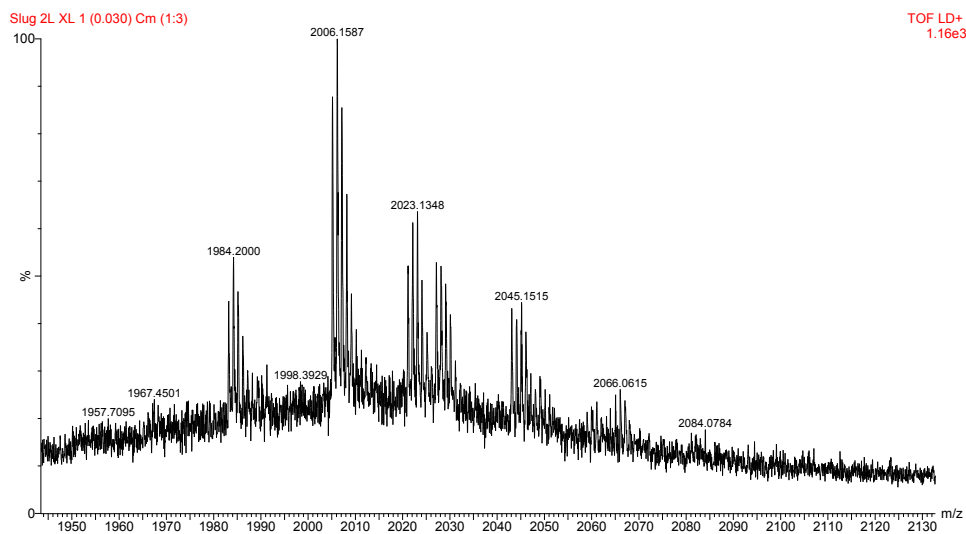


Slug Finger 1-71









## 7.2 Crystal structure data

### 7.2.1 (*E*)-(Diazene-1,2-diylbis(3,5-dichloro-4,1-phenylene))bis(methylene) bis(2-chloroacetate) (171)

Table 1. Crystal data and structure refinement for exp\_750.

Identification code	shelx
Empirical formula	C <sub>18</sub> H <sub>12</sub> Cl <sub>6</sub> N <sub>2</sub> O <sub>4</sub>
Formula weight	533.00
Temperature	150(2) K
Wavelength	1.54184 Å
Crystal system	Monoclinic
Space group	C 2/c

---

Unit cell dimensions	a = 26.8197(6) Å b = 5.44040(10) Å c = 14.4892(3) Å	$\alpha = 90^\circ$ . $\beta = 102.554(2)^\circ$ . $\gamma = 90^\circ$ .
Volume	2063.57(8) Å <sup>3</sup>	
Z	4	
Density (calculated)	1.716 Mg/m <sup>3</sup>	
Absorption coefficient	7.875 mm <sup>-1</sup>	
F(000)	1072	
Crystal size	0.256 x 0.160 x 0.034 mm <sup>3</sup>	
Theta range for data collection	3.377 to 74.226°.	
Index ranges	-31 ≤ h ≤ 32, -6 ≤ k ≤ 6, -17 ≤ l ≤ 17	
Reflections collected	8426	
Independent reflections	2074 [R(int) = 0.0242]	
Completeness to theta = 67.684°	99.9 %	
Refinement method	Full-matrix least-squares on F <sup>2</sup>	
Data / restraints / parameters	2074 / 0 / 136	
Goodness-of-fit on F <sup>2</sup>	1.058	
Final R indices [I > 2σ(I)]	R1 = 0.0246, wR2 = 0.0636	
R indices (all data)	R1 = 0.0266, wR2 = 0.0655	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.300 and -0.262 e.Å <sup>-3</sup>	

Table 2. Atomic coordinates ( $\times 10^4$ ) and equivalent isotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for exp\_750.  $U(\text{eq})$  is defined as one third of the trace of the orthogonalized  $U^{\text{ij}}$  tensor.

	x	y	z	U(eq)
C(1)	7642(1)	3350(3)	6476(1)	29(1)
C(2)	7072(1)	3173(3)	6392(1)	22(1)
C(3)	6347(1)	5615(3)	6415(1)	23(1)
C(4)	6058(1)	5574(3)	5399(1)	19(1)
C(5)	6167(1)	7338(3)	4774(1)	19(1)
C(6)	5894(1)	7329(3)	3849(1)	18(1)
C(7)	5493(1)	5676(2)	3531(1)	17(1)
C(8)	5395(1)	3915(3)	4168(1)	18(1)
C(9)	5678(1)	3828(3)	5090(1)	19(1)
N(1)	5238(1)	5788(2)	2556(1)	18(1)
O(1)	6821(1)	1343(2)	6296(1)	34(1)
O(2)	6893(1)	5435(2)	6471(1)	21(1)
Cl(1)	7900(1)	622(1)	6115(1)	43(1)
Cl(2)	6043(1)	9487(1)	3071(1)	23(1)
Cl(3)	4940(1)	1649(1)	3808(1)	21(1)

Table 3. Bond lengths [ $\text{\AA}$ ] and angles [ $^\circ$ ] for exp\_750.

C(1)-C(2)	1.508(2)
C(1)-Cl(1)	1.7655(16)
C(1)-H(1A)	0.9900
C(1)-H(1B)	0.9900
C(2)-O(1)	1.1931(19)
C(2)-O(2)	1.3344(18)
C(3)-O(2)	1.4544(17)
C(3)-C(4)	1.5070(18)
C(3)-H(3A)	0.9900
C(3)-H(3B)	0.9900
C(4)-C(5)	1.393(2)
C(4)-C(9)	1.3941(19)
C(5)-C(6)	1.3807(18)
C(5)-H(5)	0.9500
C(6)-C(7)	1.4007(19)

---

C(6)-Cl(2)	1.7324(14)
C(7)-C(8)	1.3945(19)
C(7)-N(1)	1.4323(17)
C(8)-C(9)	1.3867(19)
C(8)-Cl(3)	1.7336(14)
C(9)-H(9)	0.9500
N(1)-N(1)#1	1.249(2)
C(2)-C(1)-Cl(1)	112.28(11)
C(2)-C(1)-H(1A)	109.1
Cl(1)-C(1)-H(1A)	109.1
C(2)-C(1)-H(1B)	109.1
Cl(1)-C(1)-H(1B)	109.1
H(1A)-C(1)-H(1B)	107.9
O(1)-C(2)-O(2)	125.09(14)
O(1)-C(2)-C(1)	126.76(14)
O(2)-C(2)-C(1)	108.13(12)
O(2)-C(3)-C(4)	110.49(11)
O(2)-C(3)-H(3A)	109.6
C(4)-C(3)-H(3A)	109.6
O(2)-C(3)-H(3B)	109.6
C(4)-C(3)-H(3B)	109.6
H(3A)-C(3)-H(3B)	108.1
C(5)-C(4)-C(9)	120.14(12)
C(5)-C(4)-C(3)	119.30(12)
C(9)-C(4)-C(3)	120.54(13)
C(6)-C(5)-C(4)	119.16(12)
C(6)-C(5)-H(5)	120.4
C(4)-C(5)-H(5)	120.4
C(5)-C(6)-C(7)	121.88(13)
C(5)-C(6)-Cl(2)	118.76(11)
C(7)-C(6)-Cl(2)	119.33(10)
C(8)-C(7)-C(6)	117.79(12)
C(8)-C(7)-N(1)	124.27(12)
C(6)-C(7)-N(1)	117.77(12)
C(9)-C(8)-C(7)	121.18(13)
C(9)-C(8)-Cl(3)	117.93(11)
C(7)-C(8)-Cl(3)	120.78(10)

---

C(8)-C(9)-C(4)	119.72(13)
C(8)-C(9)-H(9)	120.1
C(4)-C(9)-H(9)	120.1
N(1)#1-N(1)-C(7)	112.57(14)
C(2)-O(2)-C(3)	115.76(11)

---

Symmetry transformations used to generate equivalent atoms:

#1 -x+1,y,-z+1/2

Table 4. Anisotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for exp\_750. The anisotropic displacement factor exponent takes the form:  $-2\pi^2 [ h^2 a^{*2} U^{11} + \dots + 2 h k a^* b^* U^{12} ]$

	$U^{11}$	$U^{22}$	$U^{33}$	$U^{23}$	$U^{13}$	$U^{12}$
C(1)	24(1)	28(1)	33(1)	-3(1)	-1(1)	5(1)
C(2)	26(1)	20(1)	17(1)	-1(1)	-1(1)	2(1)
C(3)	20(1)	30(1)	19(1)	-3(1)	1(1)	3(1)
C(4)	16(1)	21(1)	18(1)	-3(1)	2(1)	4(1)
C(5)	15(1)	19(1)	21(1)	-4(1)	2(1)	0(1)
C(6)	16(1)	16(1)	21(1)	1(1)	4(1)	2(1)
C(7)	15(1)	17(1)	18(1)	-2(1)	2(1)	3(1)
C(8)	15(1)	17(1)	21(1)	-2(1)	3(1)	0(1)
C(9)	19(1)	19(1)	20(1)	2(1)	4(1)	2(1)
N(1)	18(1)	16(1)	18(1)	0(1)	1(1)	-1(1)
O(1)	34(1)	22(1)	43(1)	-3(1)	2(1)	-4(1)
O(2)	19(1)	20(1)	23(1)	-1(1)	-2(1)	1(1)
Cl(1)	40(1)	45(1)	42(1)	-9(1)	4(1)	20(1)
Cl(2)	23(1)	23(1)	24(1)	4(1)	4(1)	-5(1)
Cl(3)	20(1)	19(1)	25(1)	0(1)	2(1)	-4(1)

---

Table 5. Hydrogen coordinates ( $\times 10^4$ ) and isotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for exp\_750.

	x	y	z	U(eq)
H(1A)	7809	3704	7142	35
H(1B)	7716	4734	6083	35
H(3A)	6271	7161	6719	28
H(3B)	6234	4224	6761	28
H(5)	6426	8531	4982	22
H(9)	5614	2585	5509	23

### 7.2.2 (*E*)-(Diazene-1,2-diylbis(3,5-dibromo-4,1-phenylene))bis(methylene) bis(2-chloroacetate (179))

Table 1. Crystal data and structure refinement for rka1402b.

Identification code	shelx	
Empirical formula	C <sub>9</sub> H <sub>6</sub> Br <sub>2</sub> Cl N O <sub>2</sub>	
Formula weight	355.42	
Temperature	296(2) K	
Wavelength	1.54184 \AA	
Crystal system	Monoclinic	
Space group	C 2/c	
Unit cell dimensions	a = 27.4395(11) \AA	$\alpha = 90^\circ$ .
	b = 5.6145(2) \AA	$\beta = 105.207(4)^\circ$ .
	c = 15.1895(5) \AA	$\gamma = 90^\circ$ .
Volume	2258.14(15) \AA <sup>3</sup>	
Z	8	
Density (calculated)	2.091 Mg/m <sup>3</sup>	
Absorption coefficient	11.136 mm <sup>-1</sup>	
F(000)	1360	
Crystal size	0.256 x 0.160 x 0.034 mm <sup>3</sup>	
Theta range for data collection	3.338 to 73.959 $^\circ$ .	
Index ranges	-33 $\leq$ h $\leq$ 25, -5 $\leq$ k $\leq$ 6, -17 $\leq$ l $\leq$ 18	
Reflections collected	3920	
Independent reflections	2210 [R(int) = 0.0210]	

Completeness to theta = 67.684°	99.5 %
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	2210 / 0 / 136
Goodness-of-fit on F <sup>2</sup>	1.012
Final R indices [I>2sigma(I)]	R1 = 0.0324, wR2 = 0.0844
R indices (all data)	R1 = 0.0401, wR2 = 0.0894
Extinction coefficient	n/a
Largest diff. peak and hole	0.753 and -0.594 e.Å <sup>-3</sup>

Table 2. Atomic coordinates ( x 10<sup>4</sup>) and equivalent isotropic displacement parameters (Å<sup>2</sup>x 10<sup>3</sup>) for rka1402b. U(eq) is defined as one third of the trace of the orthogonalized U<sup>ij</sup> tensor.

	x	y	z	U(eq)
C(1)	7597(2)	2967(11)	11485(4)	71(1)
C(2)	7038(2)	2929(8)	11379(3)	48(1)
C(3)	6347(1)	5492(8)	11350(2)	47(1)
C(4)	6058(1)	5494(7)	10362(2)	38(1)
C(5)	5669(1)	3884(7)	10036(2)	37(1)
C(6)	5383(1)	3995(6)	9133(2)	35(1)
C(7)	5489(1)	5672(6)	8528(2)	32(1)
C(8)	5902(1)	7172(6)	8866(2)	34(1)
C(9)	6175(1)	7143(7)	9765(2)	38(1)
N(1)	5235(1)	5801(5)	7577(2)	35(1)
O(1)	6771(2)	1230(7)	11278(3)	77(1)
O(2)	6883(1)	5163(5)	11438(2)	46(1)
Cl(1)	7827(1)	311(4)	11172(2)	108(1)
Br(1)	4887(1)	1611(1)	8736(1)	44(1)
Br(2)	6089(1)	9413(1)	8084(1)	49(1)



Table 3. Bond lengths [ $\text{\AA}$ ] and angles [ $^\circ$ ] for rka1402b.

---

C(1)-C(2)	1.498(6)
C(1)-Cl(1)	1.733(6)
C(1)-H(1A)	0.9700
C(1)-H(1B)	0.9700
C(2)-O(1)	1.188(6)
C(2)-O(2)	1.336(5)
C(3)-O(2)	1.452(5)
C(3)-C(4)	1.503(4)
C(3)-H(3A)	0.9700
C(3)-H(3B)	0.9700
C(4)-C(5)	1.388(5)
C(4)-C(9)	1.391(5)
C(5)-C(6)	1.391(5)
C(5)-H(5)	0.9300
C(6)-C(7)	1.398(5)
C(6)-Br(1)	1.892(3)
C(7)-C(8)	1.398(5)
C(7)-N(1)	1.431(4)
C(8)-C(9)	1.374(5)
C(8)-Br(2)	1.893(3)
C(9)-H(9)	0.9300
N(1)-N(1)#1	1.251(6)
C(2)-C(1)-Cl(1)	113.2(4)
C(2)-C(1)-H(1A)	108.9
Cl(1)-C(1)-H(1A)	108.9
C(2)-C(1)-H(1B)	108.9
Cl(1)-C(1)-H(1B)	108.9
H(1A)-C(1)-H(1B)	107.8
O(1)-C(2)-O(2)	124.4(4)
O(1)-C(2)-C(1)	127.2(5)
O(2)-C(2)-C(1)	108.4(4)
O(2)-C(3)-C(4)	110.4(3)
O(2)-C(3)-H(3A)	109.6
C(4)-C(3)-H(3A)	109.6
O(2)-C(3)-H(3B)	109.6

---

C(4)-C(3)-H(3B)	109.6
H(3A)-C(3)-H(3B)	108.1
C(5)-C(4)-C(9)	119.3(3)
C(5)-C(4)-C(3)	120.7(3)
C(9)-C(4)-C(3)	120.0(3)
C(4)-C(5)-C(6)	120.2(3)
C(4)-C(5)-H(5)	119.9
C(6)-C(5)-H(5)	119.9
C(5)-C(6)-C(7)	121.3(3)
C(5)-C(6)-Br(1)	117.1(3)
C(7)-C(6)-Br(1)	121.5(2)
C(8)-C(7)-C(6)	116.9(3)
C(8)-C(7)-N(1)	118.2(3)
C(6)-C(7)-N(1)	124.6(3)
C(9)-C(8)-C(7)	122.4(3)
C(9)-C(8)-Br(2)	117.7(3)
C(7)-C(8)-Br(2)	119.8(2)
C(8)-C(9)-C(4)	119.8(3)
C(8)-C(9)-H(9)	120.1
C(4)-C(9)-H(9)	120.1
N(1)#1-N(1)-C(7)	113.2(3)
C(2)-O(2)-C(3)	116.5(3)

Symmetry transformations used to generate equivalent atoms:

#1 -x+1,y,-z+3/2

Table 4. Anisotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1402b. The anisotropic displacement factor exponent takes the form:  $-2\pi^2 [h^2 a^{*2} U^{11} + \dots + 2 h k a^* b^* U^{12}]$

	$U^{11}$	$U^{22}$	$U^{33}$	$U^{23}$	$U^{13}$	$U^{12}$
C(1)	50(3)	75(4)	78(3)	2(3)	3(2)	12(2)
C(2)	47(2)	49(2)	41(2)	0(2)	-1(2)	0(2)
C(3)	41(2)	65(3)	31(2)	-3(2)	0(1)	4(2)
C(4)	36(2)	46(2)	29(1)	-1(1)	2(1)	5(2)
C(5)	39(2)	40(2)	30(1)	3(1)	5(1)	1(2)
C(6)	31(2)	36(2)	34(2)	-2(1)	4(1)	1(1)
C(7)	28(1)	36(2)	29(1)	-3(1)	3(1)	3(1)

C(8)	29(2)	36(2)	34(1)	4(1)	5(1)	4(1)
C(9)	30(2)	42(2)	38(2)	-7(1)	1(1)	-3(1)
N(1)	31(1)	37(2)	32(1)	0(1)	0(1)	0(1)
O(1)	72(2)	53(2)	100(3)	-13(2)	14(2)	-11(2)
O(2)	40(1)	46(2)	44(1)	1(1)	-5(1)	-2(1)
Cl(1)	101(1)	105(1)	118(1)	-7(1)	30(1)	47(1)
Br(1)	41(1)	44(1)	44(1)	-1(1)	6(1)	-10(1)
Br(2)	43(1)	53(1)	47(1)	12(1)	6(1)	-7(1)

Table 5. Hydrogen coordinates ( $\times 10^4$ ) and isotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1402b.

	x	y	z	U(eq)
H(1A)	7767	3308	12117	85
H(1B)	7675	4244	11114	85
H(3A)	6294	6990	11629	57
H(3B)	6225	4218	11667	57
H(5)	5598	2726	10423	44
H(9)	6437	8222	9973	46

### 7.2.3 (E)-(Diazene-1,2-diylbis(3,5-difluoro-4,1-phenylene))bis(methylene)-bis-(2-chloroacetate) (185)

Table 1. Crystal data and structure refinement for rka1403.

Identification code	shelx	
Empirical formula	C18 H12 Cl2 F4 N2 O4	
Formula weight	467.20	
Temperature	293(2) K	
Wavelength	1.54184 \AA	
Crystal system	Monoclinic	
Space group	C 2/c	
Unit cell dimensions	a = 19.4466(7) \AA	$\alpha = 90^\circ$ .
	b = 4.7184(2) \AA	$\beta = 103.003(4)^\circ$ .
	c = 21.6117(10) \AA	$\gamma = 90^\circ$ .
Volume	1932.17(14) \AA <sup>3</sup>	

Z	4
Density (calculated)	1.606 Mg/m <sup>3</sup>
Absorption coefficient	3.656 mm <sup>-1</sup>
F(000)	944
Crystal size	0.340 x 0.086 x 0.032 mm <sup>3</sup>
Theta range for data collection	4.199 to 74.056°.
Index ranges	-24<=h<=23, -5<=k<=5, -21<=l<=26
Reflections collected	6584
Independent reflections	1922 [R(int) = 0.0164]
Completeness to theta = 67.684°	99.9 %
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	1922 / 0 / 136
Goodness-of-fit on F <sup>2</sup>	1.036
Final R indices [I>2sigma(I)]	R1 = 0.0317, wR2 = 0.0871
R indices (all data)	R1 = 0.0372, wR2 = 0.0929
Extinction coefficient	n/a
Largest diff. peak and hole	0.181 and -0.266 e.Å <sup>-3</sup>

Table 2. Atomic coordinates ( x 10<sup>4</sup>) and equivalent isotropic displacement parameters (Å<sup>2</sup>x 10<sup>3</sup>) for rka1403. U(eq) is defined as one third of the trace of the orthogonalized U<sup>ij</sup> tensor.

	x	y	z	U(eq)
N(1)	329(1)	4992(3)	5086(1)	56(1)
C(1)	3707(1)	1560(4)	3329(1)	61(1)
C(2)	3028(1)	-55(3)	3216(1)	49(1)
C(3)	1900(1)	13(4)	3420(1)	57(1)
C(4)	1499(1)	1359(3)	3861(1)	49(1)
C(5)	800(1)	539(4)	3799(1)	57(1)
C(6)	410(1)	1697(4)	4187(1)	55(1)
C(7)	678(1)	3718(3)	4649(1)	49(1)
C(8)	1383(1)	4448(3)	4699(1)	51(1)
C(9)	1794(1)	3330(3)	4318(1)	51(1)
O(1)	2889(1)	-2192(3)	2919(1)	69(1)
O(2)	2589(1)	1253(2)	3517(1)	53(1)
F(1)	-257(1)	783(3)	4129(1)	78(1)
F(2)	1669(1)	6387(3)	5138(1)	72(1)
Cl(1)	4263(1)	350(1)	2840(1)	81(1)

Table 3. Bond lengths [Å] and angles [°] for rka1403.

---

N(1)-N(1)#1	1.248(3)
N(1)-C(7)	1.415(2)
C(1)-C(2)	1.497(2)
C(1)-Cl(1)	1.7658(16)
C(1)-H(1A)	0.9700
C(1)-H(1B)	0.9700
C(2)-O(1)	1.193(2)
C(2)-O(2)	1.3345(18)
C(3)-O(2)	1.4340(19)
C(3)-C(4)	1.501(2)
C(3)-H(3A)	0.9700
C(3)-H(3B)	0.9700
C(4)-C(9)	1.384(2)
C(4)-C(5)	1.391(2)
C(5)-C(6)	1.364(2)
C(5)-H(5)	0.9300
C(6)-F(1)	1.3457(18)
C(6)-C(7)	1.394(2)
C(7)-C(8)	1.394(2)
C(8)-F(2)	1.3448(18)
C(8)-C(9)	1.376(2)
C(9)-H(9)	0.9300
N(1)#1-N(1)-C(7)	114.58(17)
C(2)-C(1)-Cl(1)	111.99(12)
C(2)-C(1)-H(1A)	109.2
Cl(1)-C(1)-H(1A)	109.2
C(2)-C(1)-H(1B)	109.2
Cl(1)-C(1)-H(1B)	109.2
H(1A)-C(1)-H(1B)	107.9
O(1)-C(2)-O(2)	124.17(15)
O(1)-C(2)-C(1)	126.99(14)
O(2)-C(2)-C(1)	108.82(13)
O(2)-C(3)-C(4)	109.72(13)
O(2)-C(3)-H(3A)	109.7

---

C(4)-C(3)-H(3A)	109.7
O(2)-C(3)-H(3B)	109.7
C(4)-C(3)-H(3B)	109.7
H(3A)-C(3)-H(3B)	108.2
C(9)-C(4)-C(5)	119.64(14)
C(9)-C(4)-C(3)	123.09(14)
C(5)-C(4)-C(3)	117.27(15)
C(6)-C(5)-C(4)	119.73(16)
C(6)-C(5)-H(5)	120.1
C(4)-C(5)-H(5)	120.1
F(1)-C(6)-C(5)	118.11(15)
F(1)-C(6)-C(7)	118.83(14)
C(5)-C(6)-C(7)	123.03(14)
C(6)-C(7)-C(8)	115.21(14)
C(6)-C(7)-N(1)	127.68(14)
C(8)-C(7)-N(1)	117.03(14)
F(2)-C(8)-C(9)	118.63(14)
F(2)-C(8)-C(7)	117.77(14)
C(9)-C(8)-C(7)	123.59(15)
C(8)-C(9)-C(4)	118.78(14)
C(8)-C(9)-H(9)	120.6
C(4)-C(9)-H(9)	120.6
C(2)-O(2)-C(3)	114.78(12)

Symmetry transformations used to generate equivalent atoms:

#1 -x,-y+1,-z+1

Table 4. Anisotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1403. The anisotropic displacement factor exponent takes the form:  $-2\pi^2 [ h^2 a^{*2} U^{11} + \dots + 2 h k a^* b^* U^{12} ]$

	$U^{11}$	$U^{22}$	$U^{33}$	$U^{23}$	$U^{13}$	$U^{12}$
N(1)	46(1)	68(1)	55(1)	-1(1)	17(1)	2(1)
C(1)	49(1)	70(1)	66(1)	-9(1)	20(1)	2(1)
C(2)	51(1)	49(1)	49(1)	2(1)	15(1)	8(1)
C(3)	55(1)	58(1)	63(1)	-6(1)	26(1)	-9(1)
C(4)	49(1)	51(1)	50(1)	5(1)	18(1)	1(1)
C(5)	54(1)	58(1)	64(1)	-7(1)	20(1)	-9(1)

C(6)	44(1)	58(1)	65(1)	0(1)	19(1)	-6(1)
C(7)	44(1)	57(1)	49(1)	4(1)	16(1)	2(1)
C(8)	46(1)	60(1)	46(1)	-2(1)	10(1)	-2(1)
C(9)	42(1)	62(1)	51(1)	3(1)	14(1)	-2(1)
O(1)	70(1)	58(1)	85(1)	-17(1)	29(1)	1(1)
O(2)	50(1)	56(1)	57(1)	-7(1)	23(1)	-2(1)
F(1)	52(1)	85(1)	105(1)	-28(1)	34(1)	-23(1)
F(2)	54(1)	96(1)	67(1)	-28(1)	16(1)	-11(1)
Cl(1)	59(1)	103(1)	90(1)	-5(1)	36(1)	11(1)

Table 5. Hydrogen coordinates ( $\times 10^4$ ) and isotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1403.

	x	y	z	U(eq)
H(1A)	3950	1365	3770	73
H(1B)	3606	3554	3246	73
H(3A)	1942	-2009	3500	68
H(3B)	1648	296	2984	68
H(5)	599	-794	3495	69
H(9)	2262	3889	4367	61

#### 7.2.4 (*E*)-(Diazene-1,2-diylbis(4,1-phenylene))-bis-(methylene)-bis(2-chloroacetate) (191)

Table 1. Crystal data and structure refinement for rka1406b.

Identification code	shelx	
Empirical formula	C <sub>18</sub> H <sub>16.20</sub> Cl <sub>1.79</sub> N <sub>2</sub> O <sub>4.20</sub>	
Formula weight	391.45	
Temperature	296(2) K	
Wavelength	1.54184 \AA	
Crystal system	Triclinic	
Space group	P -1	
Unit cell dimensions	a = 5.6216(3) \AA	$\alpha = 62.487(7)^\circ$ .
	b = 12.9368(9) \AA	$\beta = 82.950(5)^\circ$ .
	c = 13.7836(10) \AA	$\gamma = 86.880(5)^\circ$ .
Volume	882.33(11) \AA <sup>3</sup>	

Z	2
Density (calculated)	1.473 Mg/m <sup>3</sup>
Absorption coefficient	3.276 mm <sup>-1</sup>
F(000)	405
Crystal size	0.300 x 0.205 x 0.051 mm <sup>3</sup>
Theta range for data collection	3.638 to 73.921°.
Index ranges	-6<=h<=7, -15<=k<=12, -16<=l<=12
Reflections collected	5586
Independent reflections	3405 [R(int) = 0.0196]
Completeness to theta = 67.684°	99.0 %
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	3405 / 38 / 255
Goodness-of-fit on F <sup>2</sup>	1.088
Final R indices [I>2sigma(I)]	R1 = 0.0744, wR2 = 0.2161
R indices (all data)	R1 = 0.0778, wR2 = 0.2199
Extinction coefficient	n/a
Largest diff. peak and hole	1.081 and -0.419 e.Å <sup>-3</sup>

Table 2. Atomic coordinates ( x 10<sup>4</sup>) and equivalent isotropic displacement parameters (Å<sup>2</sup>x 10<sup>3</sup>) for rka1406b. U(eq) is defined as one third of the trace of the orthogonalized U<sup>ij</sup> tensor.

	x	y	z	U(eq)
C(1)	3301(6)	591(3)	-4196(2)	40(1)
C(2)	1028(6)	379(3)	-3621(3)	49(1)
C(3)	200(6)	1007(3)	-3063(3)	45(1)
C(4)	1649(6)	1849(3)	-3066(2)	40(1)
C(5)	3916(6)	2063(3)	-3658(3)	44(1)
C(6)	4751(5)	1445(3)	-4223(2)	42(1)
C(7)	751(6)	2501(3)	-2429(3)	46(1)
C(8)	2899(5)	1750(3)	-863(2)	43(1)
C(9)	2535(7)	994(3)	366(3)	53(1)
C(10)	18387(5)	5539(3)	-4135(2)	39(1)
C(11)	16153(6)	5310(3)	-3530(3)	50(1)
C(12)	15354(6)	5932(3)	-2963(3)	49(1)
C(13)	16802(5)	6773(3)	-2977(2)	39(1)
C(14)	19056(6)	6988(3)	-3576(3)	45(1)
C(15)	19847(5)	6385(3)	-4158(2)	41(1)



C(16)	15917(6)	7475(3)	-2388(3)	47(1)
C(17)	12737(5)	6826(3)	-916(3)	43(1)
C(18)	12349(7)	6094(3)	308(3)	55(1)
N(1)	3979(5)	-100(2)	-4746(2)	45(1)
N(2)	19020(5)	4868(2)	-4708(2)	43(1)
O(1)	803(4)	1760(2)	-1260(2)	47(1)
O(2)	4676(5)	2280(2)	-1403(2)	58(1)
O(3)	15060(4)	6713(2)	-1251(2)	49(1)
O(4)	11297(5)	7441(2)	-1494(2)	63(1)
Cl(1)	5223(3)	749(1)	979(1)	55(1)
Cl(2)	9283(3)	5856(1)	810(1)	63(1)
O(5)	4550(30)	850(20)	883(19)	25(4)
O(6)	10180(20)	5779(17)	974(13)	16(4)

Table 3. Bond lengths [ $\text{\AA}$ ] and angles [ $^\circ$ ] for rka1406b.

C(1)-C(2)	1.387(5)
C(1)-C(6)	1.392(4)
C(1)-N(1)	1.429(4)
C(2)-C(3)	1.385(5)
C(2)-H(2)	0.9300
C(3)-C(4)	1.394(4)
C(3)-H(3)	0.9300
C(4)-C(5)	1.394(4)
C(4)-C(7)	1.505(4)
C(5)-C(6)	1.381(4)
C(5)-H(5)	0.9300
C(6)-H(6)	0.9300
C(7)-O(1)	1.449(4)
C(7)-H(7A)	0.9700
C(7)-H(7B)	0.9700
C(8)-O(2)	1.201(4)
C(8)-O(1)	1.355(4)
C(8)-C(9)	1.508(4)
C(9)-O(5)	1.368(10)
C(9)-Cl(1)	1.764(4)
C(9)-H(9A)	0.9700
C(9)-H(9B)	0.9700

---

C(10)-C(11)	1.384(4)
C(10)-C(15)	1.390(4)
C(10)-N(2)	1.427(4)
C(11)-C(12)	1.384(5)
C(11)-H(11)	0.9300
C(12)-C(13)	1.385(5)
C(12)-H(12)	0.9300
C(13)-C(14)	1.390(4)
C(13)-C(16)	1.505(4)
C(14)-C(15)	1.379(4)
C(14)-H(14)	0.9300
C(15)-H(15)	0.9300
C(16)-O(3)	1.449(4)
C(16)-H(16A)	0.9700
C(16)-H(16B)	0.9700
C(17)-O(4)	1.194(4)
C(17)-O(3)	1.354(4)
C(17)-C(18)	1.499(4)
C(18)-O(6)	1.387(10)
C(18)-Cl(2)	1.764(4)
C(18)-H(18A)	0.9700
C(18)-H(18B)	0.9700
N(1)-N(1)#1	1.241(6)
N(2)-N(2)#2	1.244(5)
C(2)-C(1)-C(6)	120.1(3)
C(2)-C(1)-N(1)	115.7(3)
C(6)-C(1)-N(1)	124.1(3)
C(3)-C(2)-C(1)	120.2(3)
C(3)-C(2)-H(2)	119.9
C(1)-C(2)-H(2)	119.9
C(2)-C(3)-C(4)	120.3(3)
C(2)-C(3)-H(3)	119.8
C(4)-C(3)-H(3)	119.8
C(3)-C(4)-C(5)	118.9(3)
C(3)-C(4)-C(7)	119.7(3)
C(5)-C(4)-C(7)	121.4(3)
C(6)-C(5)-C(4)	121.2(3)

---

---

C(6)-C(5)-H(5)	119.4
C(4)-C(5)-H(5)	119.4
C(5)-C(6)-C(1)	119.3(3)
C(5)-C(6)-H(6)	120.3
C(1)-C(6)-H(6)	120.3
O(1)-C(7)-C(4)	110.5(2)
O(1)-C(7)-H(7A)	109.6
C(4)-C(7)-H(7A)	109.6
O(1)-C(7)-H(7B)	109.6
C(4)-C(7)-H(7B)	109.6
H(7A)-C(7)-H(7B)	108.1
O(2)-C(8)-O(1)	125.3(3)
O(2)-C(8)-C(9)	127.3(3)
O(1)-C(8)-C(9)	107.4(3)
O(5)-C(9)-C(8)	114.3(11)
O(5)-C(9)-Cl(1)	4.3(11)
C(8)-C(9)-Cl(1)	112.5(2)
O(5)-C(9)-H(9A)	108.7
C(8)-C(9)-H(9A)	108.7
Cl(1)-C(9)-H(9A)	113.0
O(5)-C(9)-H(9B)	108.7
C(8)-C(9)-H(9B)	108.7
Cl(1)-C(9)-H(9B)	106.2
H(9A)-C(9)-H(9B)	107.6
C(11)-C(10)-C(15)	119.7(3)
C(11)-C(10)-N(2)	115.9(3)
C(15)-C(10)-N(2)	124.4(3)
C(12)-C(11)-C(10)	120.3(3)
C(12)-C(11)-H(11)	119.9
C(10)-C(11)-H(11)	119.9
C(11)-C(12)-C(13)	120.3(3)
C(11)-C(12)-H(12)	119.8
C(13)-C(12)-H(12)	119.8
C(12)-C(13)-C(14)	119.0(3)
C(12)-C(13)-C(16)	120.5(3)
C(14)-C(13)-C(16)	120.4(3)
C(15)-C(14)-C(13)	120.9(3)
C(15)-C(14)-H(14)	119.5

---

C(13)-C(14)-H(14)	119.5
C(14)-C(15)-C(10)	119.7(3)
C(14)-C(15)-H(15)	120.2
C(10)-C(15)-H(15)	120.2
O(3)-C(16)-C(13)	110.5(2)
O(3)-C(16)-H(16A)	109.6
C(13)-C(16)-H(16A)	109.6
O(3)-C(16)-H(16B)	109.6
C(13)-C(16)-H(16B)	109.6
H(16A)-C(16)-H(16B)	108.1
O(4)-C(17)-O(3)	125.6(3)
O(4)-C(17)-C(18)	126.6(3)
O(3)-C(17)-C(18)	107.8(3)
O(6)-C(18)-C(17)	127.7(8)
O(6)-C(18)-Cl(2)	15.4(8)
C(17)-C(18)-Cl(2)	112.4(3)
O(6)-C(18)-H(18A)	105.4
C(17)-C(18)-H(18A)	105.4
Cl(2)-C(18)-H(18A)	111.2
O(6)-C(18)-H(18B)	105.4
C(17)-C(18)-H(18B)	105.4
Cl(2)-C(18)-H(18B)	115.7
H(18A)-C(18)-H(18B)	106.0
N(1)#1-N(1)-C(1)	114.8(3)
N(2)#2-N(2)-C(10)	114.7(3)
C(8)-O(1)-C(7)	115.7(2)
C(17)-O(3)-C(16)	117.0(2)

Symmetry transformations used to generate equivalent atoms:

#1 -x+1,-y,-z-1 #2 -x+4,-y+1,-z-1

Table 4. Anisotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1406b. The anisotropic displacement factor exponent takes the form:  $-2\pi^2 [ h^2 a^* U^{11} + \dots + 2 h k a^* b^* U^{12} ]$

	U <sup>11</sup>	U <sup>22</sup>	U <sup>33</sup>	U <sup>23</sup>	U <sup>13</sup>	U <sup>12</sup>
C(1)	44(2)	41(2)	33(1)	-14(1)	-5(1)	0(1)
C(2)	48(2)	49(2)	47(2)	-22(1)	2(1)	-12(1)

C(3)	36(2)	52(2)	39(2)	-16(1)	4(1)	-3(1)
C(4)	43(2)	42(2)	31(1)	-13(1)	-7(1)	6(1)
C(5)	42(2)	48(2)	45(2)	-23(1)	-1(1)	-6(1)
C(6)	37(2)	49(2)	41(2)	-21(1)	-1(1)	-3(1)
C(7)	43(2)	51(2)	40(2)	-18(1)	-4(1)	8(1)
C(8)	38(2)	47(2)	46(2)	-25(1)	-2(1)	2(1)
C(9)	52(2)	60(2)	47(2)	-25(2)	-4(1)	2(2)
C(10)	40(2)	41(1)	33(1)	-16(1)	-1(1)	1(1)
C(11)	46(2)	53(2)	53(2)	-28(2)	11(1)	-14(1)
C(12)	44(2)	55(2)	45(2)	-23(1)	12(1)	-8(1)
C(13)	42(2)	39(1)	30(1)	-13(1)	-2(1)	4(1)
C(14)	41(2)	49(2)	47(2)	-26(1)	2(1)	-7(1)
C(15)	34(1)	50(2)	42(2)	-24(1)	4(1)	-4(1)
C(16)	52(2)	43(2)	40(2)	-16(1)	6(1)	1(1)
C(17)	42(2)	44(2)	48(2)	-26(1)	-7(1)	5(1)
C(18)	56(2)	59(2)	47(2)	-23(2)	4(2)	2(2)
N(1)	45(1)	48(1)	43(1)	-22(1)	-2(1)	-2(1)
N(2)	44(1)	47(1)	42(1)	-24(1)	3(1)	-3(1)
O(1)	46(1)	54(1)	39(1)	-21(1)	0(1)	-2(1)
O(2)	47(1)	68(2)	56(1)	-26(1)	4(1)	-7(1)
O(3)	55(1)	50(1)	35(1)	-17(1)	1(1)	13(1)
O(4)	48(1)	75(2)	59(2)	-23(1)	-12(1)	11(1)
Cl(1)	61(1)	53(1)	57(1)	-29(1)	-20(1)	8(1)
Cl(2)	60(1)	58(1)	72(1)	-35(1)	26(1)	-13(1)
O(5)	25(5)	24(4)	24(4)	-12(2)	2(2)	1(2)
O(6)	17(4)	18(4)	14(4)	-8(2)	-2(2)	0(2)

Table 5. Hydrogen coordinates ( $\times 10^4$ ) and isotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1406b.

	x	y	z	U(eq)
H(2)	58	-187	-3612	58
H(3)	-1331	866	-2684	54
H(5)	4885	2632	-3674	53
H(6)	6265	1598	-4617	50
H(7A)	1746	3182	-2665	55

H(7B)	-877	2762	-2577	55
H(9A)	1290	1337	679	64
H(9B)	1973	234	515	64
H(11)	15185	4735	-3503	60
H(12)	13837	5785	-2571	59
H(14)	20043	7547	-3585	53
H(15)	21350	6544	-4563	50
H(16A)	17208	7960	-2416	56
H(16B)	14628	7983	-2753	56
H(18A)	13171	5366	454	66
H(18B)	13251	6477	608	66

### 7.2.5 (*E*)-(Diazene-1,2-diylbis(3-bromo-4,1-phenylene))-bis-(methylene)-bis-(2-chloroacetate) (209)

Table 1. Crystal data and structure refinement for rka1405a.

Identification code	shelx		
Empirical formula	C18 H14 Br2 Cl2 N2 O4		
Formula weight	553.03		
Temperature	296(2) K		
Wavelength	1.54184 Å		
Crystal system	Monoclinic		
Space group	P 21/c		
Unit cell dimensions	a = 4.5764(2) Å	$\alpha = 90^\circ$ .	
	b = 29.1072(12) Å	$\beta = 103.781(4)^\circ$ .	
	c = 7.8523(3) Å	$\gamma = 90^\circ$ .	
Volume	1015.86(7) Å <sup>3</sup>		
Z	2		
Density (calculated)	1.808 Mg/m <sup>3</sup>		
Absorption coefficient	7.731 mm <sup>-1</sup>		
F(000)	544		
Crystal size	0.380 x 0.084 x 0.050 mm <sup>3</sup>		
Theta range for data collection	5.998 to 74.080°.		
Index ranges	-3 ≤ h ≤ 5, -35 ≤ k ≤ 34, -9 ≤ l ≤ 5		
Reflections collected	3428		
Independent reflections	1941 [R(int) = 0.0193]		
Completeness to theta = 67.684°	97.5 %		
Refinement method	Full-matrix least-squares on F <sup>2</sup>		

Data / restraints / parameters	1941 / 0 / 127
Goodness-of-fit on $F^2$	1.229
Final R indices [ $I > 2\sigma(I)$ ]	R1 = 0.0420, wR2 = 0.1101
R indices (all data)	R1 = 0.0465, wR2 = 0.1140
Extinction coefficient	n/a
Largest diff. peak and hole	0.488 and -0.394 e.Å <sup>-3</sup>

Table 2. Atomic coordinates ( $\times 10^4$ ) and equivalent isotropic displacement parameters ( $\text{Å}^2 \times 10^3$ ) for rka1405a.  $U(\text{eq})$  is defined as one third of the trace of the orthogonalized  $U^{\text{ij}}$  tensor.

	x	y	z	U(eq)
C(1)	6860(9)	1147(1)	5611(6)	37(1)
C(2)	5601(10)	1026(1)	3893(6)	39(1)
C(3)	3477(10)	676(1)	3529(6)	36(1)
C(4)	2576(10)	443(1)	4862(6)	34(1)
C(5)	3879(10)	563(2)	6582(6)	39(1)
C(6)	6017(10)	911(2)	6959(6)	40(1)
C(7)	9171(10)	1523(2)	6014(7)	44(1)
C(8)	9139(10)	2332(2)	6238(6)	40(1)
C(9)	7239(11)	2751(2)	5719(7)	50(1)
N(1)	361(8)	94(1)	4372(5)	37(1)
O(1)	7582(7)	1960(1)	5607(5)	49(1)
O(2)	11634(9)	2325(1)	7109(6)	68(1)
Cl(1)	9414(3)	3258(1)	6012(2)	62(1)
Br(1)	1833(1)	518(1)	1150(1)	57(1)

Table 3. Bond lengths [Å] and angles [°] for rka1405a.

---

C(1)-C(2)	1.381(7)
C(1)-C(6)	1.391(7)
C(1)-C(7)	1.503(6)
C(2)-C(3)	1.391(6)
C(3)-C(4)	1.389(6)
C(3)-Br(1)	1.896(4)
C(4)-C(5)	1.385(6)
C(4)-N(1)	1.423(5)
C(5)-C(6)	1.390(6)
C(7)-O(1)	1.463(5)
C(8)-O(2)	1.183(6)
C(8)-O(1)	1.325(5)
C(8)-C(9)	1.499(6)
C(9)-Cl(1)	1.763(5)
N(1)-N(1)#1	1.239(7)
C(2)-C(1)-C(6)	119.4(4)
C(2)-C(1)-C(7)	120.0(4)
C(6)-C(1)-C(7)	120.5(4)
C(1)-C(2)-C(3)	119.8(4)
C(4)-C(3)-C(2)	121.4(4)
C(4)-C(3)-Br(1)	120.3(3)
C(2)-C(3)-Br(1)	118.3(3)
C(5)-C(4)-C(3)	118.5(4)
C(5)-C(4)-N(1)	123.9(4)
C(3)-C(4)-N(1)	117.7(4)
C(4)-C(5)-C(6)	120.6(4)
C(5)-C(6)-C(1)	120.4(4)
O(1)-C(7)-C(1)	107.4(3)
O(2)-C(8)-O(1)	124.3(4)
O(2)-C(8)-C(9)	126.0(4)
O(1)-C(8)-C(9)	109.7(4)
C(8)-C(9)-Cl(1)	111.9(3)
N(1)#1-N(1)-C(4)	114.1(5)
C(8)-O(1)-C(7)	115.7(3)

---

---



Symmetry transformations used to generate equivalent atoms:

#1 -x,-y,-z+1

Table 4. Anisotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1405a. The anisotropic displacement factor exponent takes the form:  $-2\pi^2 [ h^2 a^{*2} U^{11} + \dots + 2 h k a^* b^* U^{12} ]$

	$U^{11}$	$U^{22}$	$U^{33}$	$U^{23}$	$U^{13}$	$U^{12}$
C(1)	26(2)	30(2)	56(3)	-3(2)	8(2)	2(2)
C(2)	40(2)	28(2)	50(3)	2(2)	14(2)	0(2)
C(3)	41(2)	30(2)	35(2)	-1(2)	5(2)	4(2)
C(4)	35(2)	23(2)	42(2)	-1(2)	7(2)	1(2)
C(5)	40(2)	35(2)	44(2)	3(2)	11(2)	1(2)
C(6)	38(2)	37(2)	42(2)	-3(2)	2(2)	0(2)
C(7)	31(2)	32(2)	65(3)	0(2)	3(2)	-1(2)
C(8)	37(2)	35(2)	44(3)	0(2)	6(2)	-5(2)
C(9)	46(3)	35(2)	64(3)	-2(2)	5(2)	-5(2)
N(1)	36(2)	32(2)	44(2)	-1(2)	11(2)	-6(2)
O(1)	37(2)	30(2)	70(2)	-7(2)	-4(2)	-2(1)
O(2)	50(2)	44(2)	93(3)	8(2)	-19(2)	-11(2)
Cl(1)	71(1)	33(1)	87(1)	-6(1)	29(1)	-8(1)
Br(1)	76(1)	53(1)	38(1)	-2(1)	6(1)	-13(1)

Table 5. Hydrogen coordinates ( $\times 10^4$ ) and isotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1405a.

	x	y	z	U(eq)
H(2)	6171	1178	2982	47
H(5)	3319	409	7493	47
H(6)	6889	986	8119	48
H(7A)	10222	1513	7242	53
H(7B)	10632	1486	5310	53
H(9A)	5757	2770	6418	59
H(9B)	6170	2726	4498	59

### 7.2.6 (E)-(Diazene-1,2-diylbis(3-bromo-5-fluoro-4,1-phenylene))-bis-(methylene)-bis-(2-chloroacetate) (221)

Table 1. Crystal data and structure refinement for rka1505.

Identification code	rka1505	
Empirical formula	C <sub>9</sub> H <sub>6</sub> Br Cl F N O <sub>2</sub>	
Formula weight	294.51	
Temperature	298(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P 21/n	
Unit cell dimensions	a = 4.5876(4) Å	α = 90°.
	b = 29.039(3) Å	β = 97.380(9)°.
	c = 7.7651(7) Å	γ = 90°.
Volume	1025.90(16) Å <sup>3</sup>	
Z	4	
Density (calculated)	1.907 Mg/m <sup>3</sup>	
Absorption coefficient	4.258 mm <sup>-1</sup>	
F(000)	576	
Crystal size	0.307 x 0.072 x 0.030 mm <sup>3</sup>	
Theta range for data collection	3.381 to 29.992°.	
Index ranges	-6 ≤ h ≤ 5, -39 ≤ k ≤ 36, -10 ≤ l ≤ 5	
Reflections collected	5191	
Independent reflections	2451 [R(int) = 0.0315]	
Completeness to theta = 25.242°	99.6 %	
Refinement method	Full-matrix least-squares on F <sup>2</sup>	
Data / restraints / parameters	2451 / 0 / 136	
Goodness-of-fit on F <sup>2</sup>	0.999	
Final R indices [I > 2σ(I)]	R1 = 0.0468, wR2 = 0.0986	
R indices (all data)	R1 = 0.0749, wR2 = 0.1126	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.573 and -0.560 e.Å <sup>-3</sup>	

Table 2. Atomic coordinates ( x 10<sup>4</sup>) and equivalent isotropic displacement parameters (Å<sup>2</sup> x 10<sup>3</sup>) for rka1505. U(eq) is defined as one third of the trace of the orthogonalized U<sup>ij</sup> tensor.

	x	y	z	U(eq)
C(1)	10402(11)	2379(2)	-380(8)	75(2)
C(2)	12132(9)	1943(1)	-320(5)	46(1)
C(3)	12521(8)	1217(1)	1040(6)	43(1)
C(4)	10712(8)	933(1)	2127(5)	38(1)
C(5)	10202(8)	1091(1)	3738(5)	38(1)
C(6)	8428(8)	834(1)	4699(5)	37(1)
C(7)	7228(8)	417(1)	4120(5)	37(1)
C(8)	7841(9)	268(1)	2504(5)	43(1)
C(9)	9510(9)	522(1)	1508(5)	44(1)
N(1)	5473(9)	177(1)	5215(5)	60(1)
O(1)	14179(7)	1857(1)	-1056(4)	64(1)
O(2)	11007(6)	1648(1)	734(4)	49(1)
F(1)	6726(7)	-132(1)	1849(4)	76(1)
Cl(1)	11912(3)	2814(1)	-1532(2)	82(1)
Br(1)	7509(1)	1084(1)	6810(1)	61(1)

Table 3. Bond lengths [ $\text{\AA}$ ] and angles [ $^\circ$ ] for rka1505.

C(1)-C(2)	1.494(6)
C(1)-Cl(1)	1.741(5)
C(1)-H(1A)	0.9700
C(1)-H(1B)	0.9700
C(2)-O(1)	1.187(5)
C(2)-O(2)	1.333(4)
C(3)-O(2)	1.435(4)
C(3)-C(4)	1.505(5)
C(3)-H(3A)	0.9700
C(3)-H(3B)	0.9700
C(4)-C(9)	1.377(5)
C(4)-C(5)	1.380(5)
C(5)-C(6)	1.389(5)
C(5)-H(5)	0.9300
C(6)-C(7)	1.383(5)
C(6)-Br(1)	1.889(4)
C(7)-C(8)	1.390(5)

---

C(7)-N(1)	1.425(5)
C(8)-F(1)	1.343(4)
C(8)-C(9)	1.371(5)
C(9)-H(9)	0.9300
N(1)-N(1)#1	1.150(7)
C(2)-C(1)-Cl(1)	112.7(3)
C(2)-C(1)-H(1A)	109.0
Cl(1)-C(1)-H(1A)	109.0
C(2)-C(1)-H(1B)	109.0
Cl(1)-C(1)-H(1B)	109.0
H(1A)-C(1)-H(1B)	107.8
O(1)-C(2)-O(2)	123.7(4)
O(1)-C(2)-C(1)	127.7(4)
O(2)-C(2)-C(1)	108.6(3)
O(2)-C(3)-C(4)	106.3(3)
O(2)-C(3)-H(3A)	110.5
C(4)-C(3)-H(3A)	110.5
O(2)-C(3)-H(3B)	110.5
C(4)-C(3)-H(3B)	110.5
H(3A)-C(3)-H(3B)	108.7
C(9)-C(4)-C(5)	119.9(3)
C(9)-C(4)-C(3)	120.4(4)
C(5)-C(4)-C(3)	119.7(4)
C(4)-C(5)-C(6)	119.2(3)
C(4)-C(5)-H(5)	120.4
C(6)-C(5)-H(5)	120.4
C(7)-C(6)-C(5)	122.2(3)
C(7)-C(6)-Br(1)	119.7(3)
C(5)-C(6)-Br(1)	118.1(3)
C(6)-C(7)-C(8)	116.6(3)
C(6)-C(7)-N(1)	117.8(3)
C(8)-C(7)-N(1)	125.7(3)
F(1)-C(8)-C(9)	117.7(3)
F(1)-C(8)-C(7)	119.9(3)
C(9)-C(8)-C(7)	122.4(4)
C(8)-C(9)-C(4)	119.8(4)
C(8)-C(9)-H(9)	120.1

---

C(4)-C(9)-H(9)	120.1
N(1)#1-N(1)-C(7)	119.0(5)
C(2)-O(2)-C(3)	116.3(3)

Symmetry transformations used to generate equivalent atoms:

#1 -x+1,-y,-z+1

Table 4. Anisotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1505. The anisotropic displacement factor exponent takes the form:  $-2\pi^2 [ h^2 a^{*2} U^{11} + \dots + 2 h k a^* b^* U^{12} ]$

	$U^{11}$	$U^{22}$	$U^{33}$	$U^{23}$	$U^{13}$	$U^{12}$
C(1)	67(3)	61(3)	103(4)	37(3)	35(3)	10(3)
C(2)	45(2)	48(2)	46(2)	9(2)	12(2)	-9(2)
C(3)	41(2)	42(2)	49(2)	7(2)	16(2)	4(2)
C(4)	36(2)	37(2)	42(2)	4(2)	9(2)	4(2)
C(5)	38(2)	35(2)	42(2)	1(2)	7(2)	-4(2)
C(6)	41(2)	39(2)	32(2)	-3(2)	7(2)	2(2)
C(7)	45(2)	34(2)	33(2)	1(2)	12(2)	-1(2)
C(8)	60(2)	31(2)	41(2)	-3(2)	16(2)	-2(2)
C(9)	59(2)	39(2)	37(2)	-1(2)	22(2)	1(2)
N(1)	91(3)	47(2)	50(2)	-8(2)	39(2)	-24(2)
O(1)	71(2)	57(2)	72(2)	3(2)	40(2)	-8(2)
O(2)	47(2)	44(2)	60(2)	18(1)	24(1)	6(1)
F(1)	126(2)	48(2)	64(2)	-22(1)	55(2)	-34(2)
Cl(1)	93(1)	57(1)	102(1)	30(1)	33(1)	-5(1)
Br(1)	83(1)	64(1)	39(1)	-16(1)	21(1)	-18(1)

Table 5. Hydrogen coordinates ( $\times 10^4$ ) and isotropic displacement parameters ( $\text{\AA}^2 \times 10^3$ ) for rka1505.

	x	y	z	U(eq)
H(1A)	10303	2484	797	90
H(1B)	8411	2319	-916	90
H(3A)	14477	1267	1650	52
H(3B)	12695	1063	-51	52
H(5)	11035	1366	4176	46
H(9)	9829	416	417	52

## 8 References

- 1 L. Lo Conte, C. Chothia and J. Janin, *J. Mol. Biol.*, 1999, **285**, 2177–2198.
- 2 T. Clackson and J. A. Wells, *Science*, 1995, **267**, 383–6.
- 3 M. R. Arkin and J. A. Wells, *Nat. Rev. Drug Discov.*, 2004, **3**, 301–17.
- 4 D. C. Burns, F. Zhang and G. A. Woolley, *Nat. Protoc.*, 2007, **2**, 251–258.
- 5 E. J. Sundberg and R. A. Mariuzza, *Structure*, 2000, **8**, R137–R142.
- 6 B. Hall, A. Limaye and A. B. Kulkarni, *Overview: Generation of Gene Knockout Mice*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2009, vol. Chapter 19.
- 7 S. L. Mansour, K. R. Thomas and M. R. Capecchi, *Nature*, 1988, **336**, 348–52.
- 8 Y. Santiago, E. Chan, P.-Q. Liu, S. Orlando, L. Zhang, F. D. Urnov, M. C. Holmes, D. Guschin, A. Waite, J. C. Miller, E. J. Rebar, P. D. Gregory, A. Klug and T. N. Collingwood, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 5809–14.
- 9 D. C. van Gent and M. van der Burg, *Oncogene*, 2007, **26**, 7731–7740.
- 10 T. Berg, *Curr. Opin. Chem. Biol.*, 2008, **12**, 464–71.
- 11 L. D. Walensky, K. Pitter, J. Morash, K. J. Oh, S. Barbuto, J. Fisher, E. Smith, G. L. Verdine and S. J. Korsmeyer, *Mol. Cell*, 2006, **24**, 199–210.
- 12 S. Kneissl, E. J. Loveridge, C. Williams, M. P. Crump and R. K. Allemann, *ChemBioChem*, 2008, **9**, 3046–54.
- 13 W. A. Velema, W. Szymanski and B. L. Feringa, *J. Am. Chem. Soc.*, 2014, **136**, 2178–2191.
- 14 J. A. DiMasi, L. Feldman, A. Seckler and A. Wilson, *Clin. Pharmacol. Ther.*, 2010, **87**, 272–7.
- 15 M. Skwarczynski, M. Noguchi, S. Hirota, Y. Sohma, T. Kimura, Y. Hayashi and Y. Kiso, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 4492–6.
- 16 F. G. Cruz, J. T. Koh and K. H. Link, *J. Am. Chem. Soc.*, 2000, **122**, 8777–8778.
- 17 F. C. Dalman, L. J. Sturzenbecker, A. A. Levin, D. A. Lucas, G. H. Perdew, M. Petkovitch, P. Chambon, J. F. Grippo and W. B. Pratt, *Biochemistry*, 1991, **30**, 5605–5608.
- 18 A. P. Pelliccioli and J. Wirz, *Photochem. Photobiol. Sci.*, 2002, **1**, 441–458.
- 19 M. S. Kim and S. L. Diamond, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 4007–10.
- 20 J. H. Kaplan, B. Forbush and J. F. Hoffman, *Biochemistry*, 1978, **17**, 1929–1935.
- 21 C. Chou and A. Deiters, *Angew. Chemie - Int. Ed.*, 2011, **50**, 6839–6842.
- 22 J. Li, G. Li, H. Wang and X. Wang Deng, *Arab. B.*, 2011, **9**, e0148.
- 23 F.-W. Li, M. Melkonian, C. J. Rothfels, J. C. Villarreal, D. W. Stevenson, S. W. Graham, G. K.-S. Wong, K. M. Pryer and S. Mathews, *Nat. Commun.*, 2015, **6**, 7852.

- 24 T. E. Swartz, S. B. Corchnoy, J. M. Christie, J. W. Lewis, I. Szundi, W. R. Briggs and R. A. Bogomolni, *J. Biol. Chem.*, 2001, **276**, 36493–500.
- 25 R. Mart, D. Meah and R. K. Allemann, *ChemBioChem*, 2015, **136**, 2178–2191.
- 26 S. M. Harper, J. M. Christie and K. H. Gardner, *Biochemistry*, 2004, **43**, 16184–16192.
- 27 O. Pieroni, a. Fissi, N. Angelini and F. Lenci, *Acc. Chem. Res.*, 2001, **34**, 9–17.
- 28 Y. Liu, M. Fan, S. Zhang, X. Sheng and J. Yao, *New J. Chem.*, 2007, **31**, 1878.
- 29 K. Fujimoto, M. Amano, Y. Horibe and M. Inouye, *Org. Lett.*, 2006, **8**, 285–287.
- 30 Y. Nakayama, K. Hayashi and M. Irie, *Bull. Chem. Soc. Jpn.*, 1991, **64**, 789–795.
- 31 M. Irie, *Chem. Rev.*, 2000, **100**, 1685–1716.
- 32 H. Cahová and A. Jäschke, *Angew. Chem. Int. Ed.* 2013, **52**, 3186–90.
- 33 M. Ibrahim, M. M. El-Nahass, M. A. Kamel, A. A. El-Barbary, B. D. Wagner and M. A. M. El-Mansy, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, 2013, **113**, 332–336.
- 34 J. Z. Vlahakis, M. D. Wand and R. P. Lemieux, *J. Am. Chem. Soc.*, 2003, **125**, 6862–6863.
- 35 R. P. Lemieux, *Chem. Rec.*, 2004, **3**, 288–295.
- 36 R. J. Mart, *Pers. Comun.*
- 37 N. S. Dokunikhin and Y. E. Gerasimenko, *Zhurnal Obs. Khimii*, 1960, **30**, 635–638.
- 38 M. Kato, *J. Am. Chem. Soc.*, 1985, **107**, 1024–1028.
- 39 Demselben, *Ann. der Pharm.*, 1834, **12**, 311–314.
- 40 J. Monteath Robertson, M. Prasad and I. Woodward, *Proc. R. Soc. London*, 1936, **154**, 189–195.
- 41 G. S. HARTLEY, *Nature*, 1937, **140**, 281–281.
- 42 J. M. Robertson, *J. Chem. Soc.*, 1939, 232.
- 43 G. S. Hartley, *J. Chem. Soc.*, 1938, 633.
- 44 H. Dahn and H. v. Castelmur, *Helv. Chim. Acta*, 1953, **36**, 638–645.
- 45 A. Shaabani and M. Zahedi, in *Journal of Molecular Structure*, 2000, vol. 506, pp. 257–261.
- 46 G. Angelini, N. Canilho, M. Emo, M. Kingsley and C. Gasbarri, *J. Org. Chem.*, 2015, **80**, 7430–4.
- 47 N. Nishimura, T. Tanaka, M. Asano and Y. Sueishi, *J. Chem. Soc. Perkin Trans. 2*, 1986, 1839.
- 48 A. Cembran, F. Bernardi, M. Garavelli, L. Gagliardi and G. Orlandi, 2004, 3234–3243.
- 49 P. Cattaneo and M. Persico, *Phys. Chem. Chem. Phys.*, 1999, **1**, 4739–4743.



- 50 P. P. Birnbaum, J. H. Linford and D. W. G. Style, *Trans. Faraday Soc.*, 1953, **49**, 735.
- 51 K. S. Schanze, T. F. Mattox and D. G. Whitten, *J. Org. Chem.*, 1983, **48**, 2808–2813.
- 52 T. J. McMillan, E. Leatherman, a Ridley, J. Shorrocks, S. E. Tobi and J. R. Whiteside, *J. Pharm. Pharmacol.*, 2008, **60**, 969–976.
- 53 J. Cadet, C. Anselmino, T. Douki and L. Voituriez, *J. Photochem. Photobiol. B.*, 1992, **15**, 277–98.
- 54 O. Sadovskii, A. A. Beharry, F. Zhang and G. A. Woolley, *Angew. Chem. Int. Ed.*, 2009, **48**, 1484–6.
- 55 A. Mourot, T. Fehrentz, Y. Le Feuvre, C. M. Smith, C. Herold, D. Dalkara, F. Nagy, D. Trauner and R. H. Kramer, *Nat. Methods*, 2012, **9**, 396–402.
- 56 M. Banghart, A. Mourot, D. Fortin, J. Yao, R. Kramer and D. Trauner, *Angew. Chem. Int. Ed.*, 2009, **48**, 9097–9101.
- 57 A. A. Beharry, O. Sadovskii and G. A. Woolley, *J. Am. Chem. Soc.*, 2011, **133**, 19684–19687.
- 58 S. Samanta, A. Babalhavaeji, M. X. Dong and G. A. Woolley, *Angew. Chem, Int. Ed.*, 2013, **52**, 14127–14130.
- 59 S. Samanta, T. M. McCormick, S. K. Schmidt, D. S. Seferos and G. A. Woolley, *Chem. Commun. (Camb).*, 2013, **49**, 10314–6.
- 60 C. Knie, M. Utecht, F. Zhao, H. Kulla, S. Kovalenko, A. M. Brouwer, P. Saalfrank, S. Hecht and D. Bléger, *Chem. - A Eur. J.*, 2014, **20**, 16492–16501.
- 61 D. Bléger, J. Schwarz, A. M. Brouwer and S. Hecht, *J. Am. Chem. Soc.*, 2012, **134**, 20597–20600.
- 62 S. Samanta, A. A. Beharry, O. Sadovskii, T. M. McCormick, A. Babalhavaeji, V. Tropepe and G. A. Woolley, *J. Am. Chem. Soc.*, 2013, **135**, 9777–9784.
- 63 C. L. Forber, E. C. Kelusky, N. J. Bunce and M. C. Zerner, *J. Am. Chem. Soc.*, 1985, **107**, 5884–5890.
- 64 S. Sawada, N. Kato and K. Kaihatsu, *Curr. Pharm. Biotechnol.*, 2012, **13**, 2642–2648.
- 65 N. Pozhidaeva, M. E. Cormier, A. Chaudhari and G. A. Woolley, in *Bioconjug Chemistry*, 2004, vol. 15, pp. 1297–1303.
- 66 H. Nishioka, X. Liang, T. Kato and H. Asanuma, *Angew. Chemie - Int. Ed.*, 2012, **51**, 1165–1168.
- 67 L. Chi, O. Sadovskii and G. A. Woolley, *Bioconjug. Chem.*, 2006, **17**, 670–676.
- 68 C. Boulègue, M. Löweneck, C. Renner and L. Moroder, *ChemBioChem*, 2007, **8**, 591–4.
- 69 M. Borowiak, W. Nahaboo, M. Reynders, K. Nekolla, P. Jalinot, J. Hasserodt, M. Rehberg, M. Delattre, S. Zahler, A. Vollmar, D. Trauner and O. Thorn-Seshold, *Cell*, 2015, **162**, 403–411.

- 70 S. Pittolo, X. Gómez-Santacana, K. Eckelt, X. Rovira, J. Dalton, C. Goudet, J.-P. Pin, A. Llobet, J. Giraldo, A. Llebaria and P. Gorostiza, *Nat. Chem. Biol.*, 2014, **10**, 813–5.
- 71 A. Nomura and A. Okamoto, *Chem. Commun.*, 2009, 1906–1908.
- 72 O. Pieroni, J. . Houben, A. Fissi, P. Constantino and F. Ciardelli, *J. Am. Chem. Soc.*, 1980, **102**, 5913–5915.
- 73 K. Nakayama, M. Endo and T. Majima, *Bioconjug. Chem.*, 2005, **16**, 1360–6.
- 74 J. Broichhagen, T. Podewin, H. Meyer-Berg, Y. von Ohlen, N. R. Johnston, B. J. Jones, S. R. Bloom, G. A. Rutter, A. Hoffmann-Röder, D. J. Hodson and D. Trauner, *Angew. Chem. Int. Ed.*, 2015, **54**, 15565–15569.
- 75 M. Bose, D. Groff, J. Xie, E. Brustad and P. G. Schultz, *J. Am. Chem. Soc.*, 2006, **128**, 388–9.
- 76 C. Hoppmann, P. Schmieder, N. Heinrich and M. Beyermann, *ChemBioChem*, 2011, **12**, 2555–9.
- 77 C. Hoppmann, V. K. Lacey, G. V Louie, J. Wei, J. P. Noel and L. Wang, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 3932–6.
- 78 J. R. Kumita, O. S. Smart and G. A. Woolley, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 3803–3808.
- 79 J. R. Kumita, D. G. Flint, O. S. Smart and G. A. Woolley, *Protein Eng.*, 2002, **15**, 561–569.
- 80 F. Bonardi, G. London, N. Nouwen, B. L. Feringa and A. J. M. Driessen, *Angew. Chem., Int. Ed.*, 2010, **49**, 7234–7238.
- 81 Z. Zhang, D. C. Burns, J. R. Kumita, O. S. Smart and G. A. Woolley, *Bioconjug Chemistry*, 2003, **3**, 824–829.
- 82 D. G. Flint, J. R. Kumita, O. S. Smart and G. A. Woolley, *Chem. Biol.*, 2002, **9**, 391–397.
- 83 Y. Zhang, F. Erdmann and G. Fischer, *Nat. Chem. Biol.*, 2009, **5**, 724–6.
- 84 R. J. Mart, P. Wysoczański, S. Kneissl, A. Ricci, A. Brancale and R. K. Allemann, *ChemBioChem*, 2012, **13**, 515–9.
- 85 F. Zhang, K. A. Timm, K. M. Arndt and G. A. Woolley, *Angew. Chem. Int. Ed.*, 2010, **49**, 3943–3946.
- 86 G. A. Woolley, A. S. I. Jaikaran, M. Berezovski, J. P. Calarco, S. N. Krylov, O. S. Smart and J. R. Kumita, *Biochemistry*, 2006, **45**, 6075–84.
- 87 H. Asanuma, T. Ito, T. Yoshida, X. Liang and M. Komiyama, *Angew. Chem. Int. Ed. Engl.*, 1999, **38**, 2393–2395.
- 88 H. Asanuma, T. Takarada, T. Yoshida, D. Tamaru, X. Liang and M. Komiyama, *Angew. Chem. Int. Ed. Engl.*, 2001, **40**, 2671–2673.
- 89 Y. Kamiya, T. Takagi, H. Ooi, H. Ito, X. Liang and H. Asanuma, *ACS Synth. Biol.*, 2015, **4**, 365–70.
- 90 [http://geneed.nlm.nih.gov/topic\\_subtopic.php?tid=15&sid=22](http://geneed.nlm.nih.gov/topic_subtopic.php?tid=15&sid=22)

- 91 D. L. Nelson and M. m. COx, *Principles of Biochemistry*, W. H. Freeman and Company, sixth eddi., 2013.
- 92 D. Voet and J. G. Voet, *biochemistry*, wiley, third., 2004.
- 93 M. A. Nieto, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 155–166.
- 94 B. Gonzalez, L. J. Schwimmer, R. P. Fuller, Y. Ye, L. Asawapornmongkol and C. F. Barbas, *Nat. Protoc.*, 2010, **5**, 791–810.
- 95 M. Garriga-Canut, C. Agustin-Pavon, F. Herrmann, A. Sanchez, M. Dierssen, C. Fillat and M. Isalan, *Proc. Natl. Acad. Sci. USA*, 2012.
- 96 J. C. Miller, M. C. Holmes, J. Wang, D. Y. Guschin, Y.-L. Lee, I. Rupniewski, C. M. Beausejour, A. J. Waite, N. S. Wang, K. A. Kim, P. D. Gregory, C. O. Pabo and E. J. Rebar, *Nat. Biotechnol.*, 2007, **25**, 778–785.
- 97 D. A. Wah, J. Bitinaite, I. Schildkraut and A. K. Aggarwal, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 10564–10569.
- 98 T. Gaj, J. Guo, Y. Kato, S. J. Sirk and C. F. Barbas, *Nat. Methods*, 2012, **9**, 805–807.
- 99 J. D. Sander, E. J. Dahlborg, M. J. Goodwin, L. Cade, F. Zhang, D. Cifuentes, S. J. Curtin, J. S. Blackburn, S. Thibodeau-Beganny, Y. Qi, C. J. Pierick, E. Hoffman, M. L. Maeder, C. Khayter, D. Reyon, D. Dobbs, D. M. Langenau, R. M. Stupar, A. J. Giraldez, D. F. Voytas, R. T. Peterson, J.-R. J. Yeh and J. K. Joung, *Nat. Methods*, 2011, **8**, 67–69.
- 100 P. Savagner, K. M. Yamada and J. P. Thiery, *J. Cell Biol.*, 1997, **137**, 1403–1419.
- 101 A. Cano, M. A. Pérez-Moreno, I. Rodrigo, A. Locascio, M. J. Blanco, M. G. del Barrio, F. Portillo and M. A. Nieto, *Nat. Cell Biol.*, 2000, **2**, 76–83.
- 102 J. Y. Shih and P. C. Yang, *Carcinogenesis*, 2011, **32**, 1299–1304.
- 103 A. A. Beharry and G. A. Woolley, *Chem Soc Rev*, 2011, **40**, 4422–4437.
- 104 Y. Yang, R. P. Hughes and I. Aprahamian, *J. Am. Chem. Soc.*, 2012, 2–5.
- 105 S. M. Kelly, T. J. Jess and N. C. Price, *Biochim. Biophys. Acta*, 2005, **1751**, 119–39.
- 106 M. Irie and M. Kato, *J. Am. Chem. Soc.*, 1985, **107**, 1024–1028.
- 107 Y. E. Dokunikhin, N. S.; Gerasimenko, *Zhurnal Obs. Khimii*, 1959, **30**, 635–638.
- 108 B. N. Kholodenko, *J. Exp. Biol.*, 2003, **206**, 2073–2082.
- 109 C. Mügge, C. Rothenburger, A. Beyer, H. Görls, C. Gabbiani, A. Casini, E. Michelucci, I. Landini, S. Nobili, E. Mini, L. Messori and W. Weigand, *Dalton Trans.*, 2011, **40**, 2006–16.
- 110 T. S. Chou, S. J. Lee, M. L. Peng, D. J. Sun and S. S. P. Chou, *J. Org. Chem.*, 1988, **53**, 3027–3031.
- 111 H. Meckler and R. J. Herr, *Org. Process Res. Dev.*, 2012, **16**, 550–555.
- 112 Y.-F. Huang, H.-P. Zhu, G.-K. Liu, D.-Y. Wu, B. Ren and Z.-Q. Tian, *J. Am. Chem. Soc.*, 2010, **132**, 9244–6.
- 113 B. S. Morgan, S. M. Roberts and P. Evans, *Tetrahedron Lett.*, 2006, **47**, 5273–

- 5276.
- 114 J. Niu, P. Guo, J. Kang, Z. Li, J. Xu and S. Hu, *J. Org. Chem.*, 2009, **74**, 5075–8.
- 115 W. J. Bailey and E. W. Cummins, *J. Am. Chem. Soc.*, 1954, **76**, 1932–1936.
- 116 A. M. Caamaño, M. E. Vázquez, J. Martínez-Costas, L. Castedo and J. L. Mascareñas, *Angew. Chem. Int Ed*, 2000, **39**, 3104–3107.
- 118 T. X. T. Luu, P. Christensen and F. D. and T. N. Le, *J. Chem.*, 2007, **45**, 105–110.
- 119 H.-Q. Li, X. Liu, Q. Zhang, S.-S. Li, Y.-M. Liu, H.-Y. He and Y. Cao, *Chem. Commun.*, 2015, **51**, 11217–20.
- 120 C. Zhang and N. Jiao, *Angew. Chemie - Int. Ed.*, 2010, **49**, 6174–6177.
- 121 K. Tanaka, W. Ohashi, H. Okada and Y. Chujo, *Tetrahedron Lett.*, 2014, **55**, 1635–1639.
- 122 H. Xie, D. Yuan and M.-W. Ding, *J. Org. Chem.*, 2012, **77**, 2954–2958.
- 123 K. Maie, M. Nakamura and K. Yamana, *Nucleosides. Nucleotides Nucleic Acids*, 2006, **25**, 453–62.
- 124 V. Sai Sudhir, R. B. Nasir Baig and S. Chandrasekaran, *European J. Org. Chem.*, 2008, 2423–2429.
- 125 T. Fricke, R. J. Mart, C. L. Watkins, A. T. Jones and R. K. Allemann, *Drug Discov. Today*, 2010, **15**, 1088–1088.
- 126 D. B. G. Williams and M. Lawton, *J. Org. Chem.*, 2010, **75**, 8351–8354.

## 9 List of Figures

<b>Figure 1.</b> Estradiol protection and photoactivation.....	4
<b>Figure 2.</b> Phototaxel and photo deprotection to paclitaxel. ....	5
<b>Figure 3.</b> Radical photo cleavage of <i>ortho</i> -nitrobenzyl ether protecting groups. ....	6
<b>Figure 4.</b> Photocycle of flavin mononucleotide (FMN) in LOV domains.....	7
<b>Figure 5.</b> Spiroyrans are photochromic molecules that change structure when exposed to different wavelengths of light. (R is poly-L-glutamic acid). <sup>27</sup> .....	8
<b>Figure 6.</b> Conformational change brought about in poly-L-glutamic acid by the different states of spiropyran (Adapted from reference 27).....	8
<b>Figure 7.</b> UV (left) and CD (right) absorbance spectra for poly-L-glutamic acid with spiropyran attached showing the light (1) and dark states (2) with dashed lines showing intermediate relaxation spectra. <sup>27</sup> .....	9
<b>Figure 8.</b> Spiropyran attached a peptide at two points to control peptide $\alpha$ -helical conformation (adapted from reference 29). ....	9
<b>Figure 9.</b> CD spectra of native peptide (black) and crosslinked peptide in light (red) and dark (blue) states in phosphate buffer (100 mM, pH 6.6) at 25 °C. <sup>29</sup> .....	10
<b>Figure 10.</b> Interaction of spiropyran <b>14</b> and histidine in solution. <sup>28</sup> .....	10
<b>Figure 11.</b> Diarylethenes can cyclise to hexadienes. <sup>30</sup> .....	10
<b>Figure 12.</b> Absorbance spectra of <b>15</b> in diarylethene (open, —) and hexatriene (closed, ----) states and the photo stationary state after irradiation at 491 nm (- · - · -) in benzene. <sup>30</sup> .....	11
<b>Figure 13.</b> Thioindigo undergoes photoinduced <i>cis/trans</i> isomerisation. ....	11
<b>Figure 14.</b> Indigo carmine, a water soluble indigo.....	12
<b>Figure 15.</b> Thioindigo molecular tweezers. ....	13
<b>Figure 16.</b> Azobenzene in the <i>trans</i> conformation (left) and the <i>cis</i> conformation (right). ....	13
<b>Figure 17.</b> Bond lengths of azobenzene from crystal structures. ....	14
<b>Figure 18.</b> Substitution of the <i>para</i> position effects the thermal relaxation of azobenzene.....	15
<b>Figure 19.</b> <i>Ortho</i> amino substitution.....	17
<b>Figure 20.</b> Compound <b>31</b> <i>N,N'</i> -(diazene-1,2-diylbis(3-methoxy-4,1-phenylene))diacetamide and compound <b>32</b> <i>N,N'</i> -(diazene-1,2-diylbis(3,5-dimethoxy-4,1-phenylene))diacetamide. <sup>57</sup> .....	18
<b>Figure 21.</b> 1,1'-((Diazene-1,2-diylbis(3,5-dimethoxy-4,1-phenylene))bis(piperazine-4,1-diyl))bis(2-chloroethan-1-one) Azonium ion stabilization. ....	19
<b>Figure 22.</b> pH dependent UV/visible absorbance spectra of compound <b>33</b> . <sup>58</sup> .....	20
<b>Figure 23.</b> <i>N,N'</i> -(diazene-1,2-diylbis(3,5-dithioether-4,1-phenylene))diacetamide. <sup>59</sup> ..	20
<b>Figure 24.</b> Tetra- <i>ortho</i> -fluoroazobenzene derivatives. <sup>61</sup> .....	21
<b>Figure 25.</b> UV/visible spectra of <i>cis</i> - <b>36</b> (solid red) and <i>trans</i> - <b>36</b> (solid blue) photostationary states and calculated spectra for pure states (dashed red and blue, modified from reference 61). ....	22
<b>Figure 26.</b> Tetra- <i>ortho</i> -halo substituted azobenzenes. ....	22

<b>Figure 27.</b> Non-symmetrical difluoroazobenzene <b>50</b> and symmetric difluoroazobenzene <b>44</b> . <sup>60</sup> .....	23
<b>Figure 28.</b> 2,6-Difluoro and tetra-2,2',6,6'-fluoroazobenzenes. ....	23
<b>Figure 29.</b> 1,2-Bis-(2-(trifluoromethyl)phenyl)diazene. ....	23
<b>Figure 30.</b> Azobenzene <i>para</i> substitution. <sup>64</sup> .....	24
<b>Figure 31.</b> The nature of <i>para</i> substituents can have strong effects on the half-lives of <i>cis</i> isomers. ....	25
<b>Figure 32.</b> Installation of a thioether in the <i>para</i> position of <b>51</b> results in <b>52</b> exhibiting very different optical properties. <sup>66</sup> .....	25
<b>Figure 33.</b> UV/visible spectra for <b>52</b> bound to DNA in buffer (10 mM phosphate pH 7.0, 100 mM sodium chloride) for the <i>trans</i> (solid) and irradiated <i>cis</i> rich states (dashed). <sup>66</sup> .....	26
<b>Figure 34.</b> Azobenzene <b>53</b> has pH dependent UV/visible absorption spectra. <sup>67</sup> .....	26
<b>Figure 35.</b> Glutathione( <b>64</b> ), an intracellular redox buffer, reducing compound <b>54</b> . <sup>68</sup> .....	27
<b>Figure 36.</b> Colchicine and photoswitchable colchicine mimic in the active and inactive state. <sup>69</sup> .....	28
<b>Figure 37.</b> Photo responsive tandem zinc fingers. <sup>71</sup> .....	28
<b>Figure 38.</b> Azobenzene amino acid <b>60</b> used for incorporation into the homodimer interface of <i>Bam</i> HI and peptide backbones.....	29
<b>Figure 39.</b> Photo switching of(S)-2-amino-3-(4-(phenyldiazenyl)phenyl)propanoic acid ( <b>61</b> ). <sup>75</sup> .....	30
<b>Figure 40.</b> Azobenzene amino acid <b>63</b> contains a pendant vinyl group for forming a second point of attachment to a peptide <i>via</i> intramolecular thiol-ene click chemistry. <sup>76</sup> .....	31
<b>Figure 41.</b> Conformational control of protein structure by azobenzene. <sup>74</sup> .....	31
<b>Figure 42.</b> Protein $\alpha$ -helix showing hydrogen bonds (left) helical structure (right). .	32
<b>Figure 43.</b> Azobenzene attached to <i>i,i+7</i> spaced residues. ....	33
<b>Figure 44.</b> Azobenzene attached to <i>i,i+11</i> spaced residues. ....	34
<b>Figure 45.</b> 4-(Bromomethyl)azobenzene reacting with thiols. <sup>80</sup> .....	34
<b>Figure 46.</b> An azobenzene bearing an iodoacetamide anchoring group. ....	35
<b>Figure 47.</b> Water soluble, cysteine reactive 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido)azobenzene. <sup>4</sup> .....	36
<b>Figure 48.</b> bZIP Proteins binding to DNA. <sup>85</sup> .....	37
<b>Figure 49.</b> D-Theroninol azobenzene derivative. <sup>66</sup> .....	38
<b>Figure 50.</b> Graphical representation of DNA transcription (modified from Ref. 90) A) DNA strands separate locally B) A transcription factor binds its cognate DNA sequence C) RNA polymerase binds to the transcription factor D) RNA polymerase begins to assemble an RNA strand based on the DNA template.....	40
<b>Figure 51.</b> DNA $\alpha$ -helix. ....	43
<b>Figure 52.</b> A single Cys <sub>2</sub> His <sub>2</sub> zinc finger is shown in blue with the sidechains of two cysteine and two histidine residues coordinating a zinc ion (green). ....	44
<b>Figure 53.</b> Zinc finger based <i>Fok</i> I dimer binding. <sup>96</sup> .....	45
<b>Figure 54.</b> Photo-decaging of a tyrosine residue. ....	47

<b>Figure 55.</b> Sequence of the product of the <i>SNAI2</i> gene in <i>Homo sapiens</i> (UniProtKB: O43623). The chosen DNA-binding helices are highlighted in red and blue. ....	47
<b>Figure 56.</b> Cartoon of <i>SNAI2</i> zinc finger residues after homology modelling their sequence onto the structure of ZIF268 (PDB 1AAY). Full schematic (left) $\alpha$ -helical regions only (right). ....	48
<b>Figure 57.</b> Double slug with photoswitch. ....	48
<b>Figure 58.</b> Elementary steps of solid phase peptide synthesis. ....	49
<b>Figure 59.</b> Cleaving of peptide from a solid support with the use of TFA (where R is the peptide). ....	50
<b>Figure 60.</b> 3'-Bis(sulfonato)-4,4'-bis(chloroacetamido)azobenzene ( <b>71</b> ) was chosen for its water solubility. ....	51
<b>Figure 61.</b> Scheme showing the synthesis of BSBCA ( <b>71</b> ). <sup>104</sup> ....	51
<b>Figure 62.</b> Mechanism for oxidative coupling of anilines to form azobenzene. ....	52
<b>Figure 63.</b> UV absorption of <b>71</b> in water in its photostationary light state (blue) and dark state (red). ....	52
<b>Figure 64.</b> Arrhenius plot for <b>71</b> in water. ....	53
<b>Figure 65.</b> Iminium ion resonance structure of azobenzene stabilised by interaction with water. ....	54
<b>Figure 66.</b> Acetamidoazobenzene crosslinker <b>49</b> . ....	55
<b>Figure 67.</b> UV absorption of <b>49</b> in DMSO in its dark state (blue) and photostationary light state (red). ....	55
<b>Figure 68.</b> Azobenzene. ....	56
<b>Figure 69.</b> UV absorption of azobenzene in DMSO in its dark state (blue) and photostationary light state (red). ....	56
<b>Figure 70.</b> Absorbance spectra of <b>27</b> (green), <b>49</b> (purple) and <b>71</b> (orange) in DMSO. ....	57
<b>Figure 71.</b> HPLC chromatograms showing absorbance at 210 nm for native peptides (green) and crosslinked peptides (orange). ....	58
<b>Figure 72.</b> UV/visible absorption spectrum of Slug Finger 1 cross-linked with <b>71</b> the dark state (blue) and the 360 nm photostationary (red) state. ....	59
<b>Figure 73.</b> UV/visible absorption of Slug Finger 2- <b>71</b> in dark (blue) and 360 nm photostationary state (red). ....	60
<b>Figure 74.</b> UV/visible absorption spectra of the Slug Double Finger- <b>71</b> - <b>71</b> in dark (blue) and 360 nm photostationary state (red). ....	61
<b>Figure 75.</b> Relaxation curve Slug Double Finger- <b>71</b> , <b>71</b> at 40 °C overlaid with a first order kinetics fit. ....	62
<b>Figure 76.</b> Schematic representation of independent Slug Double Finger- <b>71</b> , <b>71</b> relaxation. ....	62
<b>Figure 77.</b> Schematic representation of non-independent Slug Double finger- <b>71</b> , <b>71</b> relaxation. ....	63
<b>Figure 78.</b> CD spectra of different protein secondary structure motifs: $\alpha$ -helix (solid line), anti-parallel $\beta$ -sheet (long dashed line), unstructured peptide (dotted line). Modified from reference <sup>105</sup> . ....	65
<b>Figure 79.</b> CD spectra of Slug Finger 1- <b>71</b> the dark state (blue) and the 360 nm photostationary state (red). ....	66



<b>Figure 80.</b> Circular dichromism spectra of Slug Finger 2-71 in the dark state (blue) and the 360 nm photo stationary state (red).....	67
<b>Figure 81.</b> Circular dichroism Slug Double Finger-71,71 in the dark state (blue) and the 360 nm photo stationary (red).....	67
<b>Figure 82.</b> Fluorescein attached to DNA.....	68
<b>Figure 83</b> Diagram of fluorescent anisotropy apparatus optical set up.....	69
<b>Figure 84.</b> Fluorescence anisotropy binding data for Slug Finger 1-71 in both light (blue squares) and dark adapted (red squares) states. Curves represent the best fit for the data see section 6.17.....	70
<b>Figure 85.</b> Fluorescence anisotropy binding data for Slug Finger 2-71 in both light (blue squares) and dark adapted (red squares) states.....	70
<b>Figure 86.</b> Fluorescence anisotropy binding data for Slug Double Finger-71,71 in light (blue squares) and dark adapted (red squares) states.....	71
<b>Figure 87.</b> Thioindigo modified to act as a isomerisation dependant ligand for small cations. <sup>106</sup> .....	74
<b>Figure 88.</b> Oxidation of indoxyl to indigo.....	75
<b>Figure 89.</b> 5,5`-Dinitro thioindigo (98).....	75
<b>Figure 90.</b> 5,5`-Dinitrothioindigo (98) UV/visible spectra. ....	76
<b>Figure 91.</b> Idealised thioindigo photoswitch.....	76
<b>Figure 92.</b> Retrosynthesis of sulfonamidothioindigo. ....	77
<b>Figure 93.</b> Chlorination of sulfate group. ....	78
<b>Figure 94.</b> Sulphonamide formation. ....	78
<b>Figure 95.</b> Formation of sulphonamide and an unexpected byproduct from morpholine.....	79
<b>Figure 96.</b> Selective sulfonamide formation with morpholine. ....	79
<b>Figure 97.</b> Nucleophilic aromatic substitution with methylthioglycolate.....	80
<b>Figure 98.</b> Thioindoxyl ring formation using a Friedel-Crafts reaction. ....	80
<b>Figure 99.</b> Tetra- <i>ortho</i> -substituted azobenzene.....	82
<b>Figure 100.</b> Trifluoromethyl-containing azobenzene photoswitches.....	82
<b>Figure 101.</b> Retrosynthetic analysis of <i>ortho</i> -trifluoromethyl azobenzenes. ....	83
<b>Figure 102.</b> Synthetic route to <i>ortho</i> -trifluoromethyl azobenzenes. ....	83
<b>Figure 103.</b> Synthetic route to <i>ortho</i> -trifluoromethyl azobenzenes. ....	84
<b>Figure 104.</b> <sup>1</sup> H NMR (300 MHz, CDCl <sub>3</sub> ) spectrum of 4-bromo-2-trifluoromethyl aniline (129).....	85
<b>Figure 105.</b> <sup>1</sup> H NMR (300 MHz, CDCl <sub>3</sub> ) spectrum for 1,3-dibromo-6-trifluoromethylaniline from 7.4 -7.8 ppm (133).....	85
<b>Figure 106.</b> UV/visible spectra of halogenated trifluoromethyl azobenzenes. Compound 121 (blue), Compound 124 (red), Compound 122 (orange), Compound 125 (green), Compound 126 (green) and Compound 123 (light blue),.....	86
<b>Figure 107.</b> UV/visible absorbance spectra of 124 (left) and 121 (right) in their <i>trans</i> (blue) and <i>cis</i> (red) states. ....	88
<b>Figure 108.</b> UV/visible absorbance spectra of 121 in the dark <i>trans</i> state (blue) and light <i>cis</i> state (red). ....	89



<b>Figure 109.</b> UV/visible absorbance spectra of <b>122</b> , <b>123</b> , <b>125</b> and <b>126</b> in the dark <i>trans</i> state (blue) and light <i>cis</i> state (red). .....	90
<b>Figure 110.</b> Synthesis of tetra- <i>ortho</i> -fluoroazobenzene. ....	91
<b>Figure 111.</b> UV/visible absorbance spectra of <i>trans</i> - <b>137</b> (blue) and <i>cis</i> - <b>137</b> (red). .....	91
<b>Figure 112.</b> Insertion of $\beta$ -mercaptoethanol at the <i>para</i> position of <b>137</b> . ....	92
<b>Figure 113.</b> UV/visible absorbance spectra of the dark <i>trans</i> state (blue) and photostationary light state (red) of <b>138</b> . ....	93
<b>Figure 114.</b> Protecting 2-mercaptoethanol with chloro- <i>tert</i> -butyldimethylsilane. ....	94
<b>Figure 115.</b> Reaction of <b>137</b> with <i>tert</i> -butyldimethylsilyl protected 2-mercaptoethanol. ....	94
<b>Figure 116.</b> An dihydrothiophene sulfoxide thioether. ....	96
<b>Figure 117.</b> Synthesis of a model unsaturated sulfoxide linking group. ....	96
<b>Figure 118.</b> Dehydrobromination of compound <b>144</b> followed by double bond migration to yield <b>145</b> . ....	97
<b>Figure 119.</b> Debromination of <b>144</b> resulting in compound <b>142</b> with the double bond in the desired position. ....	97
<b>Figure 120.</b> Bromination of compound <b>145</b> . ....	98
<b>Figure 121.</b> Attempted insertion of a <i>para</i> sulfur with sodium hydrogen sulfide. ....	98
<b>Figure 122.</b> Attempted reaction of sodium sulphide with <i>para</i> -bromo aniline. ....	99
<b>Figure 123.</b> Grignard reagent formation for sulfur insertion. ....	99
<b>Figure 124.</b> Elimination of hydrogen bromide from <b>152</b> to form compound <b>153</b> . ....	100
<b>Figure 125.</b> Reaction of <b>153</b> with a test thiol to form compound <b>154</b> . ....	100
<b>Figure 126.</b> Ullmann ether synthesis. ....	101
<b>Figure 127.</b> Ullmann ether synthesis. ....	101
<b>Figure 128.</b> Benzylamine reaction with compound <b>157</b> . ....	102
<b>Figure 129.</b> Synthesis of compound <b>157</b> by dibromination and elimination. ....	102
<b>Figure 130.</b> Reaction of <b>157</b> with cysteine. ....	102
<b>Figure 131.</b> Synthetic pathway toward compound <b>165</b> . ....	103
<b>Figure 132.</b> Reduction of compound <b>163</b> . ....	104
<b>Figure 133.</b> Synthesis of Compound <b>171</b> tetra- <i>ortho</i> -chloro benzyl alcohol. ....	106
<b>Figure 134.</b> Oxidative dimerisation through azo bond formation. ....	107
<b>Figure 135.</b> Mechanism for base-catalysed hydrolysis of the nitrile. ....	108
<b>Figure 136.</b> Reduction of ethyl ester. <sup>121</sup> ....	109
<b>Figure 137.</b> Esterification of <b>170</b> to chloroacetate ester <b>171</b> . ....	109
<b>Figure 138.</b> UV/visible absorbance spectra of <i>cis</i> - <b>171</b> (blue) and <i>trans</i> - <b>171</b> (red). ....	109
<b>Figure 139.</b> Relaxation curve showing absorbance at 300 nm for compound <b>171</b> in DMSO at 60 °C. ....	110
<b>Figure 140.</b> ESI (left) and MALDI (right) mass spectra for the crude crosslinked FAM-Bid <sup>i,i+4</sup> - <b>171</b> showing molecular ions (calculated monoisotopic mass 3191) and sodium adducts. ....	111
<b>Figure 141.</b> UV/visible spectra of crude FAM-Bid <sup>i,i+4</sup> - <b>171</b> in the dark adapted states (red) and after irradiation with >530 nm light (blue). ....	112
<b>Figure 142.</b> Relaxation curve of absorbance at 300 nm for crude FAM-Bid <sup>i,i+4</sup> - <b>171</b> in DMSO at 60 °C. ....	112

<b>Figure 143.</b> CD spectra of crude FAM-Bid <sup>i+4</sup> - <b>171</b> in the dark adapted states (red) and after irradiation with > 530 nm light (blue).....	113
<b>Figure 144.</b> Modified synthesis of the tetra- <i>ortho</i> -chloro substituted azobenzene <b>171</b> . .....	114
<b>Figure 145.</b> Synthetic scheme for tetra- <i>ortho</i> -bromo substituted azobenzene <b>179</b> . ....	114
<b>Figure 146.</b> Synthetic scheme for tetra- <i>ortho</i> -fluoro substituted azobenzene <b>186</b> . ....	115
<b>Figure 147.</b> Synthetic scheme for tetra- <i>ortho</i> -hydroazobenzene <b>191</b> . ....	116
<b>Figure 148.</b> HPLC-DAD UV/visible spectra of <i>cis</i> - <b>191</b> (blue) and <i>trans</i> - <b>191</b> (red). ....	116
<b>Figure 149.</b> HPLC-DAD UV/visible spectra of <i>cis</i> - <b>179</b> (blue) and <i>trans</i> - <b>179</b> (red).....	117
<b>Figure 150.</b> HPLC-DAD UV/visible spectra of <i>cis</i> - <b>185</b> (blue) and <i>trans</i> - <b>185</b> (red).....	118
<b>Figure 151.</b> Crystal structures of <b>191</b> , <b>185</b> , <b>171</b> and <b>179</b> showing the dihedral angle of the azo bonds. ....	118
<b>Figure 152.</b> Crystal structures of <b>185</b> , <b>171</b> and <b>179</b> showing the angles of the aromatic rings. ....	119
<b>Figure 153.</b> Space-filling model of the compounds <b>185</b> , <b>171</b> , and <b>179</b> illustrating the increased steric demands of tetra- <i>ortho</i> -chloro- and tetra- <i>ortho</i> -bromo-substituents. .....	120
<b>Figure 154.</b> Synthetic scheme for di- <i>ortho</i> -haloazobenzenes <b>207</b> , <b>208</b> , and <b>209</b> . ....	121
<b>Figure 155.</b> Monobromination of 4-aminobenzonitrile <b>186</b> . ....	122
<b>Figure 156.</b> HPLC-DAD UV/visible spectra of <i>cis</i> - <b>209</b> (blue) and <i>trans</i> - <b>209</b> (red). ....	123
<b>Figure 157.</b> HPLC-DAD UV/visible spectra of <i>cis</i> - <b>208</b> (blue) and <i>trans</i> - <b>208</b> (red). ....	123
<b>Figure 158.</b> HPLC-DAD UV/visible spectra of <i>cis</i> - <b>207</b> (blue) and <i>trans</i> - <b>207</b> (red).....	124
<b>Figure 159.</b> UV/visible spectra of <i>trans</i> - <b>191</b> (unsubstituted, grey) <i>trans</i> - <b>209</b> (dibromo, red) and <i>trans</i> - <b>179</b> (tetrabromo, blue). ....	125
<b>Figure 160.</b> UV/visible spectra of <i>cis</i> - <b>191</b> (unsubstituted, grey) <i>cis</i> - <b>209</b> (dibromo, red) and <i>cis</i> - <b>179</b> (tetrabromo, blue). ....	125
<b>Figure 161.</b> Synthesis of dichlorodifluoroazobenzene <b>215</b> . ....	127
<b>Figure 162.</b> Synthesis of dibromodifluoroazobenzene <b>221</b> . ....	128
<b>Figure 163.</b> UV/visible absorbance spectra of <i>cis</i> - <b>215</b> (blue) and <i>trans</i> - <b>215</b> (red).....	129
<b>Figure 164.</b> UV/visible absorbance spectra of <i>cis</i> - <b>221</b> (blue) and <i>trans</i> - <b>221</b> (red).....	129
<b>Figure 165.</b> Comparison of photo switching of dichloro,difluoro <b>215</b> (left) and tetrachloro <b>171</b> (right) photoswitches after irradiation with light through a >530 nm filter for 0 (blue), 5 (purple), 10 (red) and 15 (green) minutes. ....	130
<b>Figure 166.</b> Crystal structure of <i>trans</i> - <b>215</b> showing the coplanarity of the aromatic rings. ....	131
<b>Figure 167.</b> Attempted crosslinking of compound <b>209</b> with cysteine. ....	132
<b>Figure 168.</b> Screening of crosslinking conditions reacting overnight at 40 °C. benzyl alcohol (blue), THF (yellow), methanol (grey), acetonitrile (orange), DMSO (light blue).....	133
<b>Figure 169.</b> Screening of crosslinking conditions reacting overnight at 60 °C. benzyl alcohol (blue), THF (yellow), methanol (grey), acetonitrile (orange), DMSO (light blue).....	133

---

<b>Figure 170.</b> HPLC traces of the degradation of compound <b>206</b> to benzyl alcohol compound <b>206</b> : Initial (light blue), 1 hour (orange), 2 hours (grey), overnight (yellow), and benzyl alcohol (blue). .....	134
<b>Figure 171.</b> Bromination of benzyl alcohol. ....	134
<b>Figure 172.</b> Nucleophilic aromatic substitution of benzyl bromide with sodium azide. ....	134
<b>Figure 173.</b> DIBAL reduction of benzyl azide to benzylamine. ....	134
<b>Figure 174.</b> Imminophosphrane formation from benzyl azide. ....	135
<b>Figure 175.</b> Chloroacetylation of compound <b>227</b> . ....	135
<b>Figure 176.</b> Bromination of to form hexabromide <b>229</b> . <sup>123</sup> .....	136
<b>Figure 177.</b> Chlorination of <b>178</b> to tetrabromodichloride <b>230</b> . ....	136
<b>Figure 178.</b> Formation of benzylic azide <b>229</b> from hexabromide <b>231</b> . ....	137
<b>Figure 179.</b> Reduction of crude azide displacement product postulated to contain <b>231</b> to imminophosphorane <b>232</b> . ....	137
<b>Figure 180.</b> Formation of compound <b>233</b> . ....	138
<b>Figure 181.</b> Formation of compound <b>234</b> . ....	138
<b>Figure 182.</b> Formation of chloroacetamido azobenzene <b>235</b> . ....	139
<b>Figure 183.</b> Possible azobenzene products from introducing functionalised primary amines. ....	139
<b>Figure 184.</b> UV traces of compound of <i>trans</i> (blue) and <i>cis</i> (red) <b>235</b> . ....	140
<b>Figure 185.</b> Test of crosslinking reaction conditions in the presence of cysteine. ....	141
<b>Figure 186.</b> HPLC UV traces of crosslinked cysteine (red) in both the <i>cis</i> and <i>trans</i> states (peaks A and B) with a small amount of single cysteine attached (peak C) and compound <b>235</b> (blue) for comparison. ....	141
<b>Figure 187.</b> Relaxation curve example. Raw data (blue) and the calculated data based on optimised variables (red) .....	148
<b>Figure 188.</b> Example binding plot from fluorescence anisotropy experiment. ....	151

## 10 List of Tables

<b>Table 1.</b> Table of the effects of <i>ortho</i> amino substitution on azobenzene. <sup>54</sup> .....	18
<b>Table 2.</b> Effect of <i>para</i> heteroatom substituents on the stability of the <i>cis</i> isomer of azobenzenes. <sup>64</sup> .....	24
<b>Table 3.</b> Half lives of <i>cis</i> isomers of <b>52</b> and <b>51</b> bound to DNA (5 $\mu$ m in 10 mM phosphate, pH 7, 100 mM sodium chloride). <sup>66</sup> .....	26
<b>Table 4.</b> Half-lives of the <i>cis</i> isomers of Fk-II- <b>71</b> and Fk-II- <b>70</b> .....	36
<b>Table 5.</b> Binding affinities of Bak-7- <b>74</b> and Bak-II- <b>74</b> binding to Bcl-X <sub>L</sub> . .....	37
<b>Table 6.</b> Sequences of Slug zinc fingers (Linker in red, original residues in blue, modified in green). .....	49
<b>Table 7.</b> Expected and observed masses for Slug peptides. ....	50
<b>Table 8.</b> Half-lives of <i>cis</i> - <b>71</b> in water.....	53
<b>Table 9.</b> <i>cis</i> - <b>71</b> relaxation times in DMSO. ....	54
<b>Table 10.</b> Azobenzene relaxation times in DMSO at varying temperature. ....	55
<b>Table 11.</b> Relaxation times of <i>cis</i> - <b>27</b> in DMSO at different temperatures. ....	56
<b>Table 12.</b> Table of crosslinked Slug finger peptide masses. ....	58
<b>Table 13.</b> Relaxation times of Slug Finger 1- <b>71</b> .....	59
<b>Table 14.</b> Half-lives of Slug Finger-2- <b>71</b> . ....	60
<b>Table 15.</b> Relaxation times of Slug Double Finger- <b>71,71</b> .....	61
<b>Table 16.</b> Conditions used to attempt to close the thiooxazol ring. ....	80
<b>Table 17.</b> Absorbance maxima, switching extents and thermal reversion half-lives for <b>121-126</b> . ....	87
<b>Table 18.</b> Properties of <i>para</i> -sulfur substituted azobenzenes. ....	95
<b>Table 19</b> Conditions used to attempt nitrile reduction. ....	103
<b>Table 20</b> Sequence of FAM-Bid <sup>i+4</sup> .....	110
<b>Table 21.</b> Comparison of the properties of the tetra- <i>ortho</i> -haloazobenzenes. ....	119
<b>Table 22</b> Optical and physical properties of the di- <i>ortho</i> -halo azobenzenes. ....	124
<b>Table 23.</b> Comparison between non halogenated di bromide and tetra bromide substituted azobenzenes. ....	126
<b>Table 24.</b> Photoswitching of <b>215</b> . ....	130
<b>Table 25.</b> Spectroscopic and half-life data for the mixed halogenated azobenzenes..	131
<b>Table 26.</b> Half-lives of <i>cis</i> - <b>235</b> at different temperatures.....	140