

Exploring the untapped reactivity of nitrile reagents for site-selective protein modification

Mikesh Patel

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Supervisors: Prof. James Baker and Prof. Vijay Chudasama

November 2023

Declaration

I, Mikesh Patel, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Mikesh Patel

November 2023

Abstract

Novel and effective modes of protein modification are constantly in demand for the development of bioconjugates. Antibody-drug conjugates (ADCs) are a rapidly evolving family of bioconjugates and are based on the well documented monoclonal antibody. These systems provide an excellent model to develop new conjugation techniques due to their diverse and complex nature.

The focus of this work investigates the properties and use of the nitrile bis-thiol (NBT) reaction as a new bioconjugation tool. Chapter 2 describes the use of mono-nitriles as a new generation of cysteine-to-lysine (CLT) reagents. New reagent designs have been synthesised and tested on the Fab fragment of the monoclonal antibody; trastuzumab. The CLT product generated amidine conjugates which are stable under a variety of conditions. The designs vary in reactivity *via* the modulation of the nitrile electron deficiency, eventually leading to a final functional CLT reagent and therefore Fab fragment.

Chapter 3 describes the novel chemistry and use of a new, dynamic, bridging reagent in the form of bis-nitriles. These reagents are explored for their disulfide bridging properties and relevant chemistry. The reagents react through a novel chemical cascade resulting in a trapped cyclic product. Functional analogues are then synthesised and tested on the Fab fragment, and later full antibody systems. Alternate nitrile-based bridging reagents are also synthesised and tested to further develop the understanding behind the chemistry and use of such reagents. These bridging reagents were also applied on the protein disulfide isomerase; thioredoxin to realise the potential of these reagents as selective enzymatic inhibitors.

The final chapter builds upon the foundation of the NBT reaction with ethanedithiol as a new selective reagent with potential to be developed into a biorthogonal reaction. This can have wide-spread use in areas such as rapid tag systems for use in protein labelling, pull-down assays and potentially more.

Impact statement

Bioconjugation is an ongoing field of research, which falls into a variety of applications such as the development of diagnostics and therapeutics, for example; the development of highly bespoke, antibody drug conjugates (ADCs). ADCs are comprised of two major components, the first of which is a monoclonal antibody, which presents remarkable selectivity towards a unique target. The second component is a highly cytotoxic payload, which is typically too potent for lone use due to a lack of selectivity. The combination of monoclonal antibody selectivity and the toxicity of the payload results in a highly effective treatment in the form of ADCs. As of 2023 thirteen ADCs are approved for human treatment nine of which were approved in the past five years. This shows the growth around this area of research and provides a highly rewarding field to work within.

The development of ADCs faces difficulties which can result in issues such as premature release of the linker-payload, heterogeneous species with differing drug to antibody ratios (DARs) and variable pharmacokinetics. Two key modes of modification lie within targeting cysteine and lysine residues, each of which present distinct features. The development of cutting edge bioconjugation technologies targeting these residues is highly desirable which can then be applied to a selection of proteins, including but not limited to ADCs. A key part of developing new bioconjugation technologies is the accessibility, limiting the need for practices such as protein engineering, which can be expensive and time consuming.

Nitrile moieties have been prevalent throughout protein modification, this research looks to explore new and exciting chemistry for the application of nitrile chemistry in protein modification. A recent protocol developed was based around cysteine-to-lysine transfer (CLT) in which cysteine residues were utilised as ligating hooks, followed by transfer to proximal lysine residues. The reactive centres in question were alkyl thioesters; these 1st generation reagents required long reaction times and suffered from drawbacks such as hydrolytic instability thereby lowering the efficiency. Electron deficient heteroaromatic nitriles were identified as 2nd generation CLT reagents with tuneable reactivity, overcoming the shortcomings present in the thioester reagents.

This work also introduces a new class of one-carbon bridging reagents in the form of 1,2 bis-nitriles. These reagents present a new dithiol selective approach which reacts via a novel dynamic reaction cascade. The resulting conjugates present great stability under physiological conditions and can present a great tool in bioconjugation for widespread use. Applicability is demonstrated on multiple systems with uses varying from antibody modification to enzyme inhibition.

The core, novel nitrile-bis thiol (NBT) reaction behind these reagents was also further explored. The reaction exhibited exquisite selectivity for ethanedithiol and opened several possible avenues for future applications. Biorthogonality was investigated and early studies revealed highly promising results, set to be further developed for wide-scale use.

Overall, this work provides insight into the novel NBT reaction and extends into the development of new reagents. These reagents have been proven on protein systems, highlighting their effectiveness as modification reagents for a wide variety of chemical biology scenarios.

Acknowledgments

Who would have thought I'd be at UCL for 7 years, not me. The past 3 of these years have been filled with some of the best, worst and 'that will do' moments, and as cliché as it might be I wouldn't have it any other way. The PhD was character building to say the least however it has helped me develop personally and academically.

A key contributor to the progress of my PhD is the one and only Prof Jamie Baker aka JRB. Jamie used to lecture me in 3rd year chemistry and two standouts from this were his kind nature and sub-par handwriting (sorry Jamie). He is an incredible supervisor knowing what to say and when to say it, never being too absent or overbearing. Jamies love for science is unmatched, the mumbled excited swearing and immediate apology for swearing never fails to put a smile on anyone's face. I'd like to thank him for the enthralling discussions we'd have regarding the pain nitriles cause and the fascination of what chat GPT is capable of. The biggest thanks, however, is for all the support he provided over the past few years. I know Jamie will be someone I keep in contact with, and I wish him all the luck in his football coaching and academic career.

Another large contributor to my past few years is Prof Vijay Chudasama, as with Jamie I first met Vijay in undergrad during a 2nd year lecture when he tried switching windows and accidently showed us his wedding background picture. The whole lecture theatre let out a simultaneous aww, this is the only time I've seen Vijay panic and click everything on his laptop to try move on. Since then I've known Vijay as a firm hand who tells it as it is. That being said he knows how to have a laugh and we've had great and sometimes random chats near the chem bio area. Vijay has always achieved great things and I wish this continues; I'll definitely keep in contact with Vijay, even if it is just to hear him rant about something.

Next, I'd like to thank the founder of my project and the person who set me up in the lab; Dr Nafsika Forte. I figured out pretty quickly how intelligent Nafsika is when she'd effortlessly carry out her work whilst giving me a crash course on 'how to chem bio'. The benchmark she set for me was inspiring and often when I'd get stuck, I would think 'WWND' (what would Nafsika do).

Archie, who has just as big a part in getting me started, was a key pillar in helping my science and integration into the Baker group. Our conversations developed from work to cricket to everything in between. Thank you, Archie, for giving me the start that I got as I feel like I'd definitely be a couple months behind without it.

When I think of my time in the KLB, one name is synonymous with it: Muhammed. He's been there from the start to the end, and we've seen the highs and lows throughout. Highs including

trips to simply chicken, Korean shop and the infamous Barcelona. Lows including the verge of committing fraud (which we definitely didn't) and planning experiments with questionable reagents. We are always on the same wavelength when it comes to music, food and other hobbies such as 3h online car shopping sprees. Our friendship has even blossomed into an organisation well known as Burger Boys©. We can't end the thanks to Muhammed without saying that finally…it actually is looking good brev.

Yanbo, my self-proclaimed lab sister and also a villain in disguise. When I started Yanbo would back me with a trademark 'beee niceeee' unknowingly to me. Despite this she and I never really spoke much…. Then one day it all changed. The pranks begun and she very quickly became a great outlet for when things weren't always going well. Let's not get it twisted, behind the support was always a hint of insult which was a common theme in our coffee trips each morning. All that was needed was a simple 'coffee?' message and next thing you know we'd be in line for that oat flat white and latte. Hopefully we can have a Bakers visit China trip soon and you can listen to anything other than 'we don't talk anymore' or 'attention' (PS singing is not for everyone).

Burger Boys© other founder is no other than the lab policeman and detective, Usman. Every morning you'd hear a very distinct 'well, well, well look who it is' and you know exactly who it is. Usman would always surprise with his ability to get gossip from anywhere and everywhere causally dropping it in conversation like it's the norm. He also had secrets of his own which to this day are unsolved, hopefully in the future we can solve these over more burgers and milkshakes.

This person is the human form of the French accent, its Lea Rocket. Lea started with me in the KLB 3 years ago and has always been a joy to distract from her work. This would be a hindrance to anyone else but with Lea's insane work ethic she needed someone to tell her to relax (obviously she wouldn't listen). Lea's catchphrases and struggles to remember the English words for things were and will always be amusing. For example some alternatives for set in stone…'on the rock/on the ground'. I hope to continue distracting you and know that you'll smash it wherever you go. Vasy mon garrr!

Alina was my first KLB fumehood buddy who in our first interaction terrified me after I failed to clean up on my first week. She was on the receiving end of Yanbo's 'beee niceeee' and over time the Alina I know now began to show up- still just as terrifying but twice as goofy. Over the years I've heard some bad jokes, but Alina always managed to push that bar lower…. and lower until she hit the bottom of the barrel with her trademark; 'high three'. Alina shared the same hatred of Wordle (stupid game) and would always support me with any issues I had,

(mainly chem bio… actually always chem bio). Let's look forward to more Faker meetups where the worst duo in flight club can hopefully improve.

Next is the Italian drill connoisseur, Lula. For someone who is so bubbly in person, her music choice was that of a 21-year-old roadman fresh out of prison. To the normal person that might be a bad thing but to Lula that is actually a compliment. We were paired to share a fumehood in my final year (as we were voted messiest) and it was indeed a great match. Finally, I'm glad you managed to submit before me and restore balance to the KLB timeline.

The KLB record for most nicknames belongs to the one and only Roshni…aka Roshanald, Roshnini, Roshini, RMfH, Ferrero Roshni. How can we not mention the awkward train journey after labs when a total of 3 things were said, look how far we've come. From vacuum sealing Hash to discussing Bollywood for hours its now the opposite where we can talk about the most pointless things for ages. I guess you have had a role in helping me come out of my shell at the start, so in part you are responsible for me not being able to shut up now. Loads more discussions are still pending and I'm looking forward to each and every one of them!

Standing at 9 foot 2 inches, with a vertical jump of 3 centimetres and a wingspan of a Boeing 747 its Phyll. I've known Phyll for my whole time at UCL and lets be real, it doesn't get any easier over time. I guess the friendship works because she also has the mental capacity of a 5-year-old who just learnt what the middle finger is. We became the first aiders of the lab together and ever since we've never had a serious injury…only because everyone knew the treatment by us would be worse than the pain from the injury. Phyll 'the high-performance athlete' will be someone I see often, like a doctor's appointment…except she is the illness.

The baby Baker title has been passed to the oldest 23-year-old I know, it's Charle aka Charlie. Not many people under the age of 40 say the phrase 'I don't trust google pay' but Charlie is one of them. He has fit so well into the group from day one, joining in our regular lunchtime conversations about how many toddlers would be needed to win against a bear. Keep up the spirit of the Baker group and remember when things 'don't chicken' a bossman will always help.

The caffeination couple, who make up about 65% of Pret's profits; Clíona and Ioanna. You two are regulars in our lab wide gameshow of 'What if?' and 'Do you reckon?' which usually spills into several hours of discussion. Clíona has a unique set of 'detective' skills which have come in handy more times than we'd like to admit, the only thing left for her to solve is how to lock her laptop when she leaves the room. Ioanna… well she runs the lab, but she earned that spot with an unreal drive to get things done (when I grow up, I want to be like Ioanna). I've never met someone who loves Greece like Ioanna, although I'm pretty sure the UK is her 2nd favourite island.

I could go on and on about all the amazing people in the KLB, so I will. Here are some more people I've had the pleasure of becoming friends with over the years. Nehaal, KLB's top model, you get a special thanks for trying so hard to get a conversation out of me and not giving up until I finally opened up. Steven I'll remember all those confused looks you give me when I come out with unusual sayings, and those little talks we'd have by the chem bio area about things never working. Big Dave, you were supportive not only in chemistry but also with the stresses of dealing with a PhD. I still refer to things as 'Dave scale'. The newest addition, Datt Magwell aka Cherry pie, you've extended the streak of well-matched Bakers. You and Charlie are going to do great things in the lab whilst also being the clowns of the KLB (wear this like a badge of honour). PS the back of your head was well textured. Matt Penny was a man of few words until he revealed a funny, sporty side to him which was always great to see. I appreciate you putting up with me and don't worry I know Phyll will carry on the immature behaviour which you know and love. Peter thanks for being such an informative guy who also knows how to have a laugh, you were always someone with great aspirations and I hope you achieve all you set out for. Fabien, those late nights in labs with you getting everyone hyped dancing to music will always be a great memory. Richard Spears always a laugh to be around with his holiday stories at lunch. Faiza, we had limited overlap in the lab but still managed to have great Bollywood discussions. Richard Procter, a lethal striker in football. Calise a charismatic lab mate. Toby who full of surprises. Sahra for being a solid lab DJ. Bianka who would always be concerned that me and Nehaal would 'turn to illegal sources of income'. Maybe one day you'll come around to the idea to a Red coffee (Red bull + coffee 50:50).

Outside of the PhD I've received support from friends and family. My two sisters Heena and Niralee have been there throughout alongside my parents; Joshna and Vinay. Sandra and Chloe, you two have also been there from when I used to be a little introverted kid back in year 7, thank you both. Neeku, we started by basically abusing one another in basketball, but you have become an incredible friend and have shown faith in me constantly. Finally the boys who I've made some of my best memories with and am glad to call close friends; The J man (Jesse), McLavan (Lavan) and DJ Spotify (Eddie).

Thanks to Dr Abil Aliev for providing the NMR assistance over the years and Dr Kersti Karu for the support in mass spectrometry.

A final shout out goes to The Weeknd, as the only artist all Bakers would happily listen to over 3 years without complaining.

Table of contents

Declaration ii
Abstract iii
Impact statement iv
Acknowledgments vi
Table of contents x
1. Introduction 1
1.1. Protein modification 1
1.2. Antibodies 2
1.2.1. Antibody fragments 4
1.3. Antibody drug conjugates 5
1.4. Payload 7
1.4.1. DNA-damaging drugs 7
1.4.1.1. Double strand break agents 7
1.4.2. DNA cross-linkers 8
1.4.3. Tubulin inhibitors 9
1.4.3.1. Topoisomerase I inhibitors 10
1.5. Linkers 11
1.5.1. Cleavable linkers 12
1.5.1.1. Acid cleavable linkers 12
1.5.1.2. Reducible linkers 13
1.5.1.3. Enzymatically cleavable linkers 14
1.5.2. Non-cleavable linkers 15
1.6. Bioconjugation strategies 16
1.6.1. Native antibody modification 17
1.6.1.1. Cysteine bioconjugation 17
1.6.2. Disulfide bridging reagents 21
1.6.3. Lysine bioconjugation 25
1.6.4. Emerging residues 31
1.7. Engineered antibody modification 32
1.8. Native chemical ligation 33
1.9. Cysteine to lysine transfer (CLT) 35
1.10. Nitriles 36
Aims & Objectives 41
Results and Discussion 42
2. Chapter 2: Nitrile CLT 42
2.2. Mechanistic studies 45

2.3. Electron-deficient nitriles 49
2.4. CLT bioconjugation studies 54
2.4.1. CLT of 5-(trifluoromethyl)pyrimidine-2-carbonitrile (6) 57
2.4.2. CLT of 5-(methylsulfonyl)pyrimidine-2-carbonitrile (8) 57
2.4.3. CLT of 5-nitropyrimidine-2-carbonitrile (7) 58
2.4.4. CLT of 4,6-dimethoxy-1,3,5-triazine-2-carbonitrile (9) 59
2.4.5. CLT of 5,6-dimethyl-1,2,4-triazine-3-carbonitrile (11) 61
2.4.6. CLT of 1,2,4,5-tetrazine-3-carbonitriles (12, 13 and 14) 62
2.4.7. CLT of *N*-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide (5) 66
2.5. CLT Stability studies 68
2.5.1. ELISA 68
2.5.1.1. CLT aqueous stability 69
2.6. Summary and conclusion of nitrile CLT 71
2.7. S_NAr for CLT 73
2.8. Summary and conclusion of S_NAr CLT 78
3. Chapter 3: Nitrile Disulfide Bridging Reagents 79
3.1. Bis Nitrile Bridging 79
3.1.1. Pyrazine-based Disulfide Bridging reagents 81
3.1.1.1. Chemical synthesis 81
3.1.1.2. Bioconjugation 86
3.1.2. Quinoxaline-based Disulfide Bridging reagents 95
3.1.2.1. Chemical synthesis 95
3.1.2.2. Bioconjugation 96
3.2. ELISA 100
3.3. Bis nitrile bridge stability 100
3.4. Selectivity reactions 102
3.5. Thioredoxin 104
3.6. Summary and conclusion of nitrile bridging 107
4. Chapter 4: Amino dithioacetals (ADTAs) 110
4.1. ADTA formation 110
4.2. ADTA stability 113
4.3. Summary and Further work 118
5. Conclusion 119
6. Experimental section 121
6.1. Chemical synthesis general remarks 121
6.2. Synthesis and characterisation of compounds 121
2-(Pyrimidin-2-yl)-1,3-dithiolan-2-amine (3) 121

N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide (5).. 122
5-(Trifluoromethyl)pyrimidine-2-carbonitrile (6)[248] 122
5-Nitropyrimidine-2-carbonitrile (7) 123
5-(Methylsulfonyl)pyrimidine-2-carbonitrile (8) 124
4,6-Dimethoxy-1,3,5-triazine-2-carbonitrile (9)..... 124
2-Chloro-4,6-diethyl-1,3,5-triazine (10) 125
5,6-Dimethyl-3-(methylsulfonyl)-1,2,4-triazine[215] 125
5,6-Dimethyl-1,2,4-triazine-3-carbonitrile (11)..... 126
6-Chloro-*N*-propyl-1,2,4,5-tetrazin-3-amine 126
6-(Propylamino)-1,2,4,5-tetrazine-3-carbonitrile (12) 127
N-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylacetamide (13)..... 127
N-(6-Cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylbenzamide (14)..... 128
3-(Methoxymethyl)-6-phenyl-1,2,4,5-tetrazine (15) 128
(6-Phenyl-1,2,4,5-tetrazin-3-yl)methanol (16) 129
Spiro[pyrrolo[3,4-b]pyrazine-5,2'-[1,3]dithiolan]-7(6H)-imine (25)..... 129
7,7-Bis(ethylthio)-6,7-dihydro-5H-pyrrolo[3,4-b]pyrazin-5-imine (26)..... 130
5,6-Bis(bromomethyl)pyrazine-2,3-dicarbonitrile (29)[249]..... 130
5,6-Bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile (30)..... 131
5-(4-hydroxyphenyl)pyrazine-2,3-dicarbonitrile (31)[250] 131
5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile (32) 132
Diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate (34)[231]..... 132
Pyrazine-2,3,5,6-tetracarbonitrile (35)[234]..... 133
Methyl 3-cyanopyrazine-2-carboxylate (36)[235] 133
Methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate (38)..... 134
2,3-Dicyanoquinoxaline-6-carboxylic acid (47)[251]..... 134
2,3-Dicyano-*N*-(prop-2-yn-1-yl)quinoxaline-6-carboxamide (49) 135
N-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide (48)[252]..... 135
ADTA (3) thiol stability studies..... 136
ADTA (3) incubation with ethanethiol 136
ADTA (3) aqueous stability 137
Ethanedithiol Vs. Propanedithiol NBT competition experiment 137
ADTA (3) reversibility- *N*-methyl maleimide 138
DFT general remarks and calculated co-ordinates (S5) 140
M062X/6-31+G(d,p) – 5-membered Acyclic optimised geometry..... 141
M062X/6-31+G(d,p) – 5-membered cyclic optimised geometry 142
M062X/6-31+G(d,p) – 6-membered Acyclic optimised geometry..... 143
M062X/6-31+G(d,p) – 6-membered Cyclic optimised geometry..... 144

Pyrimidine-CN 145
M062X/6-31+G(d,p) – Pyrimidine-CN 145
M062X/6-31+G(d,p) – Ethanedithiol 145
M062X/6-31+G(d,p) – Propanedithiol 146
6.3. Bioconjugation general remarks 147
LCMS general remarks 148
General procedure for the preparation of trastuzumab (HerceptinTM) Fab fragment.[253] 149
Native Fab 150
Reduced Fab 150
Trastuzumab 151
Reduced Trastuzumab 152
CLT reaction of Fab with 2-cyanopyrimidine. 154
Low equivalent reactions 154
High equivalent reactions 155
CLT reaction of Fab with 2-cyanopyrimidine (Reduce and cap) 158
Mechanistic study of nitrile CLT 160
Intramolecular selectivity competition experiment between Fab and *n*-butylamine 162
5-(trifluoromethyl)pyrimidine-2-carbonitrile 6 reactions on Fab 163
High equivalent reactions 163
Low equivalent reactions 165
5-Nitropyrimidine-2-carbonitrile 7 reactions on Fab. 166
Low equivalent reactions 166
High equivalent reactions 167
4,6-Dimethoxy-1,3,5-triazine-2-carbonitrile 10 reactions on Fab. 169
Low equivalent reactions 169
High equivalent reactions 171
5,6-Dimethyl-1,2,4-triazine-3-carbonitrile 11 reactions on Fab. 172
High equivalent reactions 172
Low equivalent reactions 174
6-(Propylamino)-1,2,4,5-tetrazine-3-carbonitrile 12 reactions on Fab. 176
Low equivalent reactions 176
High equivalent reactions 179
N-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylacetamide 13 reactions on Fab. 181
Low equivalent reactions 181
High equivalent reactions 182
N-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylbenzamide 14 reactions on Fab. 184
N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide 5 reactions on Fab. 185

N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide 5 reactions on Fab followed by SPAAC functionalisation. 188
2-Chloro-4,6-dimethoxy-1,3,5-triazine 21 reactions on Fab. 190
3-Chloro-6-phenyl-1,2,4,5-tetrazine 22 reactions on Fab. 191
3-Chloro-6-phenyl-1,2,4,5-tetrazine 22 reactions on Fab with transfer. 193
Reaction of Fab with pyrazine-2,3-dicarbonitrile 194
Reaction of Fab with 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile. 195
Reaction of Fab with 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile with CuAAC attempts. 196
Azide fluor-488 196
Benzyl azide 198
Reaction of Fab with 5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile 200
Reaction of Fab with pyrazine-2,3,5,6-tetracarbonitrile 202
Reaction of Fab with pyrazine-2,3,5,6-tetracarbonitrile with NAT attempt 203
Reaction of Fab with diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate 205
Diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate bridge hydrolysis study 206
Removal of diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate bridge 207
Reaction of Fab with methyl 3-cyanopyrazine-2-carboxylate 208
Reaction of Fab with methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate 210
Methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate bridge stability studies 211
NMM 211
Cysteine 212
Reaction of Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid 213
Reaction of Fab with 2,3-dicyano-N-(prop-2-yn-1-yl)quinoxaline-6-carboxamide. 214
Reaction of Fab with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide 216
Reaction of Fab with 2,3-dicyano-N-(prop-2-yn-1-yl)quinoxaline-6-carboxamide with CuAAC. 217
Reaction of Fab with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide and SPAAC 219
Re-bridging of full antibody with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide 220
Re-bridging of full antibody with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide and SPAAC 222
GSH stability studies – Bridge 223
GSH stability studies – CLT 229
Disulfide competition experiment (GSH Vs Fab)- 2,3-dicyanoquinoxaline-6-carboxylic acid 235
Disulfide competition experiment (GSH Vs Fab)- NMM 238

Thioredoxin (Trx) 240
Reaction of Trx with pyrazine-2,3-dicarbonitrile 241
Modified Trx stability studies 242
Bioactivity measurement of Trx and modified Trx (S82) 243
CLT SDS-page (S83) 244
Bridge SDS-page (S84) 244
Serum Stability Study (S85)[257] 245
Enzyme-linked immunosorbent assay (ELISA) – Trastuzumab against HER2 (S86) 246
7. References 246

Abbreviations

Å	Angstrom
AA	Amino acid
AAR	Alkyne-to-Antibody Ratio
ABNO	9-Azabicyclo[3.3.1]nonane N-Oxyl
ADC	Antibody-drug conjugate
ADCC	Antibody-dependent cell cytotoxicity
ADCP	Affinity peptide mediated regiodivergent functionalisation
ADTA	Amino dithioacetal
AJICAP	Affinity peptide mediated regiodivergent functionalisation
Ala	Alanine
AML	Acute myeloid leukemia
aq.	Aqueous
BBS	Borate Buffered Saline
BCMA	Anti-B-cell maturation antibody
BCN	Bicyclononyne
BCR	B-cell receptors
BME	β-Mercaptoethanol
CB	Conjugation buffer
CBT	2-Cyanobenzothiazole
CDR	Complementarity-determining region
CH	Heavy constant domain
CL	Light constant domain
CLT	Cysteine-to-lysine transfer
CuAAC	Copper-Catalysed Azide-Alkyne Cycloaddition
Cy5	Cyanine-5
Cys	Cysteine
DABCO	1,4-diazabicyclo[2.2.2]octane
DAMN	Diaminomaleonitrile
DAR	Drug-to-Antibody Ratio
DBM	Dibromomaleimide
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DDQ	2,3-Dichloro-5,6-dicyano-1,4-benzoquinone
DFT	Density functional theory
DIPEA	*N*,*N*-Diisopropylethylamine
DISN	Diiminosuccinonitrile
DM1	Mertansine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DTNB	5,5′-Dithiobis(2-nitrobenzoic acid)
DTPM	Dithiophenolmaleimide
DTT	Dithiothreitol
DVD	Dual-variable-domains
ED_{50}	Effective dose for 50% of the population
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDT	Ethanedithiol
EDTA	Ethylenediaminetetraacetic Acid

EEDQ	2-Ethoxy-1-ethoxycarbonyl-1,2-dihydroquinoline
ELISA	Enzyme Linked Immunosorbent Assay
equiv.	Equivalents
ESI	Electrospray Ionization
ET	Ethanethiol
EWG	Electron withdrawing group
FAM	Carboxyfluorescein
FAR	Fluorophore-to-Antibody Ratio
FBDP	5-Fluoro-2,4-dinitrobenzene
Fc	Fragment Crystallisable
Fv	Variable fragment
GSH	Glutathione
HAMA	Human anti-mouse antibody
HC	Heavy chain
HER2	Human Epidermal Growth Factor Receptor 2
HHLL	Heavy-heavy light-light
HL	Heavy Light
HRMS	High Resolution Mass Spectrometry
HSA	Human Serum Albumin
HSAB	Hard and soft acids and bases
IC_{50}	Half maximal inhibitory concentration
iEDDA	Inverse electron-demand Diels Alder
LC	Light chain
LCMS	Liquid Chromatography Mass Spectrometry
LCMS/MS	Tandem Liquid Chromatography Mass Spectrometry
LG	Leaving group
Lys	Lysine
mAb	Monoclonal antibody
MC	Maleimidocaproyl
MESNa	Sodium 2-mercaptosulfonate
min	Minute
MMAE	Monomethyl auristatin E
MMAF	Monomethyl auristatin F
MPAA	4-mercaptophenylacetic
MS	Mass spectroscopy
MTBE	Methyl tert-butyl ether
MWCO	Molecular Weight Cut Off
NAT	Nitrile–aminothiol
NBS	N-Bromosuccinimide
NBT	Nitrile Bis-thiol
NCL	Native chemical ligation
NGM	Next-generation maleimide
NHS	*N*-hydroxysuccinimide
NK	Natural killer
NMM	N-methyl maleimide
NMR	Nuclear magnetic resonance
PABC	*p*-Aminobenzylcarbony
PB	Phosphate Buffer
PBD	Pyrrolobenzodiazepine

PBS	Phosphate Buffer Saline
PD	Pyridazinedione
PDI	Protein disulfide isomerases
PDT	Propanedithiol
PEG	Polyethylene Glycol
PSA	Polar surface area
PTM	Post-translational modification
RDS	Rate-determining step
RT	Room Temperature
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SMCC	Succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate
S_N2	Nucleophilic Substitution, Second Order
S_NAr	Nucleophilic Aromatic Substitution
SPAAC	Strain-promoted azide-alkyne cycloaddition
TCEP	Tris(2-carboxyethyl)phosphine
TCO	*trans*-cyclooctene
TEA	Triethylamine
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin-later chromatography
TS	Transition state
UAA	Unnatural Amino Acids
Val	Valine

1. Introduction

1.1. Protein modification

Proteins are a vital component of the human body responsible for carrying out roles essential for survival. Genes provide the blueprint for proteins and peptides in the body and are sequences located along chromosomes.[1] The human genome project, completed in 2003, uncovered approximately 20,000-25,000 protein-coding genes, each gene is able to generate as many as 100 different proteins, resulting in an exceptionally complex library of proteins and peptides.[2,3] Protein modification is an important process in physiological biochemistry, contributing to this vast complexity of protein structures and functions. There are two main mechanisms by which nature expands diversity, the first of which is modification of proteins at the transcriptional level *via* alternative mRNA splicing.[4] The second is through post-translational modifications (PTMs) which are carried out *via* a range of enzymes.[5] Common examples of PTMs include acetylation, glycosylation, phosphorylation[6], alkylation and oxidation[7] (Figure 1). These modifications play key roles in regulating biological activities, such as cell apoptosis, cellular function and gene silencing.[8] Due to the intricate nature of biological processes, an irregular status of PTMs can often result in the development of disease.[9] The study of these PTMs can provide information regarding the functions of specific proteins in their innate systems, which can be applied to our understanding of disease advancement.

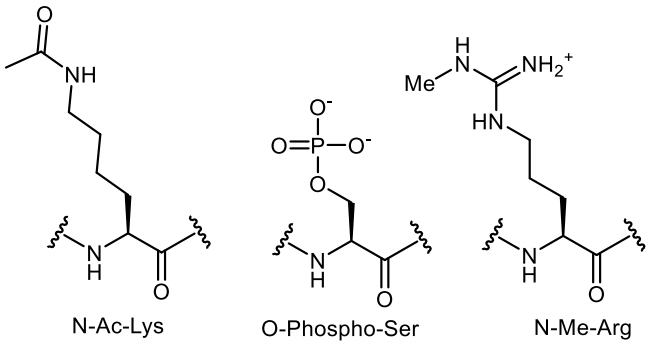


Figure 1. Examples of post-translational modification.

In addition to its importance in nature, endogenous protein modification is also highly exploited in chemical biology, particularly for bioconjugation. The covalent combination of a biological component and a chemical moiety is of huge interest. Bioconjugation has made it possible to probe and modulate biological processes *via* labelling of the proteins and peptides involved; well-known examples being fluorescent tagging[10] and targeted therapeutics.[11] A rapidly evolving sub-category of the latter include antibody-drug conjugates (ADCs).

1.2. Antibodies

Antibodies are distinctly Y-shaped, macromolecular glycoproteins generated by B-cells as part of the adaptive immune system. Antibodies are initially found anchored to the cell surface membrane of B-cells and act as B-cell receptors (BCRs).[12] Activation of these receptors *via* an antigen results in the rapid proliferation of the specific B-cell and thereby secretion of antibodies into the bloodstream. Once released into the bloodstream, antibodies permeate further into bodily fluids. Antibodies can modulate the immune response in multiple ways: 1) Pathogen neutralisation; this is the process by which unassisted antibodies bind and immobilise pathogens in the bloodstream, impeding initial pathogenic stages on host cells. The neutralisation of pathogens by lone antibodies can also result in the aggregation of organisms, and this aggregation enhances the capture of microbes in mucous and therefore clearance.[13] 2) Antibody-dependent cell-mediated cytotoxicity (ADCC); where Fc (fragment crystallisable) receptors (FcR) bearing effector cells form antigen-antibody complexes with target cells (malignant or antigen presenting cells (APC)), and in turn these complexes are recognised by natural killer (NK) cells. These cells release a family of serine proteases known granzymes which result in cell lysis.[14] This form of action plays a greater role in the clinical effect of anti-cancer mAbs (e.g. trastuzumab).[15,16] 3) Induction of phagocytosis; as with ADCC, FcR-bearing effector cells recognise antigens directly associated with pathogens or on APCs. This recognition results in antibody-dependent cellular phagocytosis (ADCP).[17]

These Y-shaped glycoproteins consist of two identical light chains (LCs) and two identical heavy chains (HCs) interlinked *via* disulfide bonds. There are two isotypes of mammalian LC; κ and λ, whereas the HC can be categorised into one or five isotypes; IgA, IgD, IgE, IgG, or IgM. Both HC and LC consist of a *N*-terminal variable domain (V_H and V_L respectively), these regions are responsible for antigen binding *via* the complementarity-determining regions (CDRs) present. LCs also consist of one *C*-terminal constant domain (C_L), in comparison HCs can contain up to three (C_{H1}, C_{H2}, C_{H3}). All antibodies also consist of an Fc region, as mentioned before the Fc region is involved in the mediation of ADCC, however the Fc region also possesses other roles such as FcRn (neonatal Fc receptor) interactions to increase the serum half-life of antibodies.[18]

The HC isotypes can vary in several ways including differences in the sequence and numbers of constant domains resulting in divergent effector functions.[19] Of these isotypes, IgG is the most abundant in human serum. IgG is further differentiated into four subclasses: IgG1, IgG2, IgG3,and IgG4 (in descending abundance). Each subclass differ in structures and functions, despite a 90% similarity on the amino acid level.[18] (Figure 2)

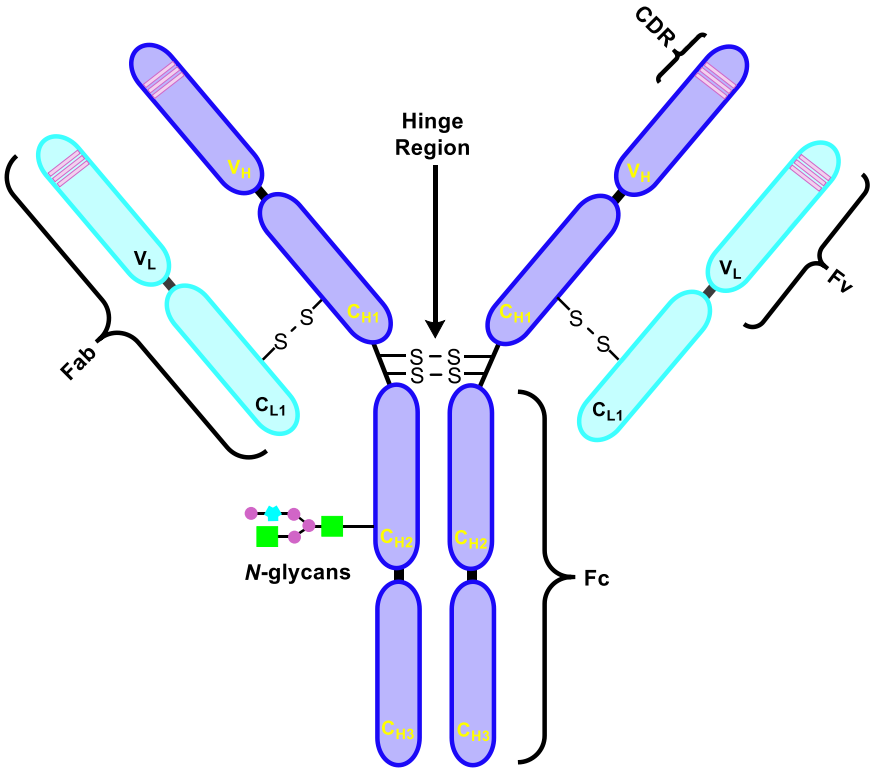


Figure 2. Structure of a human IgG1 antibody (schematic representation). The purple segments represent the heavy chains, and light blue segments represent the light chains.

A key property of antibodies is their high selectivity and affinity for their corresponding antigens, thus they are incredibly desirable as potential targeted therapeutics. In 1975, Kohler and Milstein developed the first instance of large-scale production of monoclonal antibodies (mAbs) through use of hybridoma technology. The mAbs provide a cohort of identical antibodies targeting a single antigen counterpart. Initial variations of mAbs were produced using murine proteins.[20] In 1986, the first mAb (muromonab) was approved for use in humans by U.S. Food and Drug Administration (FDA) for transplant rejection. However, due to the murine nature of muromonab, several complications arose. For example, the mAbs were highly immunogenic leading to human anti-mouse antibody (HAMA) response.[21] The HAMA response is the production of human antibodies in response to the administered murine mAbs. Murine mAbs were also found to interact poorly with the human immune system and are therefore unable to recruit immune effector functions as effectively as intended.[22] In order to engineer more human-like antibodies with desirable efficacy and lower immunogenicity, various methods of humanisation of mAbs were developed.[23] Chimeric antibodies are the first examples of engineered improvement where non-human V domains are fused with human C domains, generating mAbs with >70% human content.[22,24] The next stage of engineering was the humanisation process. Here, only the CDRs are of murine origin e.g. trastuzumab.[25,26] Finally, fully human mAbs are fully derived from human amino acid sequences.[24] Since the

development of fully humanised mAbs, Humira (adalimumab), developed for use in conditions such as rheumatoid and psoriatic arthritis, has become the world best-selling drug ($20.4 billion sales in 2020).[27] As of 2021, 131 mAb products have been brought to the market, 59 of which were of oncological use.[28]

1.2.1. Antibody fragments

Full antibodies can be divided into different fragments with varying size: $F(ab')_2$, Fab, Fv, Fc and scFv (Figure 3).[29,30] Despite all the significant advantages of full antibodies, they still exhibit limiting factors including cost of production due to PTMs such as glycosylation. Another factor is the penetration of the antibody into the target cells, such as tumours. The primary form of full-antibody transport into solid tumours is *via* diffusion - the rate of diffusion for larger molecules is inherently slower, therefore, this can hinder the efficacy of full antibodies. The lesser size of antibody fragments results in greater rates of diffusion, leading to a greater efficacy compared to full antibodies.[31,32] Fragments are also able to access cryptic antigen epitopes present in narrow cavities, which are normally poorly accessible by full antibodies e.g. G protein–coupled receptors and enzyme active sites.[33] Fragments however exhibit reduced stability and a tendency to aggregate. The reduced stability is associated with the lack of Fc region.[30] The lack of Fc region however allows fragments to bypass cytotoxic immune mechanisms such as complement-dependent cytotoxicity (CDC) or antibody-dependent cell-mediated cytotoxicity (ADCC).[34]

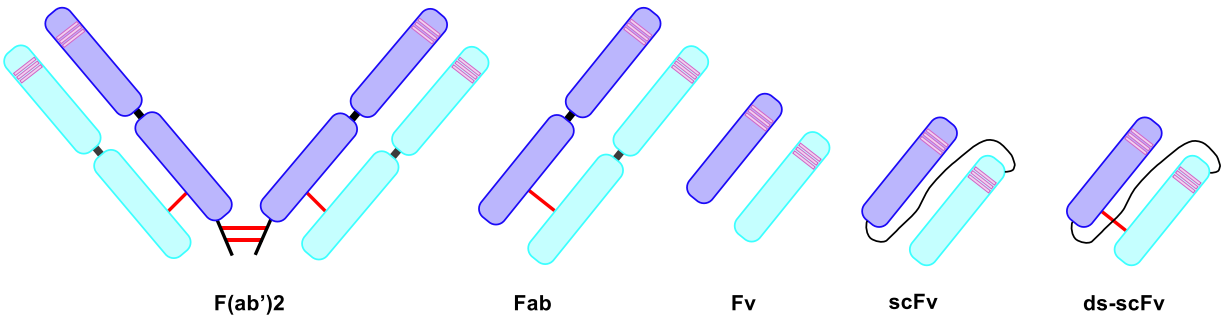


Figure 3. Structure of antibody fragments (Red lines denote disulfide bonds)

Digestion of full antibodies by the protease enzyme pepsin leads to the generation of a dimeric $F(ab')_2$ fragment *via* the degradation of the Fc region.[35] The $F(ab')_2$ fragment consists of two identical monovalent Fab fragments joined thorough a pair of disulfide bonds. Digestion of full antibodies by another protease enzyme, papain, leads to the generation of a single Fc fragment and two separate but identical Fab fragments.[35] The Fc fragment mediates the recruitment of effector functions whilst the Fab fragments are responsible for antigen binding. Each Fab fragment is made up of one full LC and the V_H and C_{H1} section of one heavy chain.[36]

Whilst protease mediated digestion is used to produce the relatively larger fragments such as $F(ab')_2$ and Fc, even smaller fragments are generated *via* recombinant techniques. These techniques have allowed the production of a range of smaller fragments by direct expression from bacterial and mammalian cells.[37] The smallest fragment which maintains comprehensive antigen binding is the Fv fragment, comprising solely of V regions (V_H and V_L). In order to improve stability of these fragments, the two V regions are connected *via* a soluble and flexible amino acid peptide linker generating a single-chain Fc fragment (scFv).[38]

Antibody fragments present pharmacokinetic advantages resulting from the increased diffusive capability, coupled with the ability to reach antigen binding sites inaccessible to full size mAbs. Production of fragments is also proven to be simpler and less costly to produce, due to the lack of glycosylation and relatively small size, which permits the use of expression systems.[30] Antibody fragments however are limited by rapid clearance, due to the renal cut off (~60 kDa). Fragments with molecular weights below this cut off suffer from rapid renal clearance. This can negatively impact the effective dosage of fragment that reaches target cells such at tumours.[39] Tactics have been employed in order to prolong the half-life of fragments which include PEGylation,[40] fragment multimerization[41] and conjugation to human serum albumin (HSA).[42]

As of 2020, seven FDA-approved fragments were available, of which more than 50% are Fab based; Abciximab, Ranibizumab, Certolizumab pegol, and Idarucizumab. Abciximab was the earliest Fab therapeutic approved in 1994.[43] This Chimeric IgG1 Fab was used for the reduction of ischemic complications during coronary angioplasty.[44] Idarucizumab is the latest Fab therapeutic approved in 2015,[43] and is a humanised Fab used for the reversal of anticoagulant effects of dabigatran, an oral thrombin inhibitor.[45] The remaining three are Fv-based: Blinatumomab, Moxetumomab pasudodox, Brolucizumab. The first Fv-based therapeutic was Blinatumomab, a murine bispecific tandem scFv approved in 2014. Blinatumomab is an antineoplastic fragment used to treat acute lymphoblastic leukaemia.[46] The most recent Fv fragment, Brolucizumab, was approved in 2019. Brolucizumab is a humanised scFv used in the treatment of age-related macular degeneration.[47] The common theme between both types of fragment development is the clear move towards more humanised forms of antibody treatments.

1.3. Antibody drug conjugates

When applied to oncological treatments, lone mAbs are often insufficiently potent and are susceptible to resistance mechanisms.[48] A potential solution comes in the form of antibody drug conjugates (ADCs). These immunoconjugates are monoclonal antibodies (mAb) with cytotoxic drugs attached *via* a carefully designed linker to produce highly efficacious yet

selective agents. This is achieved by synergising the innate selectivity of mAbs with the cytotoxicity of the drug.[49] The development of ADCs has allowed for their use as chemical probes and therapeutic agents.[50] The added selectivity of mAbs limits the off-site side effects that are typically associated with the promiscuous cytotoxic drugs utilised. This allows for expansion of the therapeutic index (the range of doses at which a therapeutic is effective without undesirable adverse events) and thus allows use of drug candidates that would otherwise be considered too toxic to be employed as lone therapy,[51] such as tubulin inhibitors.[52] Paul Ehrlich's 'magic bullet' proposal – a treatment that selectively targets only diseased tissues – is made viable through well-designed ADCs.[53]

The first ADC approved by the FDA in 2000, was gemtuzumab ozogamicin (Mylotarg), which was approved for use in acute myeloid leukaemia (AML) for patients aged over 60. Mylotarg is constructed by the conjugation of gemtuzumab (mutated anti-CD33 IgG4) to the DNA-cleaving compound, calicheamicin, *via* a cleavable hydrazone linker.[54] Issues including premature payload release, eventually saw Mylotarg being removed from the market; the linker exhibited instability *in vivo* hence leading to the premature release in circulation.[55] Mylotarg was reapproved in 2017 to meet remaining needs in CD33-positve AML patients, displaying the value of such therapeutics despite the drawbacks. Since then, a total of 14 clinically approved ADC treatments have become available; all for oncological use. [56,54] This limited list of ADCs over a decade exhibits the difficult nature of these novels agents. Despite the associated difficulties, the field of ADCs is rapidly growing.

The mechanism of action behind ADCs varies, however, the generalised mechanism involves the internalisation of the ADC, initiated *via* antigen recognition on the target cell. Endocytosis leads to the formation of an early endosome. Here, cleavable linkers are able to release payloads into the cell whereas non-cleavable linkers remain intact. The early endosome progresses to a late endosome and finally a lysosome where the ADC is broken down - it is here that non-cleavable linkers finally release their payload.[57] In addition to lysis of the internalising cell, the payload is also able to kill nearby target cells *via* the 'bystander effect'. The payload is able to diffuse out of the primary cell or is released upon cell death.[58] The bystander effect is promoted by the use of a cleavable linker and hydrophobic payload, and this effect is particularly beneficial in targeting heterogeneous antigen expressing tumours. The ADC is then able to kill both antigen-positive and antigen-negative target cells at the tumour site. [59]

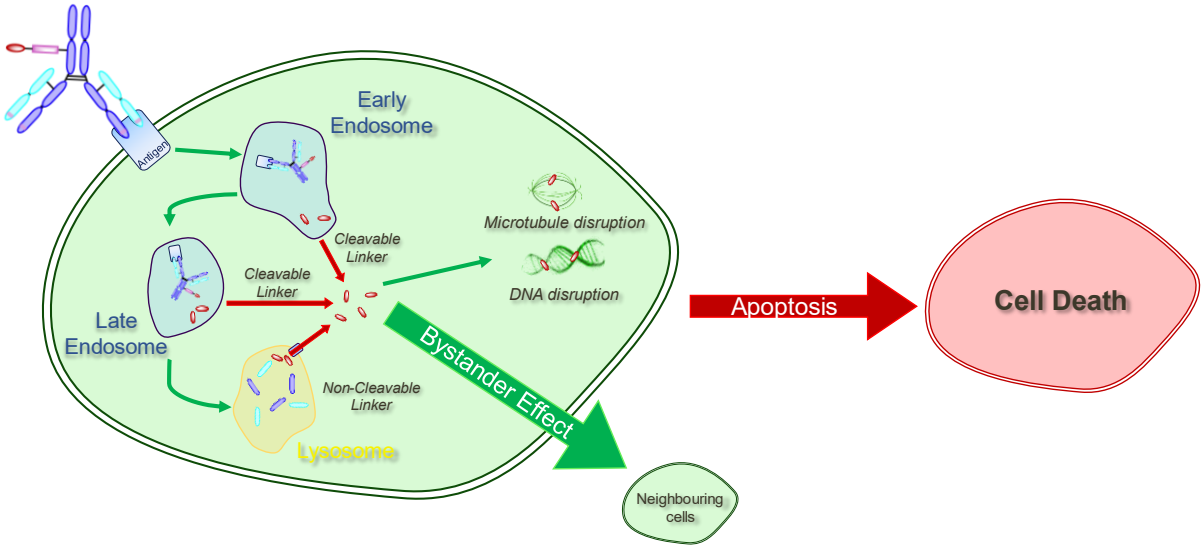


Figure 4. Typical pathway for ADC mechanism of action. Internalisation leads to the release of drug via a cleavable linker or protein digestion (non-cleavable linker release).

1.4. Payload

As mentioned earlier, the amplified selectivity provided by the mAb allows for highly toxic payloads to be utilised. These payloads should ideally employ maximal toxicity at a minimal concentration, and traditionally these payloads will have been deemed too toxic for general medical use. The payload potency must also be high enough to elicit a cytotoxic effect at low concentrations, due to the low uptake rate of mAbs. For example Tositumomab, a mAb utilised for the imaging and therapy of lymphoid malignancies, only has an average tumour uptake of 0.009% injected dose per gram of tumour.[60] These payloads often target cell proliferation *via* DNA synthesis or mitosis.[61] Alongside payload potency, hydrophobicity is another factor which must be taken into account, as increased hydrophobicity can lead to antibody aggregation and therefore accelerated clearance.[62]

1.4.1. DNA-damaging drugs

This class of drug is thought to have impacts throughout the different phases of the cell cycle. Examples include double strand break agents, topoisomerase I inhibitors and alkylators.

1.4.1.1. Double strand break agents

Calicheamicins are powerful antitumour antibiotics originally isolated from the actinomycete *Micromonospora echinospora*. Calicheamicins are found to bind to the minor groove of DNA. The initial nucleophilic attack is followed by a thiol triggered cascade, generating a highly reactive 1,4-dehydrobenzene diradical. This diradical species is able to strip hydrogen atoms from the phosphate backbone of DNA strands, resulting in the lethal scission of the DNA

double helix.[63] The two approved ADCs which use Calicheamicin (Figure 5) are Mylotarg and Besponsa.

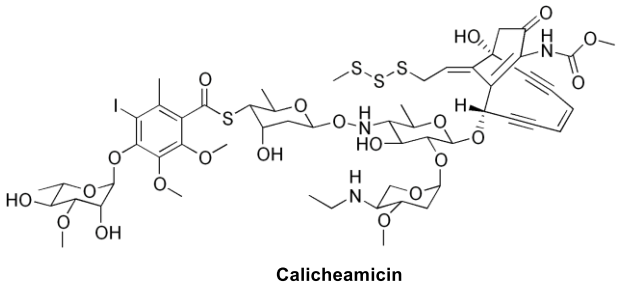


Figure 5. Chemical structure of the payload Calicheamicin

1.4.2. DNA cross-linkers

The pyrrolobenzodiazepines (PBD) are a class of drugs derived from the naturally occurring anthramycin family of antitumor antibiotics, first isolated from *Streptomyces refuineusin* in 1965.[64] The core monomer PBD contains the same skeletal structure; a substituted aromatic ring, a diazepine ring and a pyrrolidine ring with an *S*-chiral centre at the C11a position (Figure 5). The 3D shape of PBDs allows for a most favourable fit within the DNA minor groove.[65] PBDs also possess an electrophilic moiety such as an imine, which is able to react with the nucleophilic amine present on the nucleobase guanine (Figure 6).[66] DNA binding by PBD monomers triggers a series of events, such as DNA strand breakage, which leads to cell death. PBD monomers have been shown to have preference for 5'-purine-guanine-purine-3' sequences.[67]

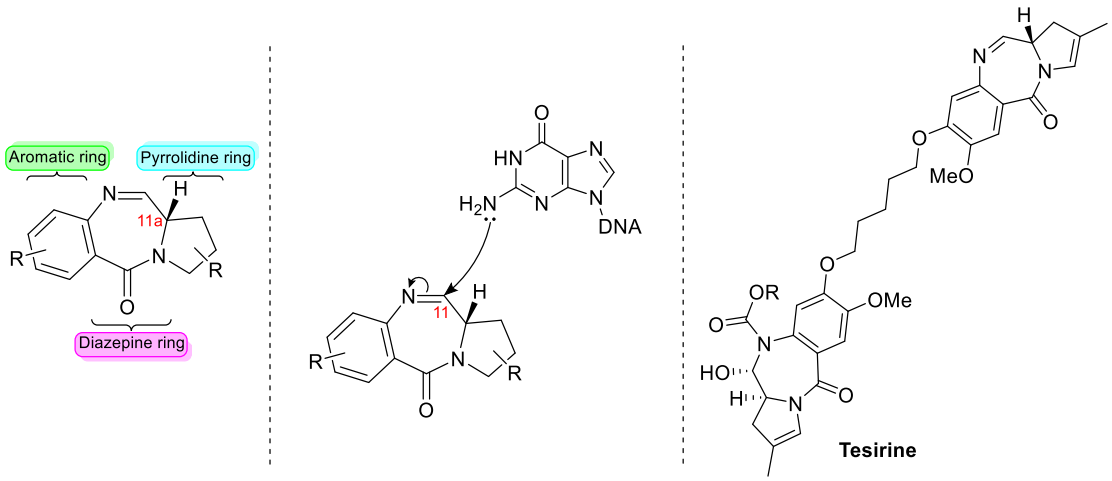


Figure 6. (a) Generalised PBD structure, (b) PBD mechanism of action, (c) chemical structure of the payload PBD dimer tesirine.

A series of synthetic PBDs have been developed since the discovery of anthramycin, joining two PBD monomers at the C7 or C8 position leads to the formation of DNA cross-linking PBD dimers. These dimers possess the ability to form both interstrand and intrastrand cross-links

in addition to the monomeric alkylation of DNA.[68] Molecular modelling of PBD cross-linking was carried out alongside thermal denaturation studies, and it was found that there was little to no impact on the secondary structure of DNA post adduct formation. The lack of structural disruption has been thought to provide PBD dimers with an outlet to avoid DNA repair enzymes, as normally repair enzymes rely on tracing distortion or helical perturbation in DNA.[69] This provides PBDs with a distinct advantage over traditional alkylating agents such as oxaliplatin and *cis*-platin as they are able to continue to kill target cells.[70] As of 2023 PBDs are found in 1 of the 12 approved ADCs currently approved; Zylonta. Zylonta utilises the PBD dimer tesirine, which was designed with desirable physicochemical properties i.e. low hydrophobicity whilst retaining a high degree of potency.[71]

1.4.3. Tubulin inhibitors

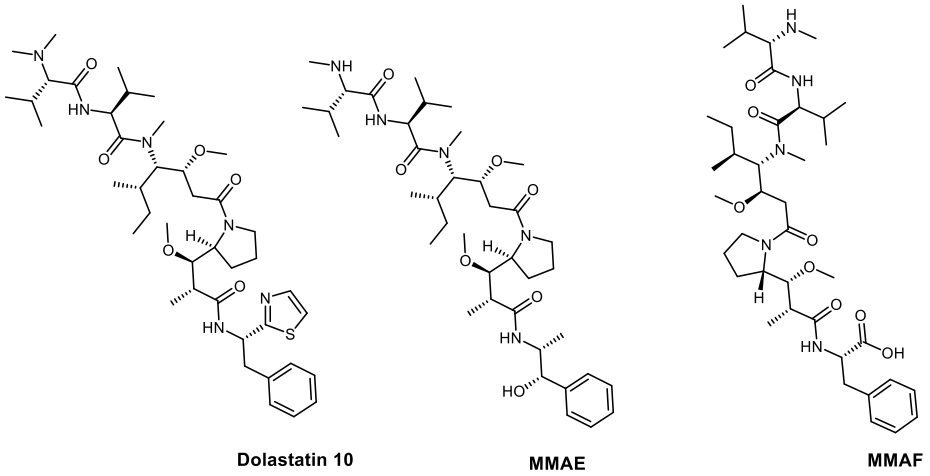


Figure 7. Chemical structure of the payloads Dolastatin 10 , MMAE and MMAF.

Microtubules are an active assembly of α- and β-tubulin heterodimers which are involved in many cellular processes, such as the maintenance of cell structure, cell division and intracellular transport. Interference of these microtubules can result in cell lysis *via* arrest of mitosis.[72] Currently, auristatin derivatives are the most commonly used payloads for ADCs (4 out of 9), which include monomethyl auristatin E (MMAE) and monomethyl auristatin F (MMAF). (Figure 7) Auristatins are synthetic derivatives of dolastatin 10, an anti-mitotic pentapeptide which inhibits tubulin-dependent GTP binding and blocks the polymerisation of tubulin in a non-competitive manner, leading to the arrest of mitosis and resultant apoptosis.[72] Adcetris, Polivy and Padcev all use MMAE as the selected payload whilst Blenrep uses MMAF.[73] MMAE has greater membrane permeability and a lower IC_{50}, whereas MMAF is more hydrophilic and therefore has a lower tendency to aggregate.[74]

Another class of tubulin inhibitors are maytansinoids, though currently only one approved ADC (Kadcyla) utilises a maytansinoid as the payload of choice (DM1). This class of inhibitor is derived from maytansine, a powerful plant toxin. Maytansinoids suppress dynamic instability of microtubules, which inhibits the innate growth and shrinking of tubules necessary for mitosis, consequently inhibiting proliferation and resulting in cell apoptosis.[75] Metabolites of DM1 (*S*-methyl DM1) have also been found to inhibit cell proliferation at very low concentrations (Figure 8).[76] Due to the hydrophobic nature of DM1, Kadcyla is prone to aggregation, utilising hydrophilic PEG linkers as a can provide a way to overcome such hydrophobicity.[77]

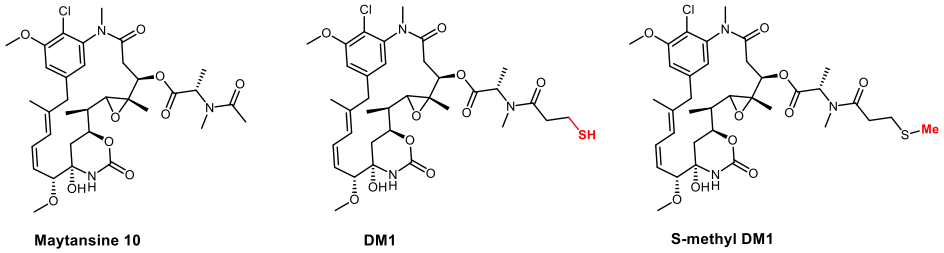


Figure 8. Chemical structure of the payloads maytansine 10, DM1 and *S*-methyl DM1.

1.4.3.1. Topoisomerase I inhibitors

DNA topoisomerases are heavily involved in the synthesis of DNA. Type I in particular cleaves one strand of duplex DNA, in order to relax DNA supercoiling generated by transcription, replication, and chromatin remodelling.[78] Topoisomerase I is thought to form a covalent linkage to the terminal phosphate of the cleaved DNA stand, while the other strand is held in close proximity *via* non-covalent interactions. The non-cleaved strand passes through the DNA break followed by subsequent re-ligation and DNA relaxation.[79] Targeting these enzymes can lead to the disruption of normal cell function and ultimately lead to cell lysis.

Trodelvy is an ADC approved for use in triple negative breast cancer. It employs an anticancer drug SN-38, an active metabolite of the prodrug irinotecan. Irinotecan is metabolised by an endogenous carboxylesterase into SN-38 which is found to be significantly more cytotoxic than the parent compound. Trodelvy targets TROP-2-expressing cancer cells and is internalised, which delivers the cytotoxic payload of SN-38 to the target cells, releasing the payload *via* a cleavable linker. The drug-to-antibody ratio (DAR) of Trodelvy is 7.6, which is 2.15-fold greater than that of Kadcyla. The higher DAR can be attributed to the lower hydrophobicity of SN-38 in comparison to other payloads, which allows for higher equivalents of payload and therefore greater amounts of drug reaching target cells without compromising the ADC's pharmacokinetics (Figure 9).

Enhertu is another approved ADC utilising topoisomerase I inhibitors, specifically DXd (deruxtecan) (Figure 9).[80] As with Kadcyla, Enhertu is employed in the treatment of HER-2 positive breast cancer, however it is found to outperform Kadcyla. The performance discrepancy originates from the nature of the DAR. As with SN-38, DXd is also less hydrophobic than DM1, and this allows for a DAR of 7.7 which is 2.18-fold greater than Kadcyla.

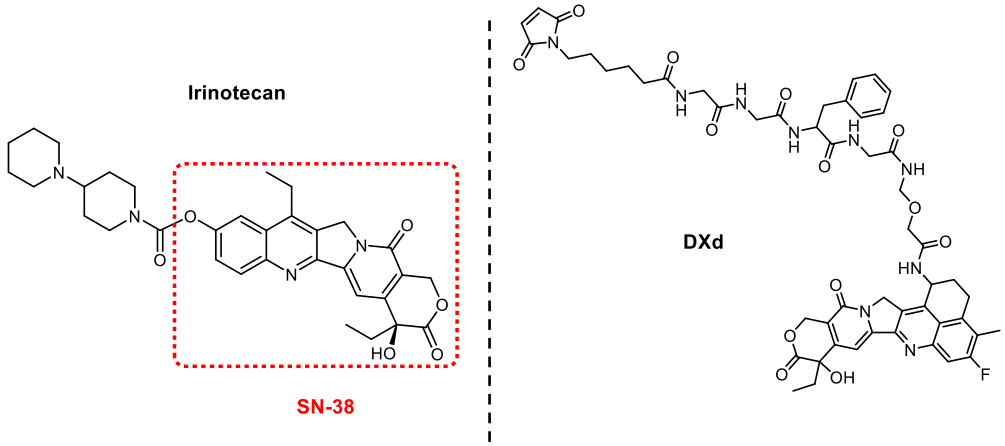


Figure 9. Chemical structure of the payloads Irinotecan, SN-38 (Highlighted red) and DXd.

1.5. Linkers

The mechanism of action behind ADCs can be associated to multiple factors, from the specific drug utilised to the type of mAb itself. Another important factor to consider is the linker which attaches the drug to the antibody (Figure 10). The linker is a vital component of both internalising and non-internalising ADCs – it must remain stable in circulation to avoid premature release of the payload. The process of internalisation of ADCs initiates release of the active payload and thereafter cell cytosis – this release can be achieved by either a cleavable or non-cleavable linker design.[81] Non-internalising ADCs instead release payloads in an extracellular fashion; ADCs of this nature are yet to be clinically approved. Linker design takes two points of attachments into consideration: the linker-antibody attachment and the linker-payload attachment. The former is crucial as it determines the DAR and therefore homogeneity of ADCs.[82]

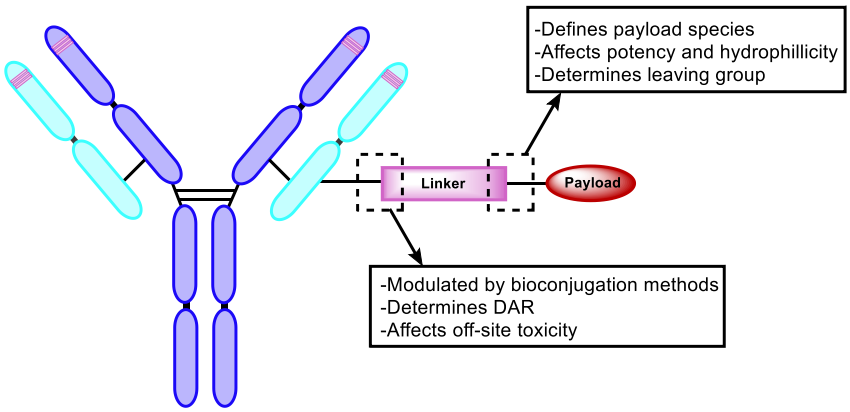


Figure 10. Generalised structure of an ADC - consisting of mAb, Linker and payload.

1.5.1. Cleavable linkers

Cleavable linkers respond to unique microenvironments within the cell resulting in payload release. These include i) acid cleavable linkers, such as hydrazone linkers, which release the drug in the low pH environments of the endosomes (pH 5.5–6.2) and lysosomes (pH 4.5–5.0); ii) reducible disulfide linkers, as the cell cytosol contains high levels of glutathione which is able to cause reductive cleavage; and iii) enzyme cleavable linkers, which employ proteolytic enzymes in the cell such as the cathepsin family of proteases.[82,55]

1.5.1.1. Acid cleavable linkers

The first generation of cleavable linkers came in the form of hydrazone, present in earlier editions of clinical ADCs. BR96-Doxorubicin (BR96-Dox) demonstrated the first example of a hydrazone linker, and is an ADC comprised of a chimeric human/murine mAb decorated with approximately eight doxorubicin molecules.[83] BR96-Dox is internalised *via* endocytosis into cell lysosomes which poses an acidic pH of 4.5 – 5.0,[84] triggering the acid catalysed hydrolysis of the hydrazone linkers. ADCs designed using traditional hydrazone linkers have been found to have issues regarding premature release due to the pH related instability, exemplified by the clinically approved ADC Mylotarg.[85] As mentioned earlier, Mylotarg was subject to temporary withdrawal from the market. The withdrawal was in part due to the *in vivo* stability issues exhibited by the hydrazone linker. ADC studies showed the linker to be stable in buffer however this was not translated over to plasma where the instability was realised.[86] Besponsa, another ADC clinically approved in 2017, also utilises an acid labile hydrazone linker.[87] Incorporation of these acid cleavable linkers in more modern ADCs fortifies the validity of these linker types. Other examples of acid cleavable linkers include carbonates, found in Trodelvy, and silyl ethers which are yet to be tested in clinical conditions.[88]

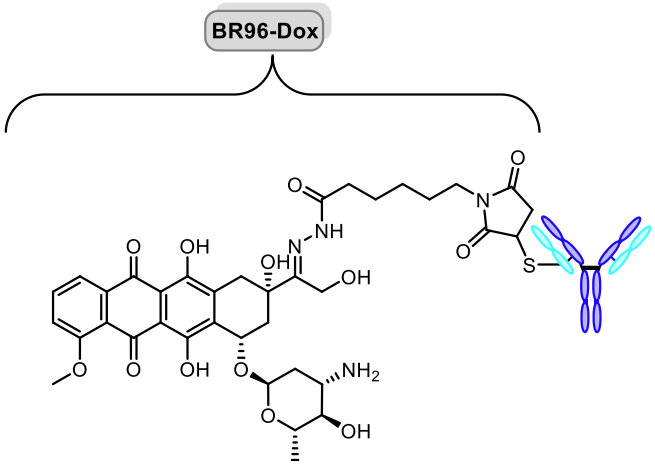


Figure 11. Structure of acid-cleavable hydrazone linkers used in BR96-Dox.

1.5.1.2. Reducible linkers

Reducible linkers utilise the difference in reduction potential between the intracellular fluid and the blood plasma. These linkers are designed to be stable to the majority of the physiological pH encountered, and instead release is triggered by the intracellular thiol-mediated reduction. The thiols displace the payload in the ADC thereby releasing the drug into the desired location. Thiol sources within the system include HSA, thioredoxin and glutathione (GSH). Despite the abundance of HSA in the blood (~600 µM),[89] the nucleophilic thiol (Cys-34) is found within a buried hydrophobic pocket of the protein (9.5-10 Å), making it difficult to access and trigger release.[90,91] GSH is the major mediator of the linker reduction, present in intracellular fluid.

Further selectivity of payload release comes from the extensive difference in cytoplasmic GSH concentration between healthy cells and malignant cells. Blood plasma also exhibits a significant difference in GSH concentration from the cytoplasm of cells, thus avoiding unselective/premature release whilst the ADC is in circulation. The normal GSH concentration ranges from 1–10 mM whilst tumour cells can contain concentrations up to 1000-fold higher.[92] This higher concentration comes from the oxidative stress stemming from rapid proliferation.[93] Tumour cells also overexpress protein disulfide isomerase (PDI) enzymes, which carry out a host of responsibilities including the formation and reduction of disulfides which may potentially contribute to payload release in malignant cells. PDIs however, are yet to be proven to contribute to the release of payload via disulfide reduction.[94] Progressive death of tumour cells also results in the release of increasing amounts of GSH leading to an intensification of drug release.[95] Optimisation of the steric hindrance surrounding the disulfide bond allows for the modulation of the efficacy.[96] Studies carried out a panel of huC242 conjugates, revealed a structure-activity relationship between hindrance and efficacy. Dimethylation on the maytansinoid payload side and no methylation on the linker side of the disulfide bond resulted in the best efficacy. The data also indicated the presence of the methyl groups increased the

hydrophobicity and consequently the cell permeability, these factors positivity impacted the ADC potency.[97] Mylotarg and Besponsa are approved ADCs which feature a linker containing both acid-labile, and reducible disulfide elements.

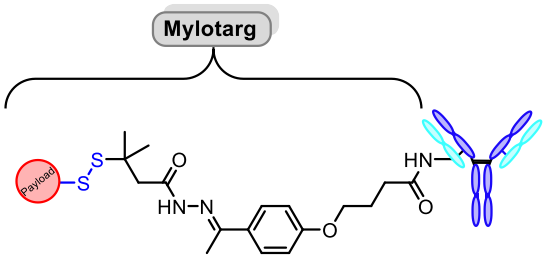
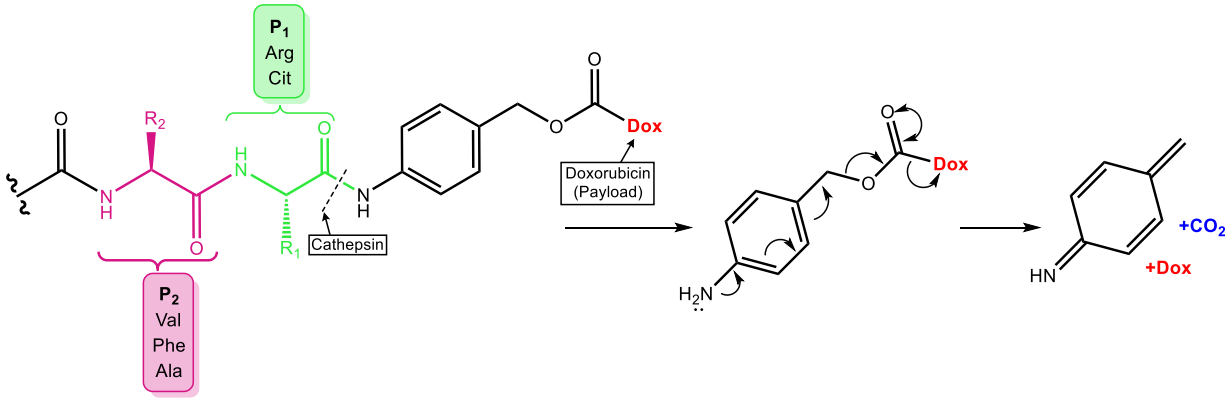


Figure 12. Structure of reducible disulfide linker (blue) used in Mylotarg.

1.5.1.3. Enzymatically cleavable linkers

These linkers take advantage of the abundance of hydrolytic enzymes able to recognise and cleave specific sequences (peptide or carbohydrate). Enzyme-cleavable linkers are the most successful and utilised varieties of cleavable linkers, and they can be found in six of the nine approved cleavable linker ADCs on the market.[98] During the initial development of these linkers, two common trends were identified; 1) a hydrophilic residue is required at P_1 and the rates of hydrolysis increases with basicity of this residue. As citrulline (Cit) is isoelectronic with Arg, and is more synthetically accessible, it is often the choice at P_1. 2) Hydrophobic residues at P_2 enable proteolytic cleavage whilst imparting plasma stability.[82]



Scheme 1. The structure and cleavage mechanism of the Aaa-Aaa-PABC-Doxorubicin (Dox) motif. Hydrolytic cleavage leads to a 1,6-elimination and traceless release of Doxorubicin.

The most common enzyme cleavable linker is the protease sensitive valine-citrulline (Val-Cit) dipeptide. Val-Cit linkers respond to lysosomal proteases, namely cathepsin B, which are proteases commonly overexpressed in tumour cells and therefore provide an ideal target mechanism for release. Cleavable dipeptides were first developed in 1998 for (Dox)-prodrugs, and the linkers were combined with *p*-aminobenzylcarbonyl (PABC) spacers in order to

compensate for the steric bulk introduced by large doxorubicin payloads.[99] PABC linkers are self-immolative spaces and undergo spontaneous 1,6-elimination upon proteolysis, which allows for the release of naked, unmodified payloads (Scheme 1).

Despite the selectivity of enzymes, there is still scope for premature release, which results from the simple nature of the dipeptide and the ability to be recognised by more than one hydrolytic enzyme. Val-Cit for example can also be cleaved by carboxylesterase 1C (Ces1C), a serum-based serine esterase.[100] The catalytic serine of CesC1 is thought to lie deep within a narrow pocket, whereas the active site of cathepsin B is much shallower in nature and therefore lends itself to the promiscuous behaviour exhibited. Chemical substitutions at the P_3 position of the Val-Cit-PABC linker provided stabilisation of the motif, providing resistance to the aforementioned CesC1 hydrolysis. (Figure 13).[100]

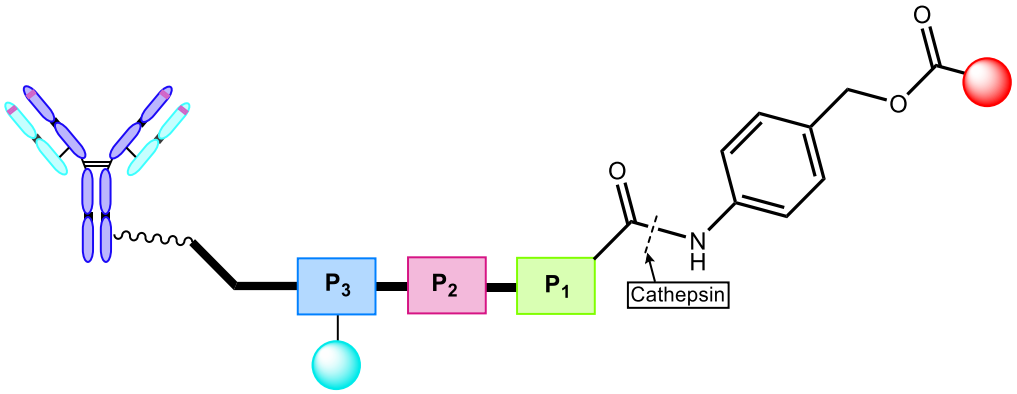


Figure 13. The structure of the Val-Cit-PABC motif. Blue sphere - chemical substitutions, red sphere - payload

Recent work has shown the viability of arylsulfate motifs as enzymatically cleavable linkers, these moieties exploit lysosomal sulfatases. A key benefit over previous peptidic linkers, such as Val-Cit or Val-Ala is the lack of inherently high hydrophobicity. The lower hydrophobicity provides several advantages including; reduced aggregation and allows the use of lipophilic payloads.[101]

1.5.2. Non-cleavable linkers

Non-cleavable linkers rely on the proteolytic degradation of the mAb scaffold post-internalisation for the release of payload, examples of such linkers include amide and thioether linkers.[102] Due to the nature of non-cleavable linkers, they often exhibit superior serum stability in comparison to cleavable linkers, and thus this greater stability also helps minimise the negative off-site effects.[73] The downside to non-cleavable linkers is the loss of the bystander effect, due to the zwitterionic properties of the catabolites.[103] Kovtun *et al*. found that ADCs with cleavable linkers, specifically reducible disulfide bonds, are effective against both antigen-positive and antigen-negative cancer cells when in close proximity to the target cells due to the bystander effect.[75] This effect is often required when the target antigen is expressed in

tumours in a heterogeneous fashion hence,[104] non-cleavable linkers are therefore applied in the treatment of haematological tumours and cancers with homogeneous levels of antigen expression, for example breast cancer.[73] In cancers where the tumour cells express the antigen more homogeneously, the bystander effect can kill nearby antigen-negative cells, illustrating the need for both variants of linker. Currently only two approved ADCs employ non-cleavable linkers: Kadcyla and Blenrep. Both Kadcyla and Blenrep utilise maleimides in their linker makeup however in differing manners.

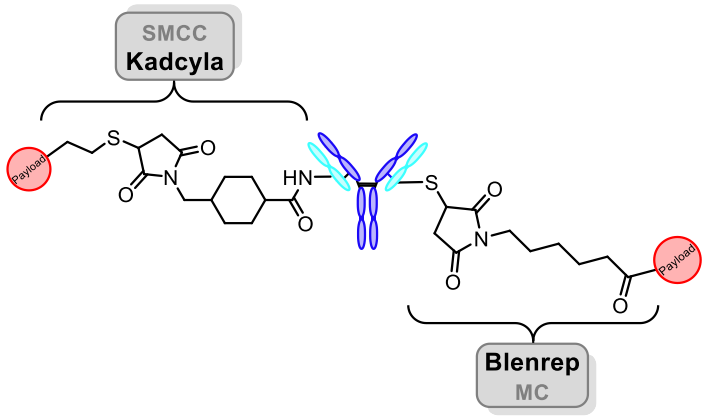


Figure 14. Linkers used in Kadcyla and Blenrep (SMCC and MC respectively).

Kadcyla utilises the well-defined linker SMCC (N-succinimidyl-4-(maleimidomethyl) cyclohexane-1-carboxylate) to connect trastuzumab to the DM1 payload. The linker attaches to the trastuzumab via a Lys residue and consists of a thioether bond courtesy of the maleimide.[87] Blenrep employs the maleimide from maleimidocaproyl (MC) directly in the conjugation to a Cys residue present on belantamab, an anti-B-cell maturation antibody (BCMA). Blenrep is a first-in-class anti-BCMA immunoconjugate utilised in the treatment of relapsed or refractory multiple myeloma.[105]

1.6. Bioconjugation strategies

Despite all the advantages of ADCs, multiple factors limit the success of progression from clinical trials to becoming approved medicines. One significant limitation is bioconjugation, the process of linking drugs to antibody. As aforementioned, Mylotarg demonstrates one of the current limitations, plasma stability. Lack of stability *in vivo* can lead to various issues regarding toxicity and therapeutic efficacy. Another limitation of current bioconjugation is the yielding of heterogeneous mixtures. Heterogeneous mixtures exhibit relatively narrow therapeutic index in addition to aggregation, unreliable toxicity, and inadequate pharmacokinetic profiles.[106]

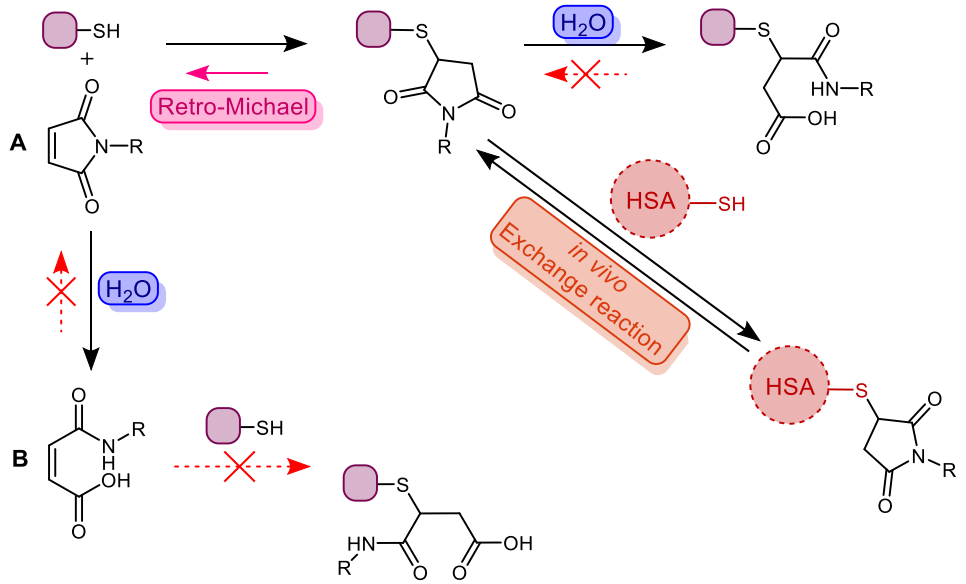
1.6.1. Native antibody modification

Bioconjugation is associated with multiple aspects of the ADC from DARs to the stability of the linker. Ideally the bioconjugation method should aim to produce a homogeneous ADC with an optimal DAR. Too low (typically <2.0) and the agent will lack efficacy due to minimal levels of cytotoxic drug, limiting the activity of the ADC;[107,108] alternatively too high of a DAR (typically >4.0) means the ADC is prone to high plasma clearance rates and aggregation due to increased hydrophobicity. It can also result in an increased likelihood to induce toxicity.[109,110] However, the most recent ADCs, Trodelvy and Enhertu, utilise DARs of 7.6 and 8.0 respectively,[111] as they incorporate more hydrophilic payload-linker systems.[112] Thus, whilst generally a DAR range of 2.0-4.0 is desired, payload choice is also a key aspect of ADC design. Early ADCs pursued lysine residues as points of bioconjugation, due to the advantages of biologically robust amide bonds being generated, limited antibody processing steps and solvent accessibility.[81] Due to the natural abundance of lysine residues, over 80 present in a typical IgG1 antibody, a statistical mixture of DARs is obtained ranging from 0-8.[109] In an attempt to reduce the heterogenicity, cysteine residues have become a focus, due to their low natural abundance (1.7%).[113]

1.6.1.1. Cysteine bioconjugation

The target within cysteine residues is a highly nucleophilic thiol, bearing a pK_a of ~8 in aqueous buffers.[114] This relatively low pK_a results in the formation of approximately 10% thiolate ion (S^-) under physiological conditions (pH = ~7). The thiol moieties are also considered 'soft' nucleophiles according to Hard and Soft Acids and Bases (HSAB) theory this is due to the polarizable nature of the sulfur atom.[115,116] Cysteine residues on IgG1 antibodies are predominantly found as disulfide bonds. Reduction of four accessible interchain disulfide bonds, exposes target thiol groups for bioconjugation. The reduction is typically carried out using reagents such as tris(2-carboxyethyl)phosphine (TCEP) or dithiothreitol (DTT). The four interchain disulfide bonds are distal from the antigen binding and Fc regions of the antibody, this is important in ensuring that antigen-binding activity and Fc-mediated antibody recycling is unaffected. This approach has been utilised for the development of an approved ADC for Hodgkin's lymphoma, Adcetris. It was found that the therapeutic index significantly increases when the drug load is reduced from eight to four, due to reduction in plasma clearance and aggregation.[108] Notably, modification of the disulfide bridges in this manner also leads to the loss of these interchain linkages which thereby may impact the overall integrity of the antibody.[117,48]

Maleimides are widely used to target cysteine residues and can be found in the majority of current ADCs. Maleimides consist of a highly reactive alkene group, which acts as an electrophile and undergoes thio-Michael addition to yield a thioether bond.[118]

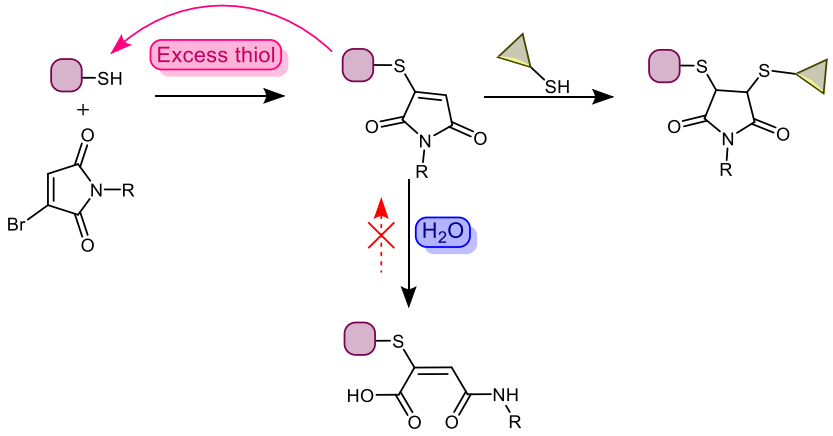


Scheme 2. General cysteine modification with traditional maleimide. Maleimides are able to undergo thio-Michael addition to yield a thioether bond. The modified maleimide then hydrolyses to yield a stable thiol modification. The modified maleimide is also able to reverse *via* a retro-Michael type reaction leading to limited in vivo stability.

The root of the reactivity is in part due to two characteristics: 1) the strain of the 5-member ring 2) the presence of two electron withdrawing carbonyls alpha to the alkene.[119,120] Unlike other modifying agents, maleimides do not change the overall net charge of the biomolecule.[121] The optimal pH for maleimide-thiol conjugation is pH 7, and at this pH the rate of reactivity is 1000-fold greater than that of the reaction with amines.[122] The range of chemoselectivity sits between pH 6.5-7.5, as above pH 8.5, competitive aza-Michael additions start to occur.[118] Typically these maleimide ADCs were observed to have limited *in vivo* stability due to the thiosuccinimide linkage undergoing a retro-Michael type reaction, resulting in premature drug loss, leading to off-site toxicities and a loss in efficacy.[123] Under aqueous conditions maleimide rings are susceptible to hydrolysis, which increases with pH. Hydrolysis of the thiosuccinimide ring to succinamic acid results in a stable conjugate that is no longer able to undergo retro-Michael type reactions or *in vivo* exchange. The yield of the stable hydrolysed product can be limited up to 50%, this is due to the competing retro-Michael pathway which cannot be negated.[124] Hydrolysis may provide stability, however, this is only possible if it occurs after the initial conjugation. If hydrolysis occurs to the unreacted maleimide (Scheme 2- **A**) it results in the formation of maleic amide (Scheme 2- **B**) which is unreactive to thiols.

To combat this, maleimide reagents have been designed with electron withdrawing *N*-substituents to increase the rate of thiosuccinimide hydrolysis, which removes the ability of the ADC linker to undergo retro-Michael addition increasing *in vivo* stability and limiting payload loss (Scheme 2).[125]

Baker, Caddick and co-workers reported on next generation maleimides (NGMs),[126] these reagents incorporate a leaving group (e.g. bromine) across the maleimide double bond. The leaving group allows for an addition-elimination sequence to take place which retains the double bond. The retained bond provides two points for chemical attachment, as it is then able to undergo a Michael addition with a stoichiometric measure of second thiol. These reagents were found to be extremely rapid and efficient, treatment of the L111C mutant of the Grb2 adaptor protein with 1 equivalent of *N*-methylbromomaleimide gave 100% conversion after 1h at 0 °C, pH 8.[127] These conjugates are reversible; as when exposed large concentrations of free thiol, conjugate addition occurs at the double bond which is followed by cleavage *via* a retro-Michael pathway. Thiomaleimides are also susceptible to hydrolysis as with traditional maleimides, and the resulting maleamic acid is stable to thiol and retro-Michael addition. The hydrolysis is accelerated by incubating the thiomaleimide conjugate at an elevated pH and temperature (e.g., pH 8, 37 °C). [128,129]



Scheme 3. General cysteine modification with next-generation bromomaleimides. These reagents are able to undergo a two-step dual modification reaction and are stabilised *via* a hydrolysis step.

As with maleimides, other Michael acceptors are a staple for cysteine bioconjugation. The rate-determining step for such reagents is the initial addition of thiol to the unsaturated alkene or alkyne of Michael acceptors. Activation of these reagents *via* the incorporation of electron withdrawing groups increases the electrophilicity of the Michael acceptor which can therefore improve the conjugation efficiency.[130] An example of these activated Michael acceptors was demonstrated by Bernardes and co-workers in 2016, who introduced an electron withdrawing carbonyl group adjacent to an acrylic reagent to generate rapid and irreversible

carbonylacrylics.[131] When tested on reduced trastuzumab, the carbonylacrylic reagent was found to yield conjugates with eight modifications.[132]

Similar to Michael acceptors, other electrophilic reagents can also be activated by the addition of EWGs such as heteroaromatic rings. These electron deficient rings react with cysteine *via* the nucleophilic aromatic substitution (S_NAr) pathway. Reaction at the site of a leaving group (LG) results in the formation of an anionic σ-complex, better known as a Meisenheimer complex, followed by elimination of the LG and re-aromatisation.[133] A recent piece of work by Martin *et al*. revealed a rapid heteroaromatic sulfone-cysteine conjugation system. A library of heteroaromatic sulfones were compiled of which the benzothiazole rings exhibited the fastest cysteine reactivity, and these reagents form stable adducts.[134] These electron deficient heteroaromatic rings are yet to be tested on complex systems such as antibodies.

Transition metals are not commonly used for bioconjugation, due to the need for stringent reaction conditions such as aqueous media, low temperature, and mild pH. In 2015 however, Buchwald *et al*. developed a novel organometallic palladium (II) reagent through a careful choice of ligand. Palladium complexes containing chloride, bromide and iodide counterions were all found to produce the desired product while also being stable to ambient conditions, allowing for long term storage (4 °C for over four months under air). This class of reagent is also synthetically attractive as it is possible to exploit features of the payloads chemical structure, e.g. synthesis of a vandetanib palladium complex through the insertion of vandetanib via the aryl bromide present in the structure. Partially reduced trastuzumab was successfully modified with the vandetanib palladium complex, to generate an ADC with a 4.4 DAR in just 30 minutes.[135]

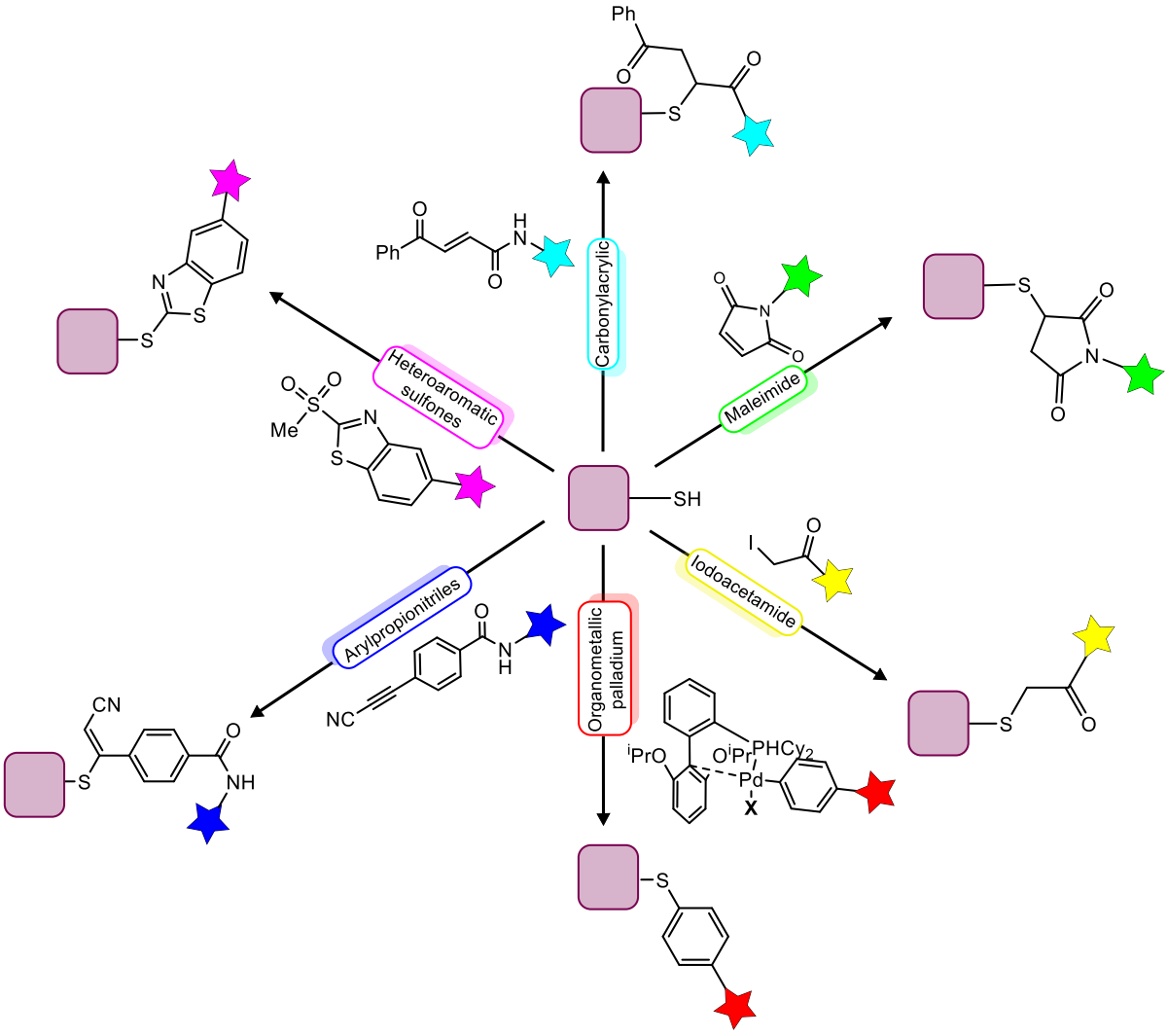


Figure 15. Summary of a selection of cysteine modification techniques. Light blue = carbonylacrylic reagents, green = maleimide reagents, yellow = Iodoacetamide reagents, red = organometallic palladium reagents, dark blue = Aryl propionitrile reagents, pink = heteroaromatic sulfones.

Alkyl halides are also found as common thiol-reactive compounds with iodoacetamide being the most popular. Iodoacetamide undergoes a nucleophilic attack to irreversibly form a thioether bond. Unlike maleimides which exhibit impressive cysteine selectivity, iodoacetamides have been found to cross-react with other amino acids such as methionine, tyrosine, lysine, glutamic acid or histidine.[136,137] This introduction provides an insight into a select few cysteine modifications, however, alternate techniques and reagents can be found (Figure 15) and more are still being developed to date.

1.6.2. Disulfide bridging reagents

Another way to target cysteine residues is to bridge two cysteines, thus re-forming the disulfide linkage. Solvent accessible disulfides often do not play a role in the biological activity of

proteins, but instead provide stability to the protein, making them potential targets for modification.[138,139,140] Re-bridging reagents have the potential to preserve the innate stability provided by disulfide bridges, resulting in homogeneous ADCs with a defined DAR (Figure 16). Disulfide bonds are easily reduced with reducing agents such as DTT or TCEP, exposing two free electrophilic thiol groups for modification.[141,142]

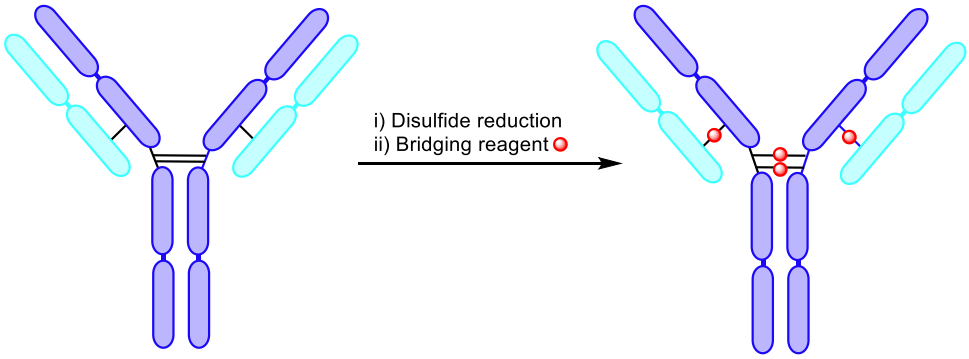
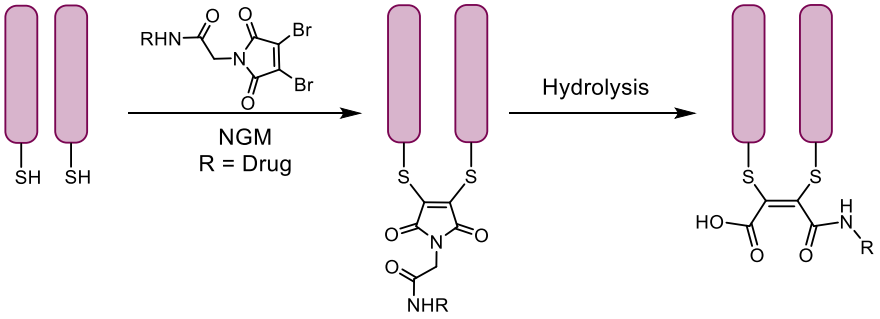


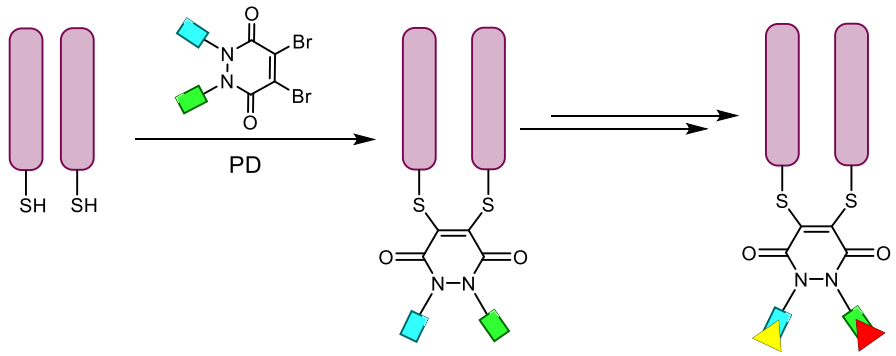
Figure 16. Ideal disulfide bridging reaction on a mAb.

Following on from the development of NGMs for single cysteine modification, Baker and Caddick *et al*. proposed two leaving groups, on the maleimide double bond, could enable two consecutive addition-elimination sequences with a reduced disulfide, to generate a rigid two-carbon bridge. Use of NGMs as re-bridging reagents was exemplified by the reaction of reduced trastuzumab with dithiophenolmaleimide (DTPM)-MMAE, which implements thiophenol as the leaving group. After reacting for 1 h at 37 °C, the resulting conjugate was subjected to hydrolysis *via* incubation at pH 8.4 for 72 h to generate a thiol stable ADC, with a DAR of 3.89. The hydrolysed ADCs offer antiproliferative effects on tumour cells both *in vitro* and *in vivo*.[143] Post-conjugation hydrolysis is vital for robust serum stability, however this can be a lengthy process (72 h). In order to optimise this step, Baker and Chudasama *et al*. developed a dibromomaleimide (DBM) platform for accelerated post-conjugation hydrolysis. It was found that introduction of electron-withdrawing C-2 linkers onto the maleimide nitrogen optimised post-conjugation hydrolysis to just 1 h at pH 8.5.[144]



Scheme 4. NGMs for disulfide bridging.

Chudasama, Baker *et al*. also reported on the use of dibromopyridazinediones (PDs) as adaptable disulfide re-bridging reagents.[145] PDs contain four potential attachment points, two of which can be bioorthogonal handles, and can be exploited for dual-labelling of mAbs such as trastuzumab (Scheme 5).[145] Like NGMs, PDs also undergo a dual two addition-elimination reaction yielding a rigid two-carbon bridge, which is more resistant to hydrolysis and thiol cleavage reactions. An example use of these tetra-functional reagents was demonstrated by Robinson *et al*. who combined the anti-cancer drug doxorubicin (Dox) and a cyanine-based fluorophore (sulfo-Cy5) to trastuzumab with a PD linker. PDs enable efficient re-bridging of disulfide bonds in mAbs, whilst also benefiting from immediate increased stability without the need for post-conjugation hydrolysis.[146]



Scheme 5. Dibromopyridazinediones (PDs) for disulfide bridging.

Despite the advancements made with cysteine bridging reagents, there were still a few drawbacks which limit the efficacy. The bridging reagents can result in disulfide scrambling (Figure 17) to form half-antibodies and as a result induces protein unfolding and aggregation.[147,148,144] To overcome this Bahou *et al*. developed an *in situ* bioconjugation protocol combating the drawback of disulfide scrambling. Use of dibromopyridazinediones in an optimised protocol allows for the modification of native trastuzumab with over 90% homogeneity whilst avoiding disulfide scrambling.[149] Low reaction temperature (4 °C) and an optimised number of equivalents (20 eq.) allowed for the production of a trastuzumab conjugate with a DAR of 4.

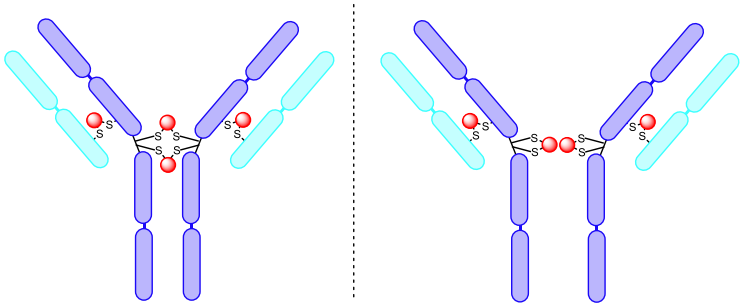
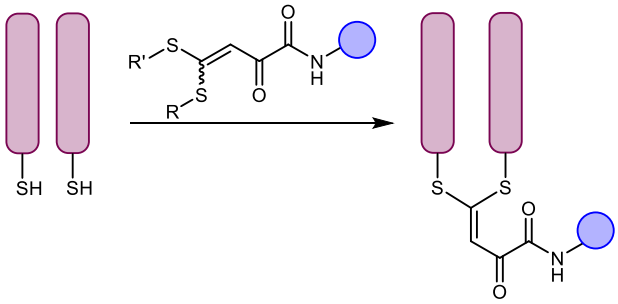


Figure 17. Disulfide scrambling resulting in half-antibody formation (Right), successful disulfide conjugation (Left).

A recent technique for disulfide bridging was reported by Poulsen *et al*. in 2022. The group previously described STEF reagents as vinylogous thioesters, able to undergo selective and irreversible conjugation to lysine residues. The conjugation occurs *via* a transfer reaction, mediated by a proximal cysteine residue. Further development of these STEF scaffolds, through the introduction of an additional sulfide substituent, reduces the electrophilicity and therefore reactivity towards lysine residues, whilst retaining reactivity towards cysteines. The newly developed compounds, classified as oxSTEF reagents, are consequently able to undergo two distinct thiol-exchange reactions. First examples of disulfide bridging were shown on the macrocyclic peptides octreotide and lanreotide, where selective disulfide re-bridging was achieved.



Scheme 6. oxSTEF reagents for disulfide bridging

The oxSTEF reagents were therefore implemented in more complex systems, the first of which was the 32-residue hormone, salmon calcitonin. Despite the positive results prior, the reaction with salmon calcitonin resulted in the formation of two major products, thought to be due to the *N*-terminal cysteine present in the salmon calcitonin disulfide bond. The *N*-terminal cysteine was able to under cyclisation with the oxSTEF reagents, forming a 5-member ring. oxSTEF reagents may therefore be incompatible for bridging with systems containing *N*-terminal cysteine residues, however, could also provide exciting new scope for use in site-selective *N*-terminal cysteine modification. Human growth hormone (hGH) and egg-white lysozyme both contain internal disulfide bonds which replicated the selective disulfide re-bridging when treated with oxSTEF reagents. Trastuzumab was selected as the system with the greatest complexity, where re-bridging experiments revealed selective labelling of the reduced native mAb, however further work is yet to be done regarding the intricacies of the modification e.g. DAR and conversion.[150]

Despite the advantages of disulfide re-bridging strategies regarding selectivity there are still drawbacks within these systems. The strategy is yet to be fully validated in clinical trials and whilst the disulfide bond it is retained chemically, the chemical stability of each bridge is variable and therefore can impact the overall tertiary structure when placed in comparison with the native disulfide bond. This introduction provides insight into a select few bridging reagents however, a variety of other reagents are also available (Figure 18).

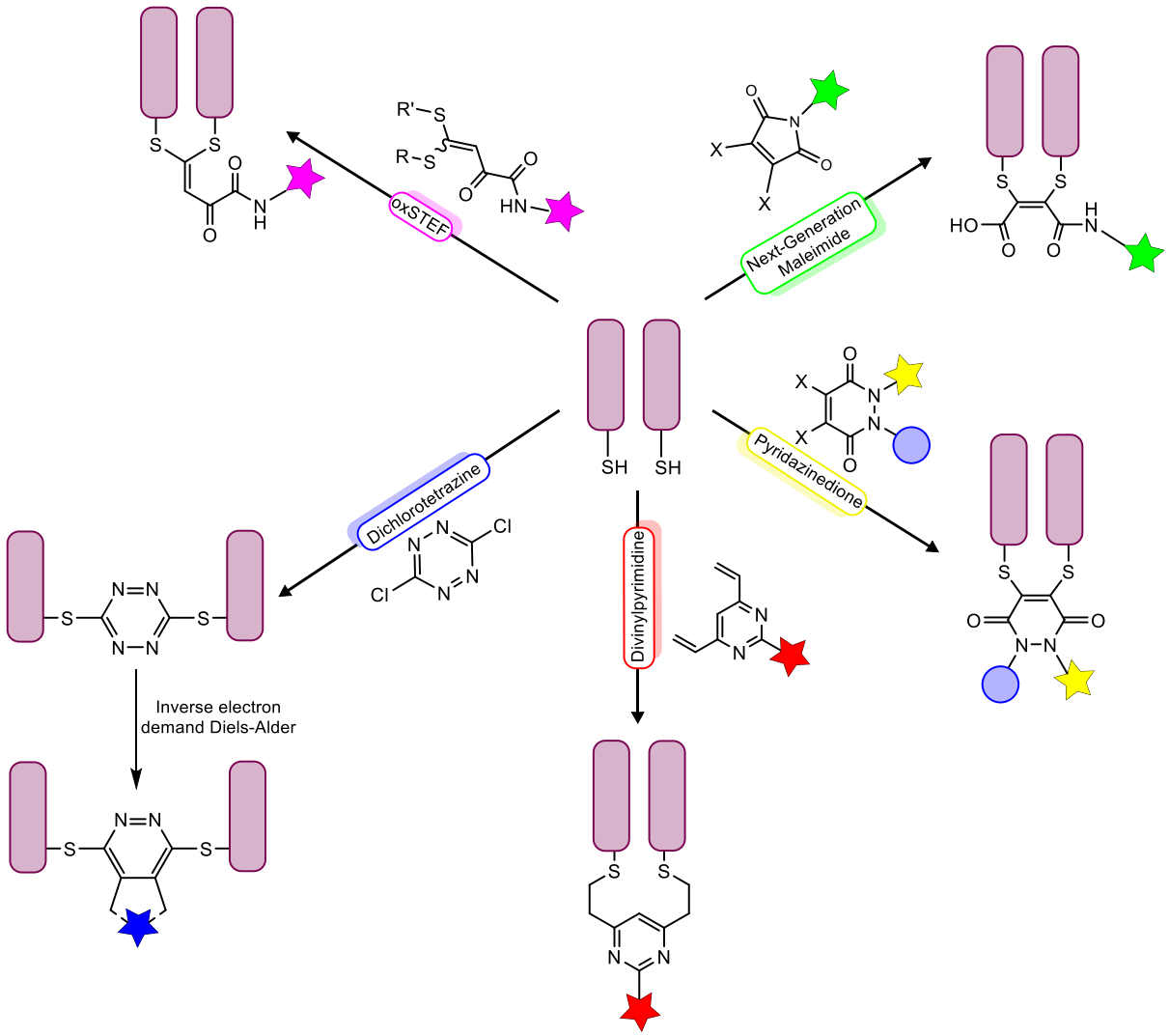


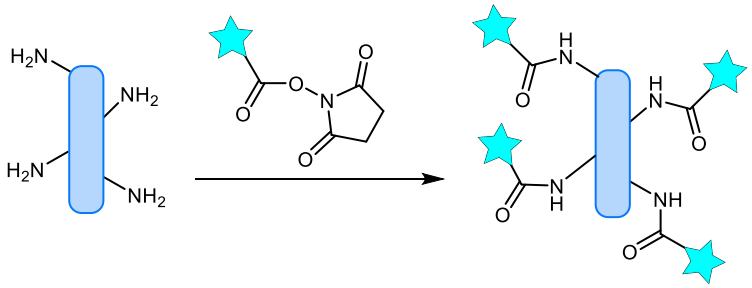
Figure 18. Summary of a selection of cysteine bridging approaches. Pink = oxSTEF reagents, green = Nest-generation maleimides (NGMs), yellow = pyridazinediones, red = divinyl pyrimidines, blue = dichlorotetrazines.

1.6.3. Lysine bioconjugation

Lysine is one of the most abundant amino acids in proteins, containing a linear four-carbon chain and a nucleophilic primary ε-amine group at the terminus with a pK_a of around 10.5.[151] Alongside arginine and histidine residues, lysine contributes to the overall net positive charge of proteins due to the dominant ionic character of the side chain ε-amine, which favours formation of the positively charged ammonium groups at physiological pH. This characteristic often results in high levels of solvent exposure, making them easily accessible for chemical modification.[118] On average, IgGs contain a high abundance of lysine residues (~80) of which almost half are available for modification.[152] With such a large number of accessible lysines, it is a challenge to obtain selective, controlled and consistent modification. This leads to the formation of heterogeneous mixtures where antibodies present a different number of

attachments at different sites.[153] Heterogeneity in ADCs can cause issues relating to sub-optimal pharmacokinetics, efficacy, and safety profiles.[154] Common lysine modification reagents include acylating agents, which are able to generate stable amide bonds. Reagents used for lysine conjugation include NHS esters, sulfonyl fluorides, iminoboronates, diazonium salts, isothiocyanates and many more.[155] Typically, basic pH conditions (pH > 8) are required to achieve the best yield due to ε-amino groups having a high pK_a (pK_a ~10.5). The degree of modifications can be regulated by limiting the reagent dose, for example, equivalents or sub-stoichiometric reaction conditions which can lead to a decrease in modification yields.[156]

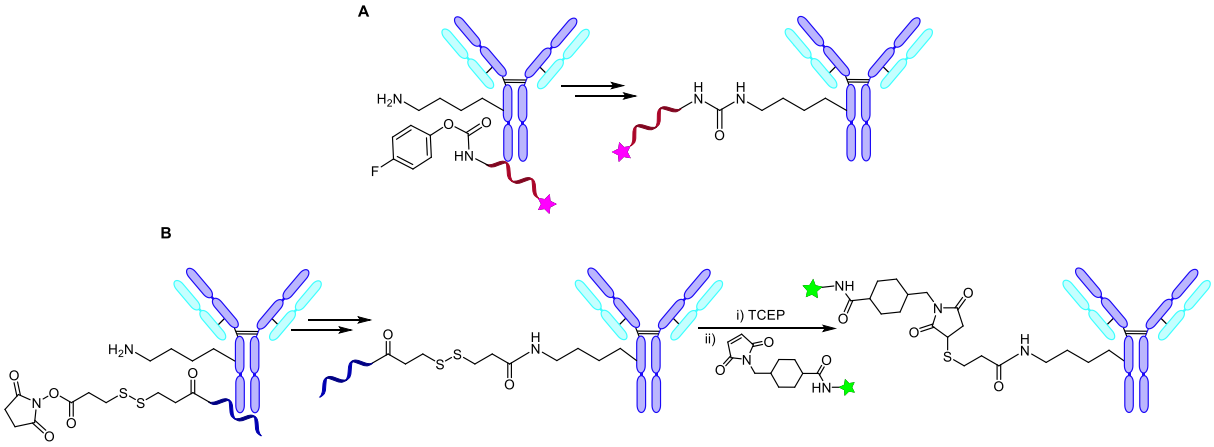
N-hydrosuccinimide (NHS) esters were first introduced in 1963 by Anderson *et al*. for the formation of peptide bonds.[157] They have become popular acylation reagents which undergo rapid reaction with free amines to form a stable amide bond.[158] NHS esters contain an activated acyl group and are typically formed by the reaction of a parent carboxylic acid with NHS in the presence of a coupling agent e.g. carbodiimides or DCC.[122] NHS esters have been employed in the production of several clinical ADCs, namely Mylotarg, Kadcyla and Besponsa. The lysine modification on each of these ADCs results in heterogeneous products, as mentioned prior this can lead to multiple issues, exemplified by the temporary withdrawal of Mylotarg due to off-site toxicity and difficulty identifying a dose regimen.[54] NHS esters are also known to suffer from chemoselectivity issues, and exhibit cross-reactivity with other accessible nucleophilic residues (e.g. Cys, Tyr, Ser).[158]



Scheme 7. NHS esters for Lys conjugation.

Site selectivity remains a challenge in the production of lysine conjugates, though strategies have been developed in attempts to overcome previous limitations through novel chemistry and methodologies. One of these strategies was developed by the biotech company Ajinomoto, who reported a new method of affinity peptide mediated regio-divergent functionalisation, termed AJICAP. The group identified three peptides through a phage display system panning against human IgG1. One of the peptides, comprised of 17 amino acids, presented a high affinity for the Fc region of the human IgG (K_d= 9 nM). Co-crystal structures specified the distance between the affinity peptide and lysine residues, providing the

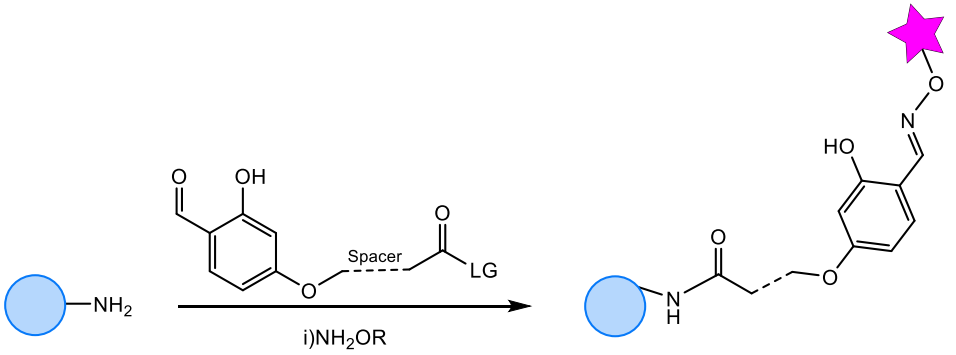
information required for the linker design. Dithiobis(succinimidylpropionate), a useful tool in protein labelling, provided an appropriate length linker (12Å) with an incorporated NHS ester and was therefore employed in the modification of the affinity peptide. Once the antibody affinity peptide conjugate (AAPC) is obtained, the NHS ester is placed in proximity to a lysine residue with which it reacts to form the amide bond. As NHS esters are prone to hydrolysis at neutral pH, the method is conducted at an acidic pH (5.5), this feature is also important for preventing unspecific conjugation. After 1 h at 22 °C, ADCs with an average DAR of 1.9 were generated.[159]



Scheme 8. Examples of affinity-guided strategies for site-selective lysine modification (**A**) Xiao *et al*. protocol (**B**) AJICAP protocol developed by Mendelsohn *et al*.

Another peptide affinity strategy was developed by Xiao *et al*. who reported a ncAA-encoded affinity peptide-based protocol for lysine modification on IgG. The group synthesised ncAA-containing peptides based on the B domain of protein A (FB protein) derived from *Staphylococcus aureus*.[160] The FB protein was found to be a small, stable protein that only binds to IgGs in their Fc regions (C_H2−C_H3 junction), avoiding the potential disruption of the CDR or Fc-receptor binding sites. An electrophilic ncAA-containing FB domain was generated *via* the site-specific incorporation of 4-fluorophenyl carbamate lysine (FPheK), an ncAA which is able to react with a proximal lysine to form a stable urea cross-link. Six FB protein mutants were developed of which only the E25FPheK variant provided the desired FB−heavy chain complex after 48 h at 37 °C, pH 8.5. Fluorescein-labelled trastuzumab conjugates prepared with the reported strategy demonstrated 98% conversion. These conjugates also retained their HER2 binding ability as fluorescent imaging showed fluorescence only in HER2+ cells.[161]

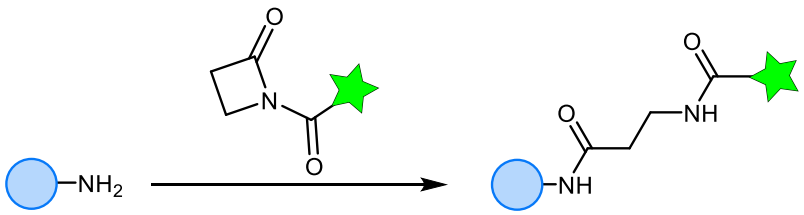
A more recent strategy developed by Rai *et al*. utilises a novel 'linchpin' platform, which unlike the AJICAP protocol, employs a reversible chemical tether. The technique was first applied to the modification of histidine residues in 2018 before the application was propagated to lysine modification.[162] The method relies on the design of the linchpin compound which is based upon a F_K^1-spacer-F_K^2 structure. Two chemoselective handles present differing characteristics: 1) F_K^1 undergoes a rapid, unselective, and reversible reaction with all solvent accessible lysines regulating the micro-concentration of the other handle near a proximal Lys residue 2) F_K^2 reacts irreversibly with lysine at a relatively slow rate which governs the site selectivity. The reactivity of F_K^2 also ensures that the design of the linker regulates the site of conjugation as opposed to the order of lysine reactivity, as only lysine residues suitably distanced from each other will undergo labelling. Labelling experiments on RNase-A identified site-selective modification of K37, considered the most reactive lysine residue, yielding a bioconjugate with high homogeneity. The linchpin platform was then applied to trastuzumab, conjugation at 25 °C, pH 7.0 for 16 h, followed by the functionalisation with emtansine (25 °C, pH 7.0) resulted in the synthesis of trastuzumab emtansine. Unambiguous analysis of the synthesised ADC identified the sites-of-conjugation as K169 (LC) and K395 (HC). The synthesised ADC also showed marginally improved growth inhibition of HER2+ SKBR-3 cells (45%) in comparison to Kadcyla (38%), whilst also showing no significant impact on HER2- cells.[163]



Scheme 9. Linchpin platform for Lys conjugation.

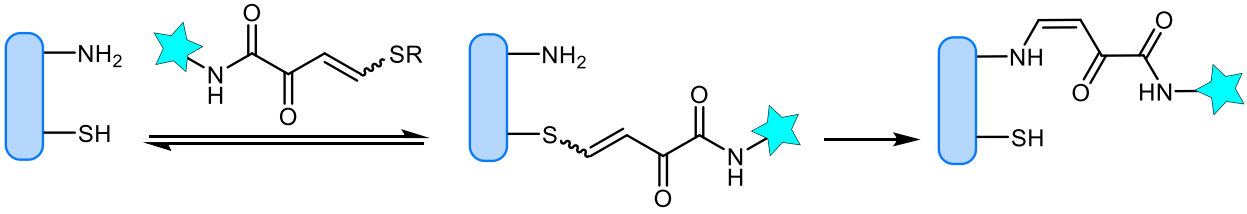
Rader *et al*. developed a strategy targeting a natural lysine residue with distinctive chemical properties. The humanized anti-hapten mAb h38C2, possesses a lysine residue at the bottom of an 11-Å deep hydrophobic pocket. This lysine residue is more nucleophilic at physiological pH, due to its abnormally low pK_a (~6.0) compared to solvent-exposed lysines (~10.5).[164] The group utilised β-lactam functionalised linkers, which undergo nucleophilic ring opening to form an irreversible amide bond with the buried h38C2 lysine residue. In order to harness the lysine reactivity of h38C2 for payload attachment and enable tumour targeting, dual-variable-domains (DVDs) were developed, which combined the Fv of h38C2 with trastuzumab. The

DVD species was then modified with a MMAF functionalised β-lactam for 4 h at 22 °C, pH 7.4 to generate an anti-HER2 DVD-ADC with a DAR of 2.0.



Scheme 10. β-lactam functionalised linkers for Lys conjugation.

A recent and highly relevant methodology for lysine modification was developed by Poulsen *et al*. who synthesised and demonstrated the use of STEF-reagents, as electrophilic warheads in chemical biology. A key aspect of these reagents is the reactivity towards both thiol and amine groups *via* an addition-elimination mechanism. Thiol STEF-reagents were hypothesised to reversibly react with cysteine thiol residues, forming a mixture of *cis* and *trans* isomers, followed by subsequent irreversible reaction with amine nucleophiles to exclusively generate the *cis* isomer (Scheme 11). This was successfully demonstrated with small molecule experiments alongside lysate systems such as MCF7. Iodoacetamide capping strongly reduced STEF labelling, supporting the idea of intermediate cysteine conjugation sites prior to irreversible amine conjugation. To investigate the application of STEF reagents in bioconjugation the methodology was applied on HSA. Here, two major sites of modification were detected based on the number of mass spectral scans: K414 and K199. However, both of these residues are located in shallow hydrophobic grooves and are not proximal to the free thiol (Cys 34). The cumulative data indicates that whilst cysteine proximity tends to determine the major lysine reactivity, other target motifs exist, providing possible scope for reagent tuning.[165] These reagents are yet to be proven on antibody systems which may provide additional challenges.



Scheme 11. STEF reagents for Lys conjugation.

Despite all the benefits of current lysine modification approaches, there are limited strategies that provide both convenient and efficient homogeneous lysine conjugates. Only one of the last five ADCs that have been approved has utilised lysine conjugation, whilst the remaining four have employed homogeneous cysteine conjugation.

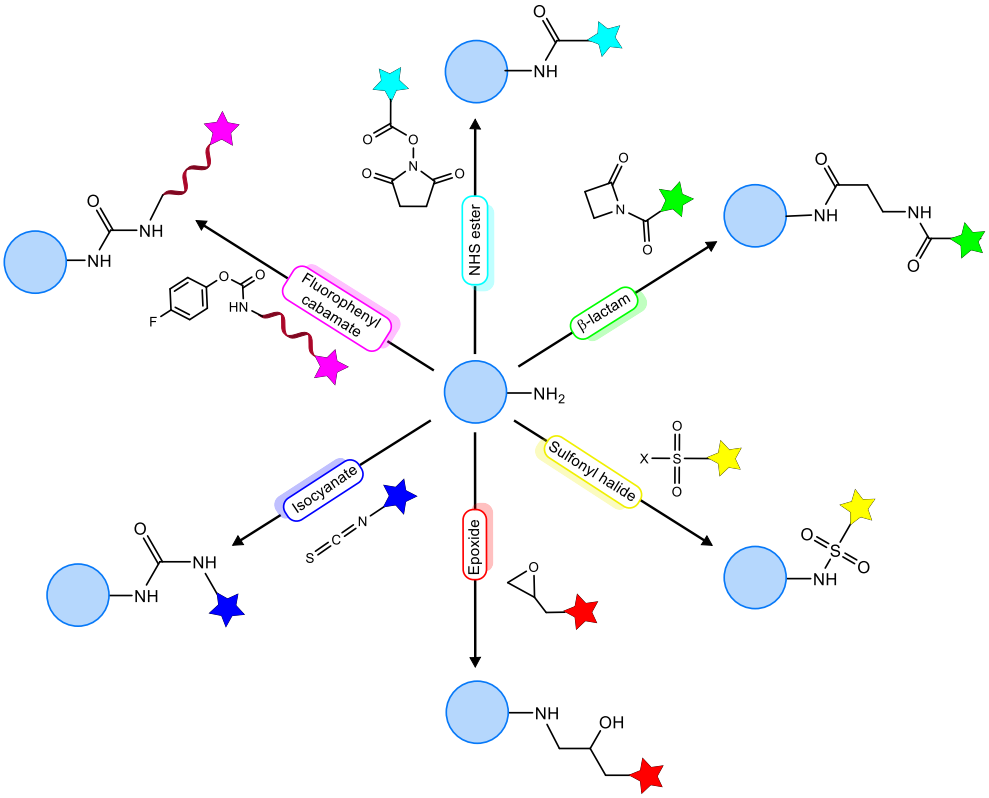
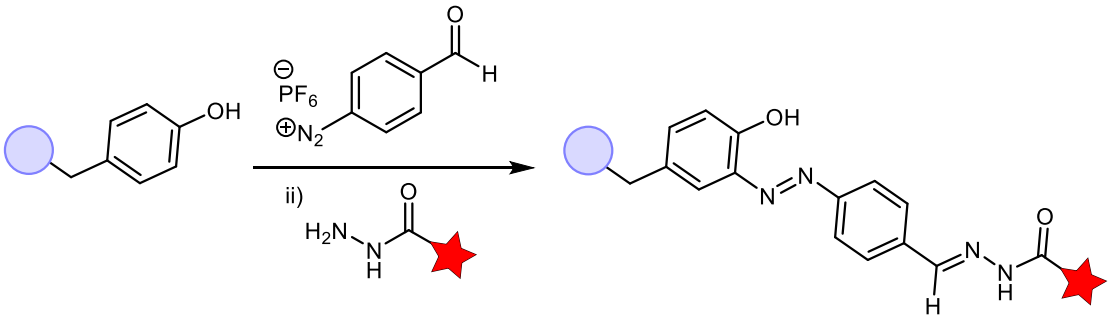


Figure 19. Summary of a selection of lysine modification techniques. Light blue = NHS esters, green = β-Lactam, yellow = sulfonyl halide reagents, red = epoxide reagents, dark blue = isocyanate reagents, pink = Fluorophenyl carbamate (peptide affinity tags).

Whilst there are an ample number of lysine modification reagents for wide-use protein bioconjugation, successful site-selective methods investigated for ADC synthesis are limited, as seen by the scarcity of recently approved lysine-conjugated ADCs. This introduction provides an insight into a select few lysine modification techniques, more can be found (Figure 19) and more are still being developed to this day.

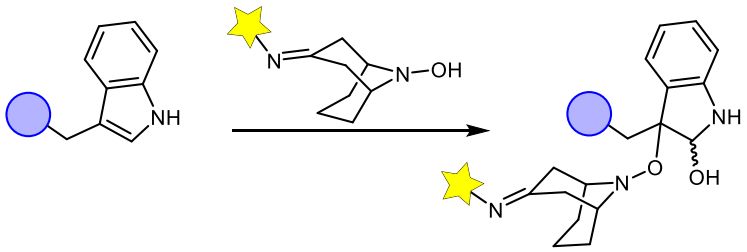
1.6.4. Emerging residues

Cysteine and lysine residues have been at the forefront of native antibody conjugation due to their solvent accessibility and physiochemical properties. As of recently, other residues have come to light, providing possible advantages over traditional target residues. One of these residues is tyrosine which consists of an amphiphilic phenol group. The amphiphilic nature results in a residue that is often partially buried in the surface of the proteins with the hydroxyl group exposed, tyrosine is scarcely expressed on protein surfaces making it a good target for site-selective modification. [166] In 2012 Barbas *et al*. first described the bench stable 4-formylbenzene diazonium hexafluorophosphate (FBDP) reagent which was reported for the selective modification of trastuzumab. FBDP generates conjugates with a free aldehyde handle which was then modified with biotin–oxyamine to give an average DAR of 1.8 biotin tags.[167]



Scheme 12. Diazonium salts for tyrosine modification.

Tryptophan is another residue that is rarely targeted, as of 2021 only two reported methods achieved functionalisation. Tryptophan is the least abundant (~1%)[168] and surface-exposed proteinogenic amino acid. In 2016 Kanai *et al*. reported the selective labelling of tryptophan in aqueous medium with the stable organoradical 9-azabicyclo [3.3.1]nonane-3-one-N-oxyl (keto-ABNO). The reactive species was thought to be an oxoammonium cation, derived from keto-ABNO *via* acid promoted disproportionation and oxidation by NO_x, which is generated from $NaNO_2$ and acetic acid. The conjugation then proceeds through the nucleophilic attack of the enamine nucleophile in the indole ring of tryptophan, followed by the attack of water to generate the resulting imine. This strategy was applied to the synthesis of an antibody−small molecule conjugate, the commercially available anti Aβ1−16 antibody (6E10) was conjugated with a keto-ABNO fluorescein derivative.[169,166] Recent data from the same group has also shown the modification of tryptophan in the production of trastuzumab-gold nanocluster conjugates. These conjugates, exhibit binding affinities with HER2 antigen that are 4-fold smaller (ED_{50} = 1.65 nM) when compared to the naked mAb (ED_{50} = 6.65 nM), this is an improvement over an analogous Lys conjugate which exhibited a binding affinity 18-fold lower (ED_{50} = 0.37 nM).[170]



Scheme 13. ABNO mediated C-H activation for tryptophan modification.

1.7. Engineered antibody modification

Antibody engineering has become a well-developed practice in the development of ADCs with high levels of homogeneity. Engineered antibodies are generated by recombinantly inserting additional surface-exposed residues that can then be targeted. These insertions are designed to ensure no structural or functional impairments occur, which may impact antigen binding, internalisation or structural integrity of the native mAb.[171] A well-established example of antibody engineering was described in 2008 by Junutula *et al*. who developed THIOMABs. These antibodies consist of two engineered cysteine residues on the mAb trastuzumab. The cysteine residues were carefully placed within the light and heavy chains of the fab constant domains to ensure minimal impact on the antigen binding or in Fc-mediated effector functions of the antibody. Preliminary studies revealed the engineered cysteine residues were capped as mixed disulfides, with cysteine or glutathione molecules, which presumably formed during production. Prior to use, the THIOMABs must therefore undergo partial reduction to expose the engineered cysteine thiols.[96] Use of THIOMABs allowed for the production of ADCs with an average DAR of 1.9 and more than 90% homogeneity. Unnatural ammino acids (UAA) are also implemented for site-selective modification, the incorporation of which, is carried out during the transcription stage of antibody synthesis.[172] One such example, communicated by Schultz *et al*., incorporated *p*-acetylphenylalanine into an anti-CXCR4 antibody to produce DAR 2 ADCs *via* oxime ligation.[173]

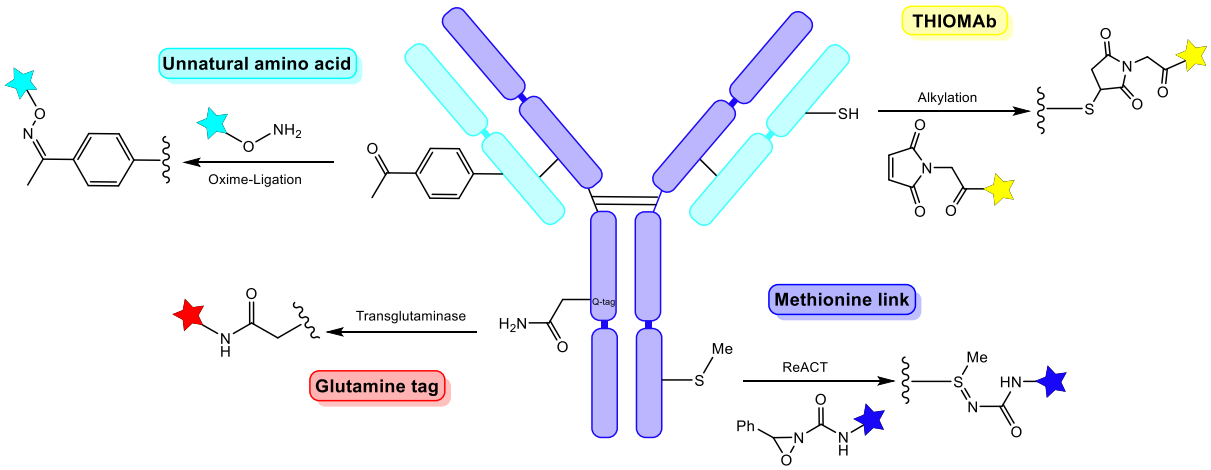
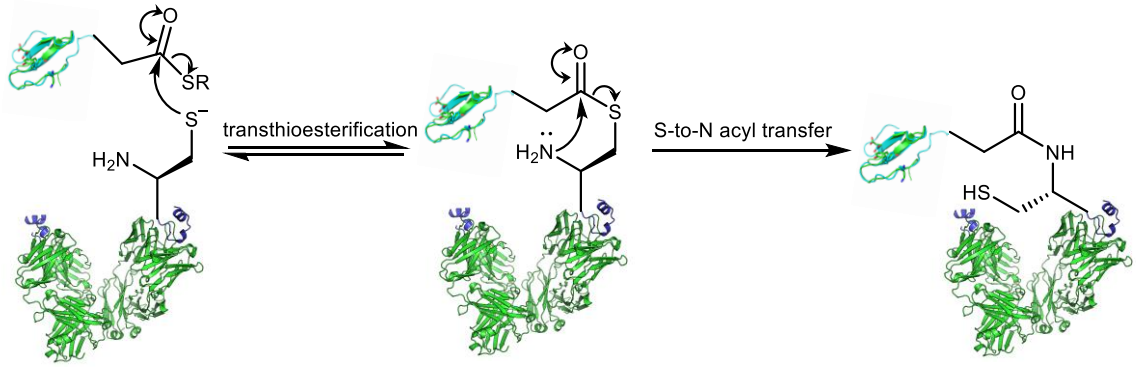


Figure 20. General figure highlighting typical methods for engineered antibody modification; yellow = introduction of cysteine residues (THIOMABs), dark blue = methionine insertion, light blue = unnatural amino acid insertion, red = introduction of a glutamine tag for transglutaminase functionalisation.

Whilst engineered ADCs have been shown to outperform heterogeneous systems, they are often limited by production costs due to the added complexity and unique expression systems, which are yet to be proven in clinical trials.

1.8. Native chemical ligation

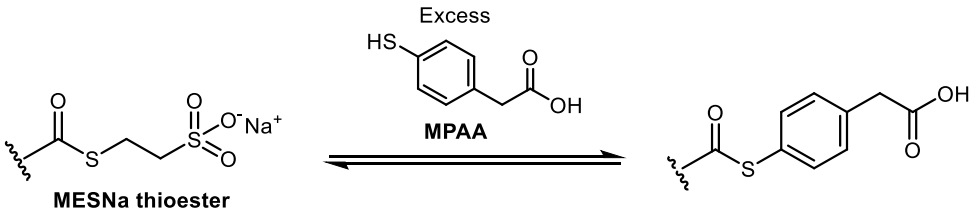


Scheme 14. General mechanism of NCL. Step 1 is a reversible transthioesterification followed by step 2; a spontaneous intramolecular *S*-to-*N* acyl transfer.

Native chemical ligation (NCL) is a process developed in 1994 by Kent *et al.* as a methodology for peptide ligation and has since become a staple in protein synthesis and chemical biology strategies. NCL allows for the direct and chemoselective reaction between two unprotected peptides. The reaction takes place between an *N*-terminal cysteine and a *C*-terminal thioester.[174] This strategy has been applied in multiple protocols from the synthesis of a HIV-1 protease dimer[175] to human interleukin-8.[176] Mechanistically, there are two steps: 1) the initial step is a reversible transthioesterification rection in the presence of an added thiol, which is

normally considered to be the rate-determining step (RDS), 2). The second step involves a spontaneous intramolecular *S*-to-*N* acyl transfer, resulting in the formation of a native peptide bond. This second step is irreversible under the reaction conditions, as the transfer is thermodynamically favourable due to the formation of the stable amide bond relative to the thioester bonds (Scheme 14).[177]

The proximity between the amino group and thioester moieties provides an additional driving factor, which is referred to as 'entropic activation'.[178] The proximity increases the local concentrations of the nucleophile and electrophilic carbon centre of the thioester. A broad range of ring transitions states (TS) sizes had been studied, with the most common and fastest of being the 5-membered TS.[179] This is expected as an increase in the TS ring size will result in a reduction of the 'entropic activation'.

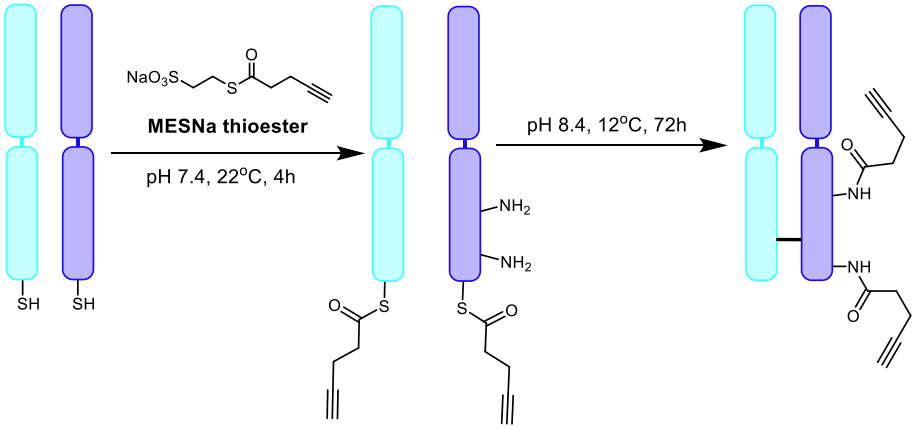


Scheme 15. Thiol-thioester exchange with aryl thiols (MPAA).

The rate of transfer could be altered by a modification of parameters. The amino acid involved plays a role in the kinetics, as Dawson *et al.* reported that the Gly thioester resulted in the fastest transfer (<4 hr), whereas amino acids such as Val, Ile, and Pro all presented less favourable conditions and therefore slower intermolecular ligation (>24 hr) due to the added steric bulk.[176] This was overcome with the use of selenoesters in place of thioesters. This comes down to the pK_a values of selenocysteine (5.2–5.6) and cysteine (8.2), with the lower pK_a of SeH resulting in a superior leaving group and therefore faster ligation.[180] Tuning the pH can also result in an accelerated rate, an example coming from Kent *et al.* who reported an increase in pH from 6.1 to 7.2 for ligation at the Glu-Cys site, which limited the formation of side product by 9% (from 11% to 2%).[181] The rate of ligation can also be influenced by the type of thiol utilised. Alkyl thioesters, e.g. 2-Mercaptoethanesulfonic acid (MESNa) thioester, can undergo 90% ligation post 24 hours, however aryl thioesters, such as 4-Mercaptophenylacetic acid (MPAA) thioester, only require 30 min.[182] However, alkyl thioesters are usually utilised due to their more appealing synthesis and handling, and they are converted into aryl thioesters *in situ via* addition of excess aryl thiol resulting in thiol exchange[183] (Scheme 15).

1.9. Cysteine to lysine transfer (CLT)

With the aim of investigating a new bioconjugation strategy, which would provide a selective homogeneous conjugate that is stable in circulation, Forte *et al.* developed a novel methodology known as Cysteine-to-Lysine transfer (CLT) (Scheme 16). The strategy achieves selectivity by utilising reduced cystine residues as ligating 'hooks' which then deliver an acyl group to proximal lysine residues through *S*-to-*N* transfer. This aims to provide the selectivity of cysteine modification with the stability of lysine modification.

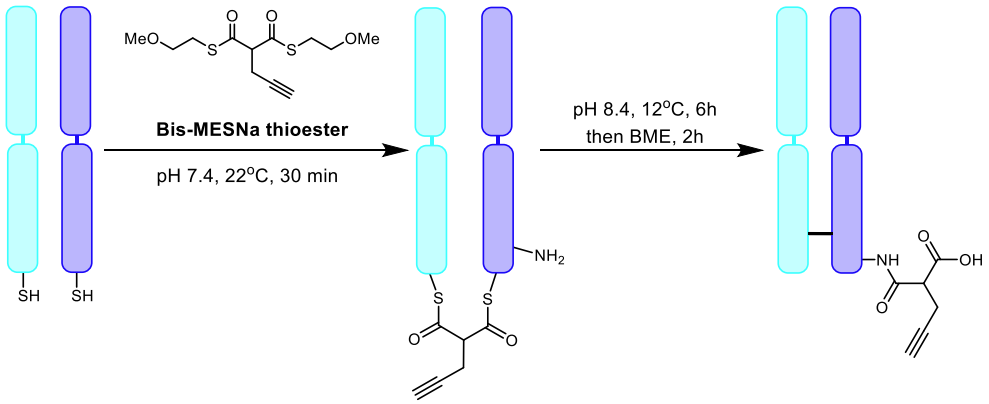


Scheme 16. CLT strategy applied with MESNa thioester. Step 1 involves reversible transthioesterification followed by step 2; a spontaneous *S*-to-*N* transfer.

The methodology was developed on the Fab fragment of trastuzumab, a HER2-targeting breast cancer drug.[184] The Fab fragment consists of a single disulfide bridge connecting a light-chain (LC) and heavy-chain (HC).[185] The crystal structure of Fab revealed the presence of 26 lysines, three (K136, K221, K225) of which were neighbouring the disulfide bridge on the HC, whilst no proximal lysines were located on the LC.[186] Thioesters were selected as chemical candidates due to their reactivity profile. Thioesters are approximately 100-fold more reactive towards amines relative to the corresponding oxoesters. Thioesters also undergo hydrolysis 20% slower than the corresponding oxoesters, and both of these characteristics demonstrate the advantage of thioesters over oxoesters for CLT.[187] Upon testing for the selectivity of aryl thioesters, it was uncovered that the reaction of this species with native Fab resulted in unselective addition to lysine residues. Due to this background conjugation, the reactivity of the thioesters had to be modulated, and alkyl thioesters were therefore chosen for further work. The MESNa alkyne thioester (Scheme 16) was found to be unreactive towards native Fab, even at high equivalents, whilst retaining reactivity towards reduced Fab, showing its cysteine selectivity over lysine residues. An alkyne handle allowed for post-conjugation functionalisation *via* click chemistry. The optimised conditions for the transthioesterification were found to be 100 eq. pH 7.4 at 22 °C over 4 h. The initial conditions for the *S*-to-*N* acyl transfer were as follows: pH 7.4 at 22 °C over 72 h which provided significant transfer. By

increasing the temperature to 37 °C and the pH to 8.0, the transfer reaction was pushed to completion, however the rate of competing hydrolysis also increased. In order to minimise hydrolysis and ensure completion, a high pH, low temperature approach was taken at pH 8.4 12 °C for 72 h. The cysteine residues were found to spontaneously re-oxidise, reforming the disulfide bridge and providing a traceless conjugation which maintains native mAb stability. The final conditions resulted in an average alkyne-to-antibody ratio (AAR) of 1.5 with 75-80% transfer efficiency. Three possible lysine residues were confirmed all of which reside on the HC: K136, K221, K225.

From this, it was envisaged that a bis-thioester would provide added rigidity due to formation of a bridged system which may help further control regioselectivity. Here 2-methoxyethanethiol was utilised to incorporate water solubility whilst limiting polarity, and therefore purification issues associated with the use of MESNa. Reaction with reduced Fab resulted in a bridged conjugate in just 30 min at 22 °C, and the bis-thioester also retained cysteine selectivity and was unreactive towards native Fab.



Scheme 17. CLT strategy applied with an alkyl bis-thioester. Step 1 involves dual transthioesterification followed by spontaneous *S*-to-*N* transfer and thioester cleavage.

The increased rate of transthioesterification was associated to the increased electron-withdrawing effects of the additional carbonyl. Treatment with BME for 2 hours revealed traces of mono-thioesters which had formed due to competing hydrolysis, thus the final AAR was revealed as 0.8 (Scheme 17).[186]

1.10. Nitriles

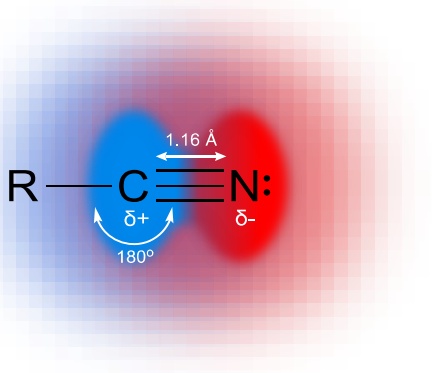


Figure 21. Schematic representation of the nitrile (cyano) group.

Nitriles are versatile functional groups which have found application within over 70 approved drugs in the pharmaceutical field e.g. Cimetidine,[188] Vildagliptin and Letrozole.[189] Also known as cyano groups, nitriles provide multiple advantageous properties which can improve compound pharmacokinetics and pharmacodynamics.[190] Nitrile bonds are linear in geometry connecting a nitrogen atom to a sp hybridised carbon atom through a triple bond. The sp carbon is electrophilic in nature due to a significant electron deficiency, as a result of the superior electronegativity of the nitrogen atom and high dipole moment present in the triple bond. The nitrogen lone pair enables the moiety to act as a hydrogen bond acceptor group, whilst the triple bond π-system enables lipophilic interactions. [189,191]

The electrophilic carbon centre is able to reversibly react with nucleophilic residues within protein systems, such as serine or cysteine to form imidates or thioimidates respectively.[192] As the reactivity of amino acid side chains is dependent on its pK_a, cysteine (~8) residues are favoured in comparison to serine (~13). Catalytic cysteine residues are found as part of a catalytic dyads such as Cys/His. These dyads promote the formation of the highly nucleophilic thiolate species which in turn increases their reactivity. The formation of the covalent nitrile adduct is a two-step process where the formation of ionic pair (Cys^-/His^+) is followed by the nucleophilic attack on the nitrile carbon by the cysteine thiolate. Alongside the nucleophilic attack, a proton transfer occurs from the protonated histidine to the newly formed thioimidate in order to neutralise and therefore stabilise the nitrogen atom. Computational studies with a nitrile based cruzain inhibitor indicate the nucleophilic attack and proton transfer occur in a concerted manner.[193]

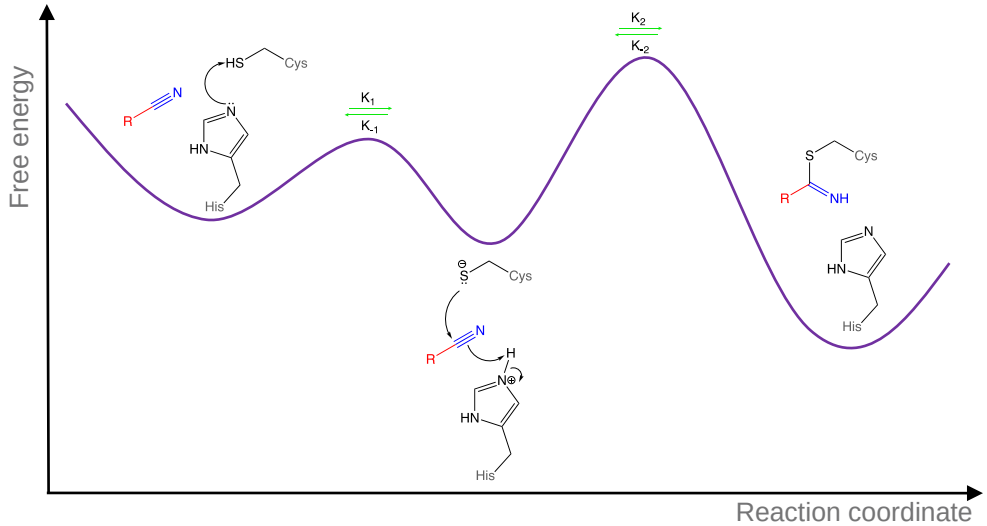


Figure 22. Schematic representation of a catalytic Cys/His dyad involving a reversible covalent inhibitor containing a nitrile warhead. The first step represents the acid–base reaction, followed by the noncovalent bound state and lastly the covalent bound state. Adapted from reference.[193]

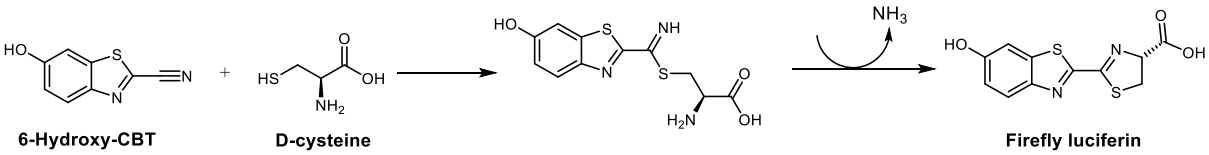
Whilst cysteine residues are commonly targeted regardless of the catalytic ability, serine is almost exclusively nucleophilic at the protein active site, resulting from a charge relay mechanism in a catalytic triad. One such triad found in serine proteases e.g. chymotrypsin, is formed of three coordinated amino acids: serine, histidine and aspartate.[194] The triad works in a similar fashion to the catalytic dyad, the difference comes from aspartate which hydrogen bonds to histidine. Aspartate helps to orientate histidine and neutralise the developing charge on the imidazole ring,[195] enabling histidine to act as a powerful base for the deprotonation of serine.[196,197]

In 2004, Altmann *et al.* reported the use of purine nitriles as potent inhibitors of the cysteine protease cathepsin K. The nitrile moiety reacts reversibly with Cys25 to form a thioimidate bond. Substitution of the nitrile with an alkyl group such as methyl was found to be detrimental to the activity of the inhibitor, thereby indicating the importance of the cyano group.[198] In 2007, Altmann *et al.* then further developed nitrile inhibitors of cathepsin K to find more active and selective reagents. The focus then shifts from purine to pyrimidine nitriles where introduction of electron-withdrawing groups (EWGs) at position 5 resulted in highly potent inhibitors. This electron-withdrawing effect reduces the electron density surrounding the carbon in the nitrile therefore increasing its electrophilicity and reactivity.[198]

Whilst the traditional formation of thioimidates is reversible, Lin *et al*. reported on the development of internal cysteine tagging, through a bespoke peptide tag, which was able to produce stable thioimidates with 2-cyanobenzothiazole (CBT) reagents. The small protein ubiquitin (Ub) was selected as the core due to its innate lack of cysteine residues, Ub was modified with select peptide sequences attached to the C-terminus. As a negative control, a sequence consisting of a cysteine flanked by two glycine residues was produced, which as expected was unable to form a stable adduct with 6-amino CBT. Unbiased investigation into sequence environment was carried out *via* a randomised 11-mer peptide phage library, X_5-Cys-X_5 where X is any natural amino acid. This revealed the greatest conversion to be with the peptide sequence named CX10R7. The reasoning behind the success of CX10R7 was thought to be the stabilising effects by polar and charged residues neighbouring the central cysteine residue.[199]

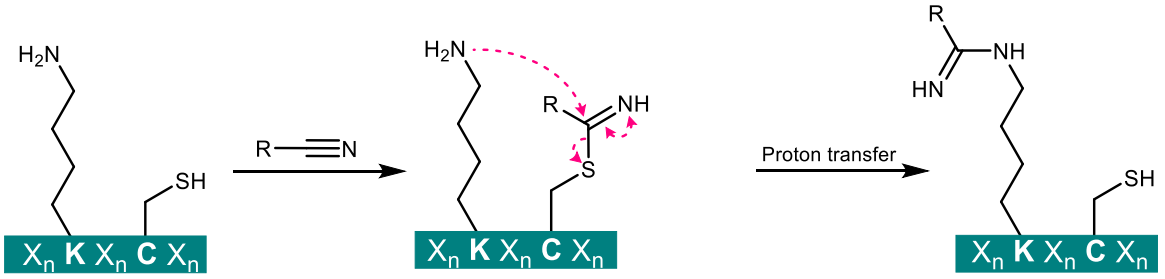
Nitriles are also known to undergo a 2-step condensation reaction with 1,2-aminothiols. Using this, Field *et al.* developed the synthesis of firefly luciferin, an oxidisable substrate which is implemented in firefly bioluminescence. 6-hydroxy-CBT is combined with *D*-cysteine. The cysteine thiol reversibly attacks the nitrile carbon to form a thioimidate intermediate. This is then attacked irreversibly by the adjacent amino group to form a tetrahedral intermediate, followed by the loss of ammonia to form the 4,5-dihydro thiazole ring present in luciferin.[200]

The loss of ammonia provides a driving force due to the entropic gain provided (Scheme 18). This is contrasting to NCL where the thiol transfers the acyl group to the amine.



Scheme 18. Synthesis of firefly luciferin.

Following from this, Rao *et al.* investigated an *in vivo* reaction between 6-amino-2-CBT and cysteine at physiological pH and room temperature which proceeds with a 2nd order rate constant of 9.2 M^{-1} s^{-1}. This was applied to develop a method to selectively label bioengineered *N*-terminal cysteine peptides *in vivo*.[201] Since then, multiple groups have utilised this reaction to develop labelling strategies. For example, Lin *et al.* have reported the use of the CBT motif as a cysteine labelling agent through a genetically encodable peptide tag.[199] In 2018, Bertozzi *et al.* exploited this nitrile reactivity to develop a peptide motif which reacts with cyanobenzothiazole *via* an internal cysteine−lysine relay. They developed a small genetically encodable peptide tag with the ability to ligate CBT derivatives selectively and irreversibly. The final peptide tag, named CBTag was 11 amino acids long and was able to undergo an irreversible intramolecular *S*-to-*N* transfer.[202] CBTag incorporates a lysine in close proximity to cysteine, which is optimal positioning for transfer (Scheme 19). The efficiency of transfer was suboptimal with a yield of 17-38% over 48 h, depending on the conditions. Thus, there is still room for development regarding the rate and efficiency of transfer. Foden *et al*. have also shown this type of transfer reactivity in thiol catalysed coupling of peptide nitriles with amines.[203]



Scheme 19. Reaction between CBTag and CBTNAc resulting in a thioimidate intermediate. This is followed by an irreversible intramolecular *S*-to-*N* transfer.

Studies have been done to investigate the stability of thioimidates. Schmir *et al.* looked at the hydrolysis of four thioimidate esters at multiple pH values ranging from 0-10.[204] They report an unusual feature regarding the pH-rate profiles for **A** and **B** (Figure 22) - there is rapid disappearance of thioimidate at alkaline pH. A parallel increase in hydroxide ion and resultant drop in protonated thioimidate should result in a consequential drop in the rate of

hydrolysis.[205,206] A more likely pathway is a hydroxide ion catalysed elimination of thioimidate to the nitrile and thiol.

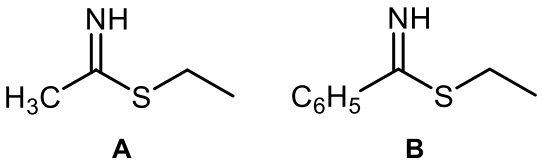
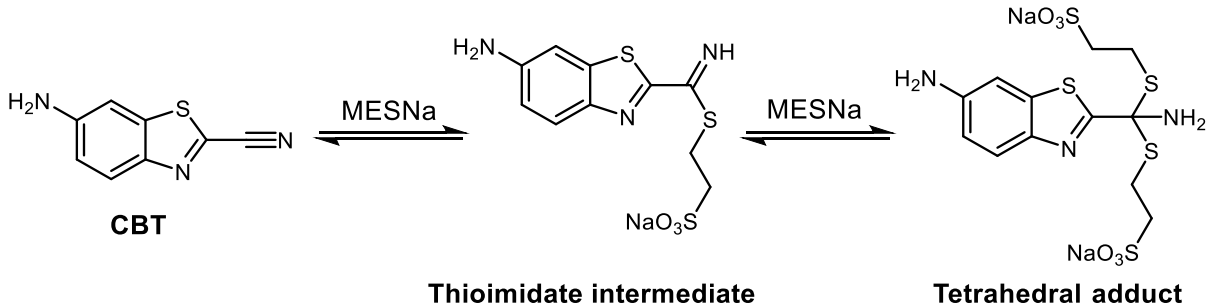


Figure 22. Thioimidate ester structures.

Pfizer has investigated the proton-catalysed synthesis of amidines from thioimidates. Therein, they discuss that the rate of amidine formation is likely to be a function of the concentrations of protonated thioimidate and free amine.[207] At high pH, the neutral thioimidate fragments to the nitrile and thiol, as seen by Schmir *et al.*[204] On the other hand, too low of a pH results in loss of free nucleophilic amine. The effect of this retro-nitrile formation was shown to be insignificant when aromatic amines were utilised, this effect however was dominant with aliphatic amines. Use of an acetate buffer allowed for amidine formation using methylamine and ethylamine, while providing decent yields (>80%) despite a longer reaction time of 72 h.[207]

Most recently, Bayley *et al*. reported the use of CBT compounds for peptide modification as before however, they also report on further reactivity of thioimidates. Therein, they utilised ESI-MS and protein nanoreactors to detect the presence of tetrahedral adducts which form upon reaction of two thiol ligands (MESNa) to CBT. Both additions of thiol are reversible, and the monoadduct proved to be approximately 80-fold longer-lived than the corresponding tetrahedral adduct (Scheme 20).[208] The multi-thiol reactivity exhibited by CBT molecules opens avenues for their use as small-molecule walkers: molecules that are able to undergo translational motion on a track containing multiple thiols.[209] In addition to multi-thiol walker systems, the CBT molecules also show promise for CBT-based walker systems with alternating cysteine and lysine footholds through the formation of amidine groups via a tetrahedral intermediate.

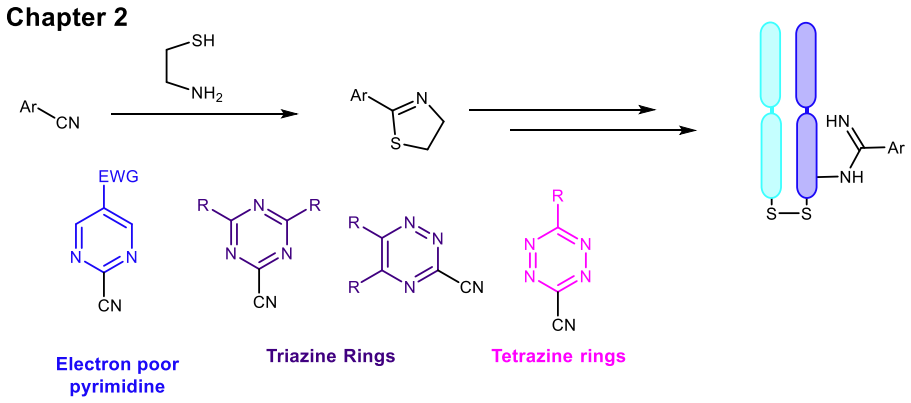


Scheme 20. Reversible addition of thiol ligands (MESNa) to CBT to form a tetrahedral adduct.

Aims & Objectives

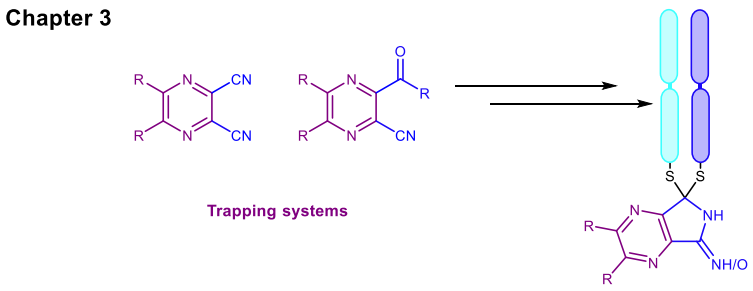
The main aim of this work is to explore the use of the nitrile functional group for the site-selective modification of proteins. The project will aim to identify nitrile reagents with novel designs and chemistry with a focus on cysteine-to-lysine transfer (CLT) and disulfide bridging reagents.

Chapter 2 centres around the development of a new generation of CLT reagents. Previous work investigated thioesters as appropriate cysteine-to-lysine transfer reagents, exhibiting desirable cysteine selectivity and acyl transfer reactivity for NCL. Despite this initial success of developing site-selectively modified lysine conjugates, there were issues regarding hydrolysis of starting materials and intermediates, leading to incomplete transfer (75–80%), with suboptimal reaction times (>7h). Early work done by Bertozzi et al. established nitriles as a functional group able to undergo *S* to *N* transfer. Electron-deficient mono-nitriles were therefore proposed as a novel type of CLT reagent, incorporating nitrogen-containing heteroaromatic as the core structure. It was hypothesised that the electrophilic nitrile would rapidly react with the free thiols on a reduced Fab antibody fragment, sequentially transferring onto proximal lysine residues. The design of these reagents will look to explore the scales of reactivity respective to CLT and aim to introduce accessible points of functionality.



Scheme 20. Chapter 2 showing the development of a new generation of CLT reagents.

Chapter 3 focuses on the design and synthesis of a novel one-carbon disulfide re-bridging system based on an electron-deficient bis-nitrile motif. It was hypothesised that incorporation of an electrophilic centre adjacent to the nitrile group would act as a 'trapping' system, allowing for successful disulfide re-bridging. The dynamic nature of these reagents also allows for the exploration of additional uses for these reagents as well as investigation into the novel chemistry behind the trapping system. These systems will then be applied on different proteins (e.g. trastuzumab Fab, full trastuzumab and thioredoxin) alongside small molecule systems (mono- and bis-thiols).



Scheme 21. Chapter 3 illustrating a novel one-carbon disulfide re-bridging system.

The aim of Chapter 4 was to develop the understanding of a novel potential bioorthogonal click reaction between an ethanedithiol moiety and an electron-deficient mono-nitrile (e.g. 2-cyano pyrimidine). This nitrile bis-thiol (NBT) reaction presents remarkable selectivity which could potentially be implemented in a wide variety of work from peptide tags to protein pulldown systems.

The greater part of this work is carried out on the Fab fragment of HER2+ targeting breast cancer drug trastuzumab. This is due to the desirable platform of a single solvent accessible disulfide bond, surrounded by proximal lysine residues. This structure allows for the protein to be utilised for the work described in both Chapters 2 and 3.

Results and Discussion

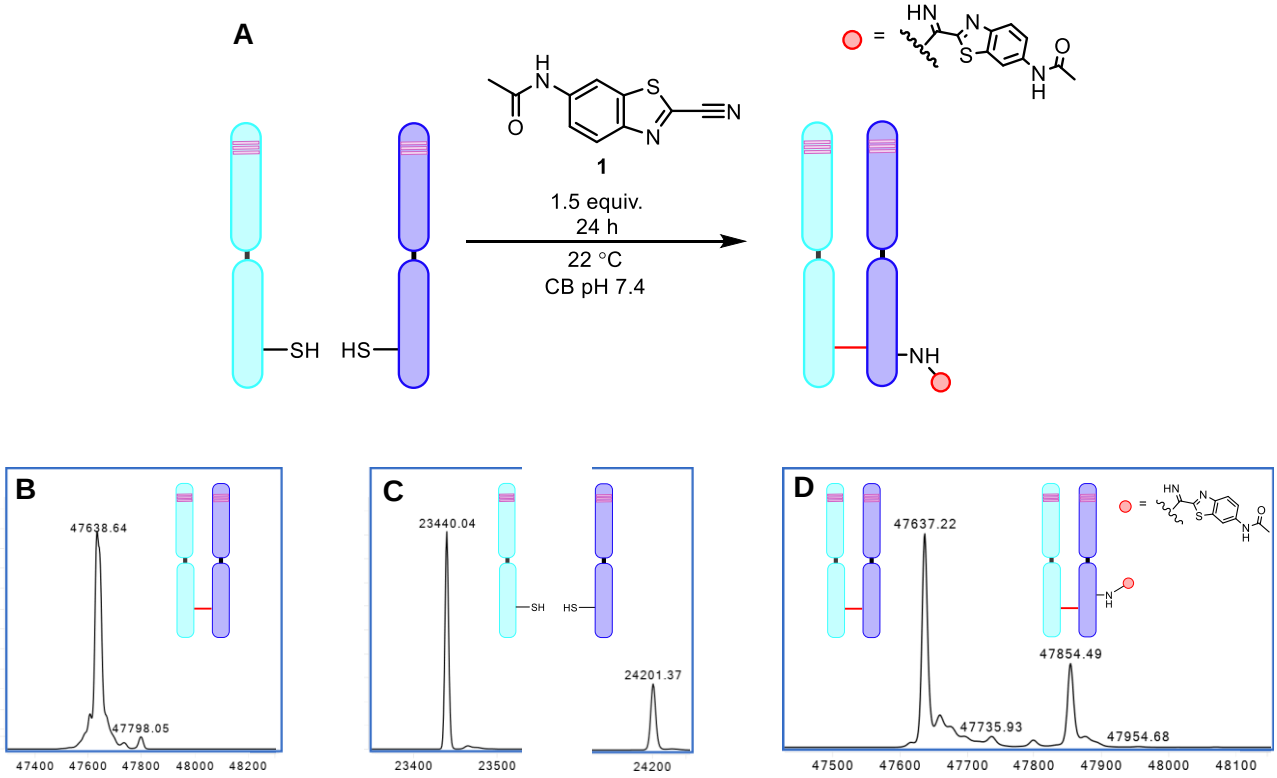
2. Chapter 2: Nitrile CLT

This chapter looks to develop a new generation of cysteine-to-lysine transfer (CLT) reagents. The reagents aim to provide controlled addition to the Fab fragments whilst retaining a high degree of stability. A variety of cores are to be investigated, attempting to generate a homogeneous sample. Different conditions are trialled in order to gain a greater understanding behind the chemistry of these reactions and also the applicability toward protein modification. The reagent design aims to also incorporate useful functional handles in order to accommodate a variety of payloads. The new generation of CLT reagents will also look to surpass the prior generation of CLT reagents, thioesters, with regards to both yield of modification (>80%) and reaction time (<10h).

2.1.1. Previous work

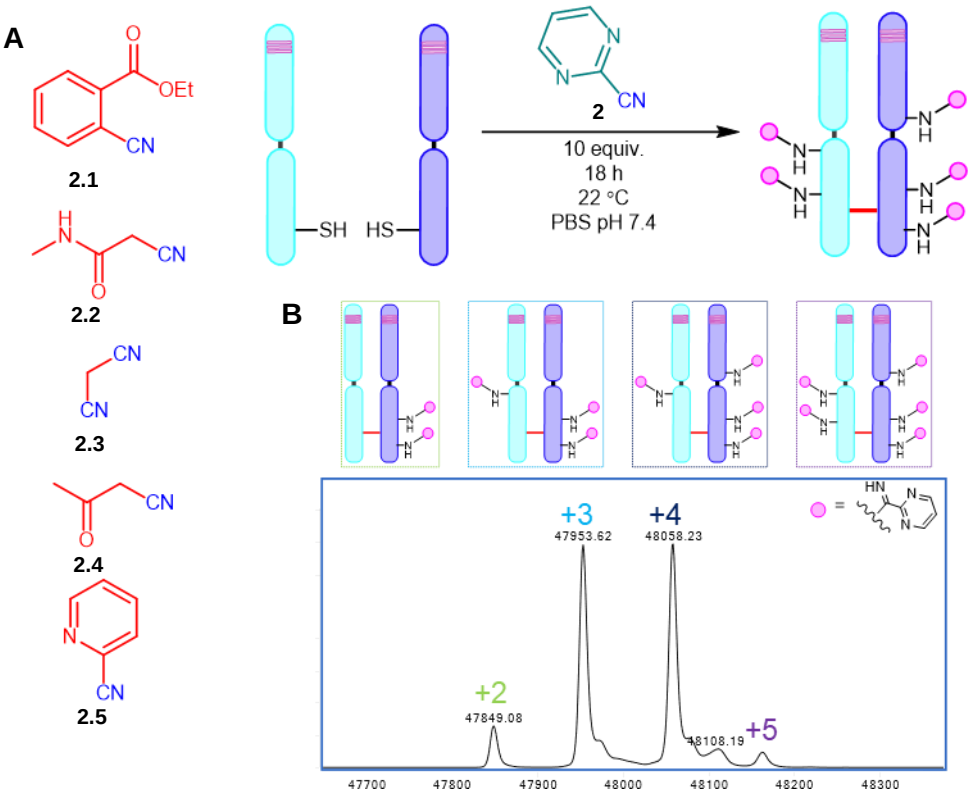
Initial work done by Dr. Nafsika Forte identified cyanobenzothiazole (CBT) (**1**) as the first nitrile CLT reagent. The reagent was found to present exquisite selectivity for cysteine thiols over other solvent accessible residues present on trastuzumab Fab. Low equivalent (1.5 equiv.) reactions over an extended period of time were shown to yield modifications of one addition. The reaction mixture did not show any HC/LC upon mass spectrometry, indicating the disulfide

bridge had been regenerated, this was the first indication that the modification was not present on the cysteine thiols but had indeed transferred (Scheme 22).



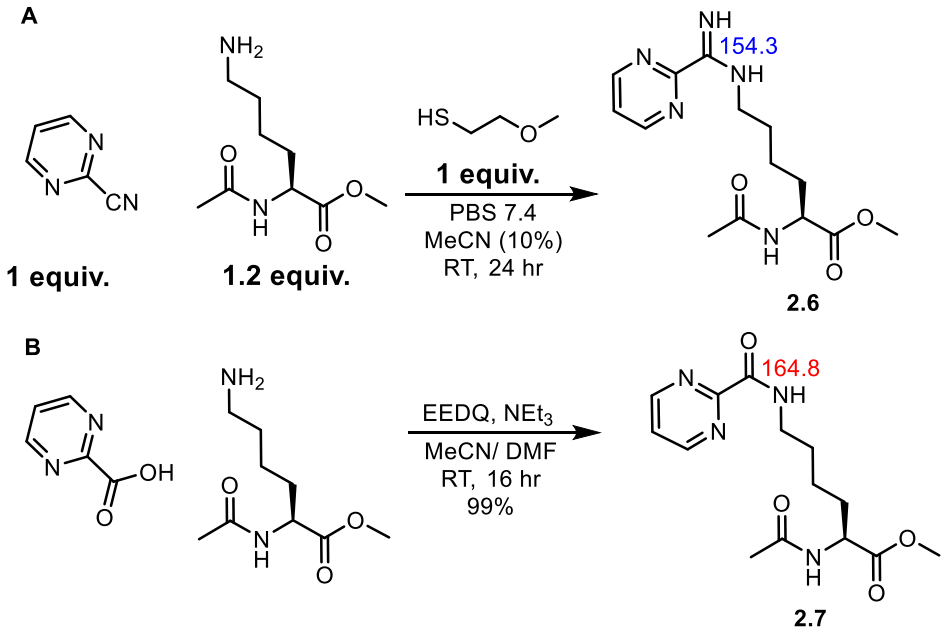
Scheme 22. (A) CLT reaction of CBT **1** reactions on reduced Fab, 1.5 equiv., 24 h, 22 ºC, pH 7.4. (**B**) LCMS of non-reduced Fab. (**C**) LCMS of Reduced Fab (**D**) LCMS of CBT **1** reaction conjugate, Mass of native Fab = 47638 Da. Mass of native LC = 23440 Da. Mass of native HC = 24201 Da. Expected mass of addition = 217 Da.

With this result utilising CBT, a small library of nitrile compounds were trialled from which 2-cyano pyrimidine (**2**) was identified as a reagent of high interest due to its reactivity profile. Upon reaction with reduced Fab, **2** was able to generate a statistical distribution of up to 5 additions (Scheme 23). This preliminary result sets the starting point for the synthesis of potential nitrile CLT reagents, it was hypothesised that electron-deficient nitriles were required for rapid and efficient CLT. Higher equivalents were empolyed in comparison to the prior CBT in order to push the reaction further, therefore highlighting the standout reagents. The structures seen in red provided little to no modifcation of the Fab fragment (Scheme 23).



Scheme 23. (A) Reaction scheme for the CLT reaction of 2-cyano pyrimidine **2** reaction on reduced Fab. Compounds illustrated in red failed to show signs of CLT (**2.1** - **2.5**). **(B)** LCMS of 2-cyano pyrimidine **2** reactions on Fab. A) Reduced Fab B) 2-cyano pyrimidine **2**, 10 equiv., 18 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 105 Da. +2 (Green box), +3 (Light blue box), +4 (Dark blue box), +5 (Purple box) refer to number of additions.

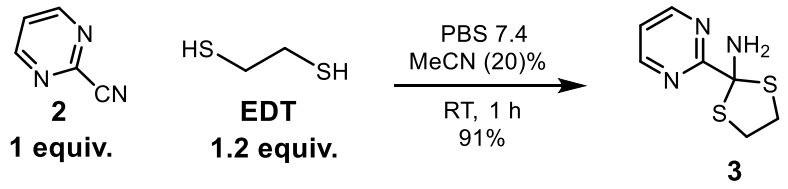
In order to further understand the CLT reagent reactivity, a small molecule study was also conducted, in which **2** was reacted with a mixture of *N*(α)-acetyl-L-lysine methyl ester and 2-methoxyethane-1-thiol, mimicking lysine residues and cysteine thiol respectively. When incubated with the amine alone no reaction occurred with **2** even after 24 h. The isolated species was thought to be an amidine (**2.6**) however, the possiblity of an amide formation was also feasible. The amide (**2.7**) was therefore synthesised *via* an amide coupling reaction to compare the shifts of the quaternary carbon. The amide shift (164.8) was found to be 10 ppm downfield when compared to the product (154.4) in identical solvent systems (Scheme 24).



Scheme 24. Amidine and amide comparison. (**A**) Synthesis of the amidine product (**2.6**) highlighting the C=N shift at 154.3 ppm. (**B**) Synthesis of the amide product (**2.7**) highlighting the C=O shift at 164.8 ppm.

2.2. Mechanistic studies

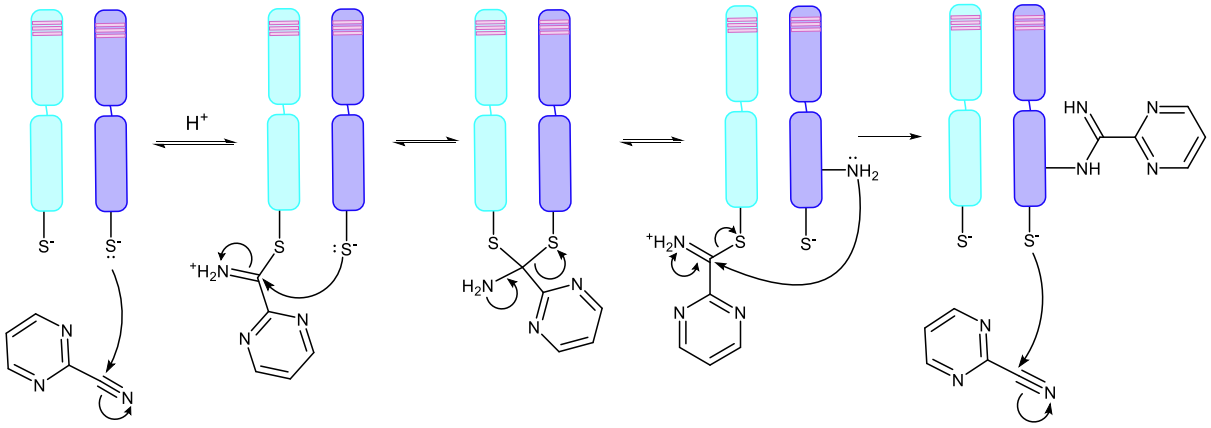
Following on from the previous work, a key part of the project revolves around the understanding of how these reagents react on the protein systems. In order to emulate the disulfide bond present on Fab, ethanedithiol, a simple commercially available bis-thiol, was reacted with **2** under buffered conditions resulting in the nitrile bis-thiol (NBT) reaction (Scheme 25). The result of this reaction was an impressive yield of the amino dithioacetal (ADTA), **3**. The NBT reaction is further discussed in Chapters 3 and 4.



Scheme 25. Reaction scheme for the EDT-NBT reaction, synthesising ADTA **3**.

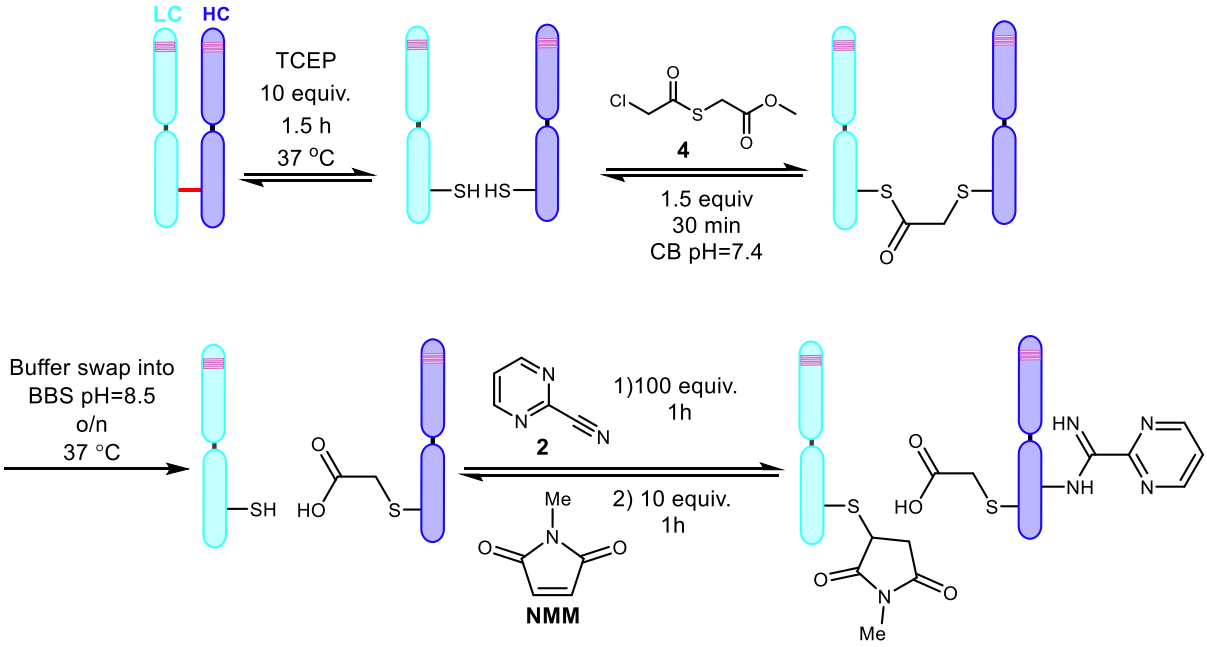
The NBT reaction, alongside the work done by Bayley *et al*., suggests that an ADTA type species could form during the mechanism of nitrile CLT. Control experiments on non-reduced Fab (100 equiv., 1 h, 22 ºC, pH 7.4), demonstrate that 2-cyano pyrimidine does not react unselectively with lysine residues, yielding only native Fab. The hypothesised mechanism behind nitrile CLT, assumed the nucleophilic free thiol reversibly reacts with the electrophilic carbon to form a thioimidate intermediate, which can undergo further reversible attack by a second free thiol (Scheme 26). In order to determine the importance of this second

nucleophilic attack in relation to CLT, one of the thiols involved in the disulfide bridge was capped before the addition of nitrile **2**. This would indicate whether the formed bridged species is required for CLT.



Scheme 26. Proposed mechanism behind nitrile CLT. Initial attack of the cysteine thiolate reversibly generates a thioimidate species. A second nucleophilic attack is also possible, generating a short lived tetrahedral intermediate. The reactive species responsible for CLT is hypothesised to be the protonated thioimidate, which upon transfer generates the amidine conjugate, leaving the thiols free for further CLT reactions.

As before, the disulfide bond was reduced using TCEP, to expose the two free thiols. In order to obtain controlled capping, methyl 2-[(2-chloroacetyl)sulfanyl]acetate (**4**) was utilised (kindly provided by Dr. Alina Chrzastek).[210] 1.5 equiv. of **4** successfully re-bridged the Fab disulfide in 30 min at 22 °C, pH 7.4. Due to the nature of the bridging reagent, one side forms a stable alkylated cysteine moiety, whilst the other side forms a hydrolytically labile thioester moiety *via* S_N2 displacement and transthioesterification respectively. The formed bridged product was buffer-swapped into pH 8.5 borate buffered saline (BBS) at 37 °C overnight, in order to hydrolyse the labile thioester bond at the LC. This hydrolysis step provides a preferential capping of the HC over the LC (~5:1). To this partially capped conjugate, was added 2-cyano pyrimidine (100 equiv., 1 hr, 22 °C) followed by addition of another cysteine capping reagent NMM (10 equiv., 1 hr, 22 °C) (Scheme 27).



Scheme 27. Mechanistic study of nitrile CLT. Reduction of native Fab with TCEP ii) Addition of methyl 2-[(2-chloroacetyl)sulfanyl]acetate to re-bridge Fab iii) Buffer-swap into BBS pH 8.5 to induce hydrolysis at the LC iv) Addition of 100 equiv. 2-cyano pyrimidine v) Addition of 10 equiv. *N*-methyl maleimide (NMM).

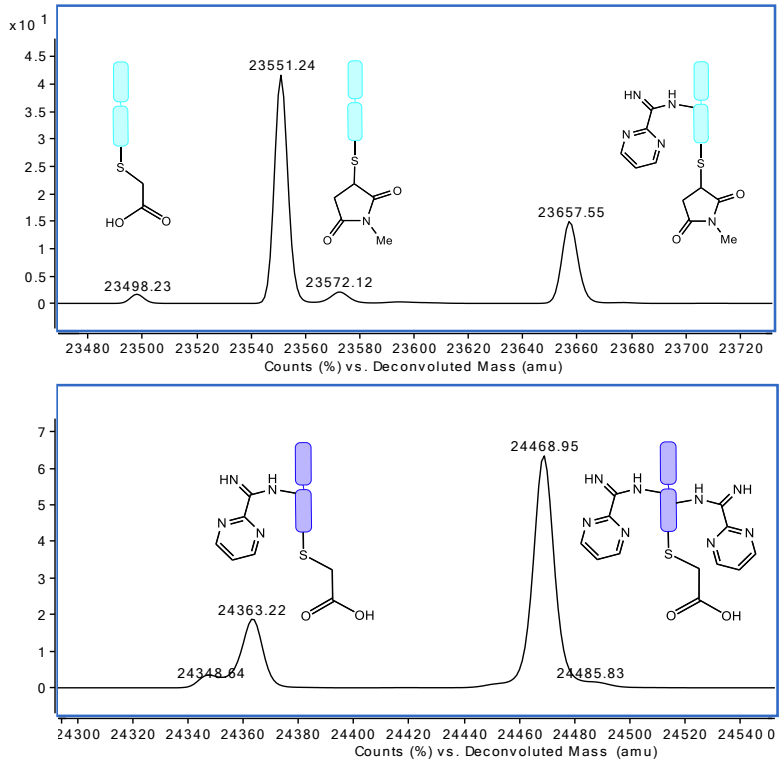
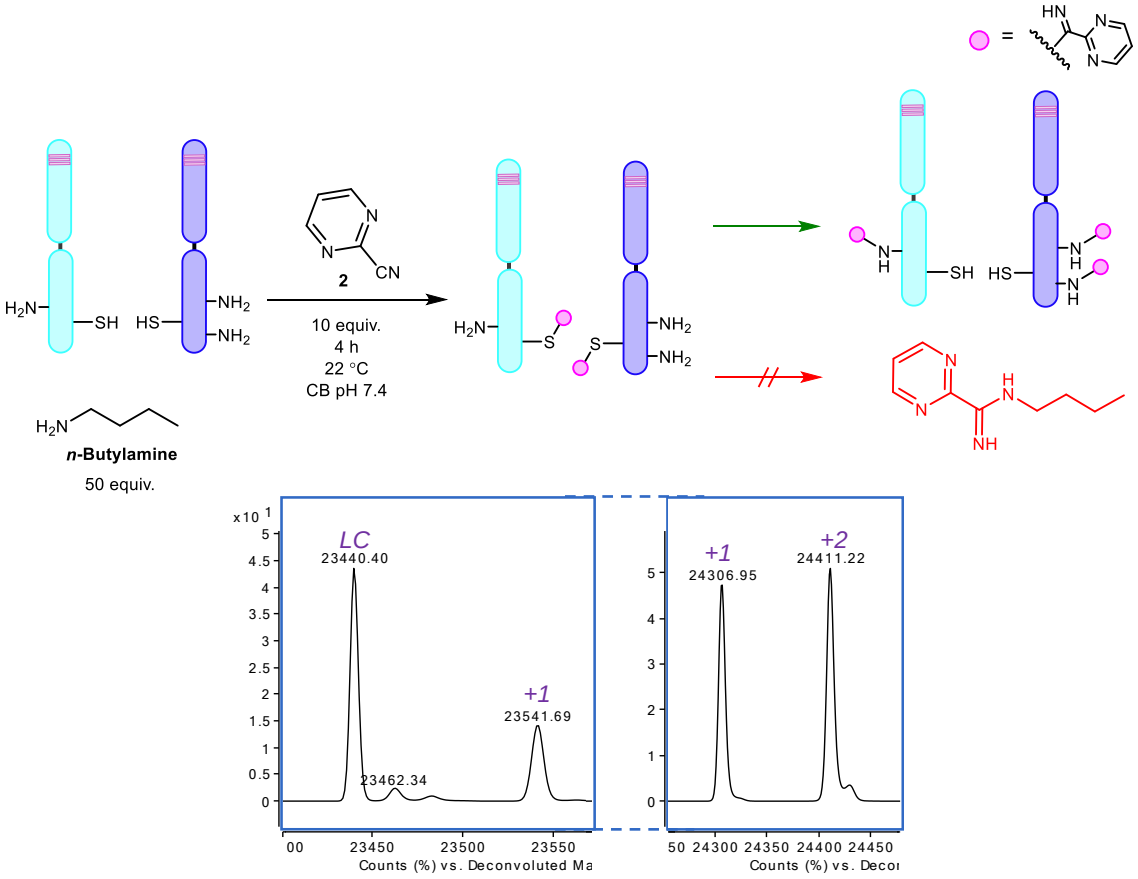


Figure 23. LCMS of mechanistic study; Mass of native LC = 23440 Da. Mass of native HC = 24201 Da.

From the MS data, it can be concluded that the second free thiol is not required for nitrile CLT. In the LC region (Figure 23), it can be seen that the major peak present is the maleimide capped species. There is also maleimide capped LC with an addition of one nitrile indicating CLT also occurs on a LC lysine residue. As for the HC region (Figure 23), both species present indicate capping by the chloroacetyl thioester with up to +2 additions of nitrile *via* LC cysteine to HC lysine transfer. The capping of both cysteines involved in the disulfide bond by either the chloroacetyl thioester or NMM reveals that nitriles are not attached to the cysteine thiols. In order to determine the intramolecular selectivity of nitrile CLT, a large excess (50 equiv.) of free amine (*n*-butylamine) was added to the reaction mixture. This large excess failed to inhibit CLT, alluding to the idea that the positioning of lysine residues on Fab is better suited for intramolecular *S*-to-*N* transfer (Scheme 28).



Scheme 28. Intramolecular selectivity competition experiment between Fab and *n*-butylamine. 2-cyano pyrimidine **2**, 10 equiv., *n*-butylamine, 50 equiv. 4 h, 22 ºC, pH 7.4. Mass of native LC = 23440 Da. Mass of native HC = 24201 Da. Expected mass of addition = 105 Da. +1 and +2 refer to number of additions.

2.3. Electron-deficient nitriles

A common feature of nitrile CLT was the addition of multiple molecules resulting in the formation of a statistical distribution. The distribution was thought to form due to the highly reversible nature of thiol addition to nitriles, leading to inconsistent loading onto the free cysteine thiols. The synthetic aims therefore looked at developing the reagent design in order to limit the formation of the statistical distribution. It was hypothesised that more electron-deficient nitriles could limit this distribution, as the equilibrium would shift in favour of the thioimidate intermediate. The shift would aim to 'load' the nitriles on to the free thiols prior to transfer by limiting the reversal and therefore potentially the statistical distribution. The reagent design would also look to introduce areas of functionality to assist with post conjugation modifications.

2.3.1. Chemical Synthesis of Pyrimidine analogues

Initial development of reagent design was based on the pyrimidine core previously identified. Initially, a series of pyrimidine nitriles with electron-withdrawing groups (EWGs) were synthesised. The strength of the EWG was quantified by reference to the literature Hammett constants (σ). These values are obtained from ionisation constants of benzoic acids derivatives in solution. These constants take both the inductive and resonance effects of the substituent into consideration.[211] Whilst these constants are determined only according to benzoic acid and its derivatives, the general trend can also be applied to other aromatic rings substituted with other functional groups, such as the nitriles of interest (Figure 24).

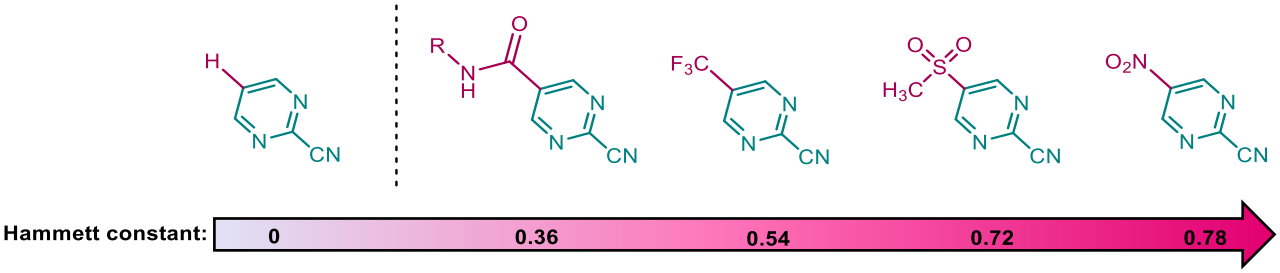
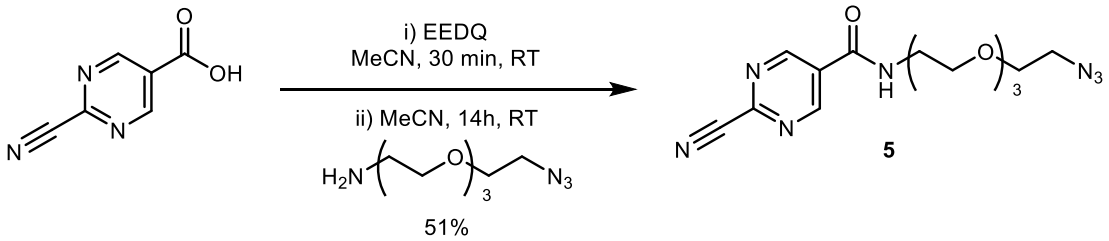


Figure 24. Schematic representation of *p*-EWGs on the pyrimidine ring.[211]

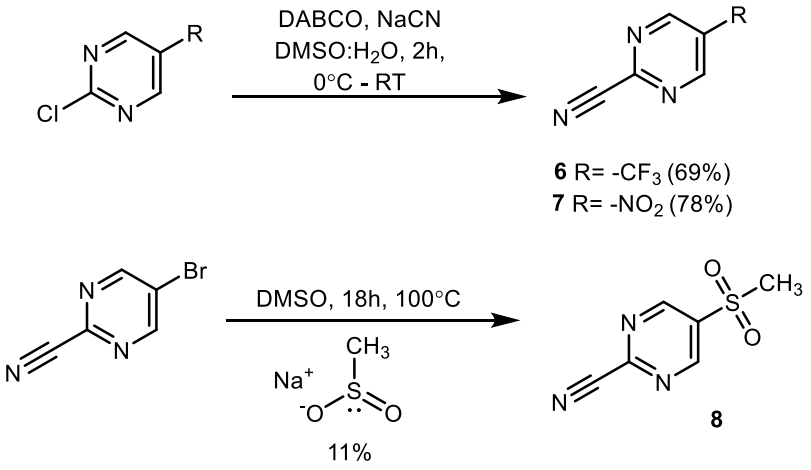
5 was synthesised *via* an *N*-Ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline (EEDQ) based coupling reaction with the commercially available 2-cyanopyrimidine-5-carboxylic acid. The selected amine was a PEGylated amine-azide, this choice provided a point of functionalisation alongside increased water solubility (Scheme 29).



Scheme 29. Synthesis of *p*-amide pyrimidine nitrile **5**.

6 and **7** were both synthesised *via* a similar cyanide S_NAr reaction of the commercially available aryl halide starting material. Due to the highly electron-poor ring, S_NAr provides a viable synthetic route as it is able to exploit the electrophilicity of the target carbon electrophile and stabilise the Meisenheimer complex formed during the reaction.[212] Initial attempts to synthesise these reagents were carried out at room temperature in air however, the extreme electron-deficiency of these systems saw degradation of organic matter upon addition of sodium cyanide. The reactions were therefore successfully conducted at 0 °C - RT under an N_2 atmosphere to protect any sensitive intermediates (Scheme 30).

8 was also synthesised *via* S_NAr reaction of the commercially available starting material 5-bromopyrimidine-2-carbonitrile. Unlike **6** and **7**, the substitution was carried out at the position *para* to the nitrile moiety with sodium methyl sulfinate as the selected nucleophile. The poor yield (11%) and harsh conditions (100 °C, 18 h) required signify the poor nucleophilicity of sulfur in the +2 oxidation state (Scheme 30).

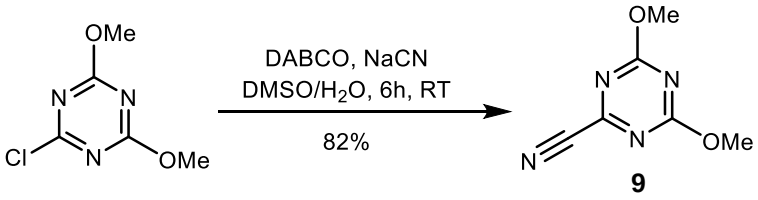


Scheme 30. Synthesis of pyrimidine nitriles **6**, **7** and **8**.

2.3.2. Chemical Synthesis of 1,3,5-Triazine nitriles

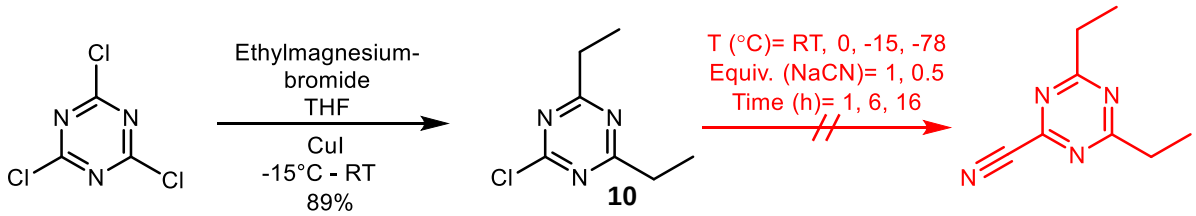
As seen before (pyridine vs pyrimidine) a greater number of nitrogens in the ring system results in greater electron-withdrawing ability, which in turn increases the electron-deficiency of the nitrile bond through inductive and mesomeric effect. One such ring is the 1,3,5-triazine which make it a target core for electron-deficient nitriles.

A 1,3,5-triazine nitrile **9** was synthesised from the commercially available 2-chloro-4,6-dimethoxy-1,3,5-triazine *via* S_NAr in a good yield (Scheme 31).



Scheme 31. Synthesis of 1,3,5-triazine nitrile **9.**

The alkoxy groups provide a potential point of dual-functionalisation in future designs, and these groups also possess a slight electron-withdrawing effect *via* a mild inductive effect when placed in the *meta* position indicated by the positive Hammett constant (0.12). Despite the lack of effect on the *meta* position other carbon substituted triazine rings were also investigated in order to limit overall electron donation into the ring. Following an adapted literature protocol, the diethyl substituted 2-chloro-1,3,5-triazine was synthesised from cyanuric chloride in excellent yield *via* Grignard mediated substitution.[213] Attempts to insert the nitrile handle were unsuccessful under a variety of conditions (Scheme 32). Each set of conditions led to the degradation of starting material with no signs of product formation. Investigation and analysis of the crude mixture yielded no further information into the identity of these degradation products, and as a result the carbon substituted triazine **10** was no longer pursued.



Scheme 32. Attempted synthesis of 1,3,5-triazine nitrile**.**

2.3.3. Chemical Synthesis of diene nitriles

Alongside greater electron-withdrawing effects, multi-heteroaromatic rings, namely 1,2,4-triazines and tetrazines, are able to undergo inverse electron-demand Diels Alder (iEDDA) reactions (Figure 25). This [4+2] cycloaddition *S*-tetrazines with various dienophiles represents an ideal bioorthogonal reaction. iEDDA reactions are one of the fastest modern bioorthogonal reactions with rate reaching 10^6 $M^{-1}s^{-1}$, 10000-fold greater than that of traditional CuAAC cycloadditions. iEDDA reactions are also selective and biocompatible, meaning these reactions can take place in aqueous conditions without the need for a metal catalyst e.g. Cu(I) which result in toxicity and purification concerns.[214]

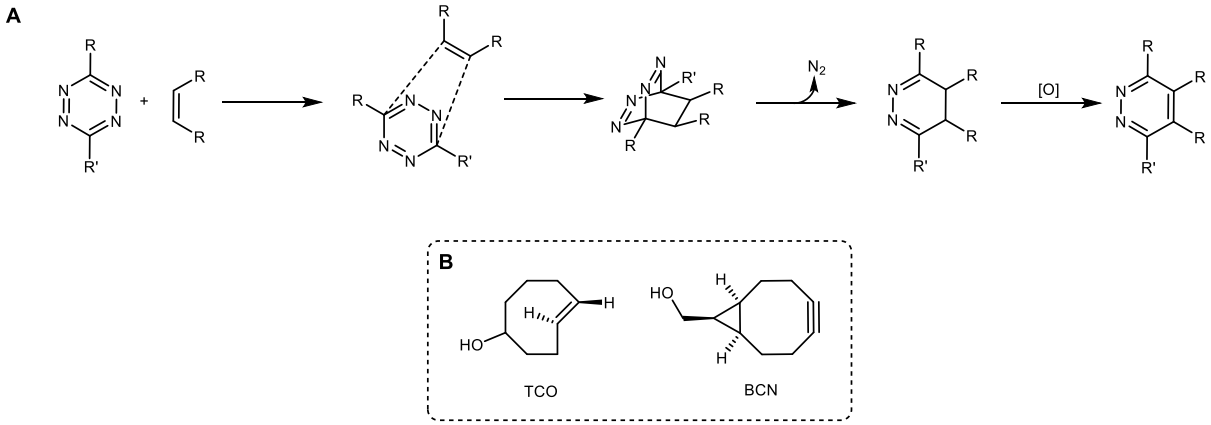
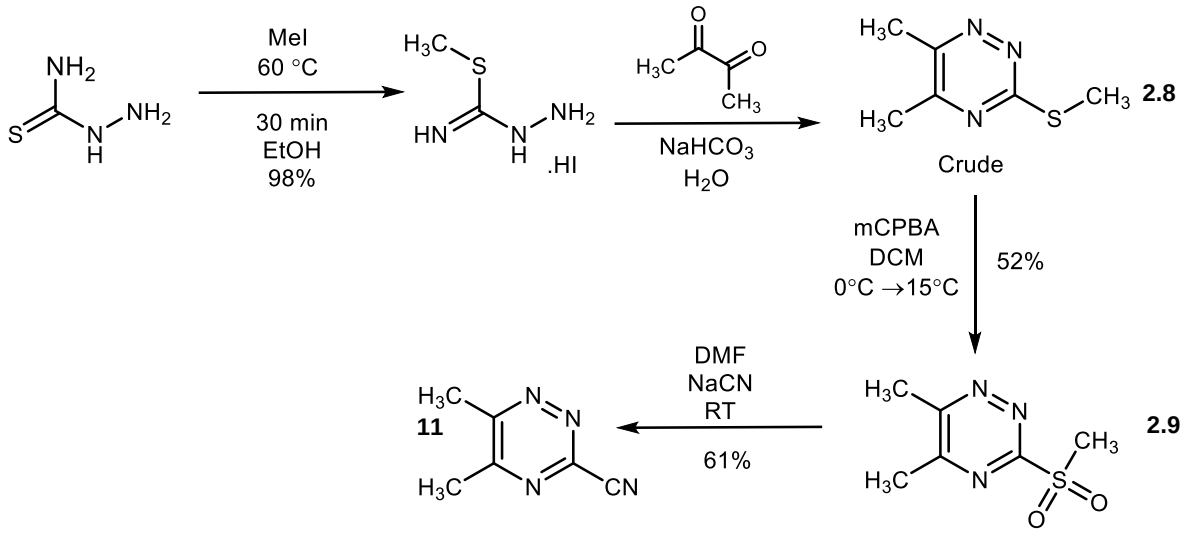


Figure 25. inverse electron-demand Diels Alder (iEDDA). **A**) Mechanism of iEDDA reactions with tetrazine rings as dienes. **B**) Structures of common dienophiles in bioconjugation.

5,6-dimethyl-3-(methylsulfonyl)-1,2,4-triazine was synthesised *via* a literature protocol,[215] where thiosemicarbazide was methylated to yield the desired *S*-methyl isothiosemicarbazide hydroiodide.

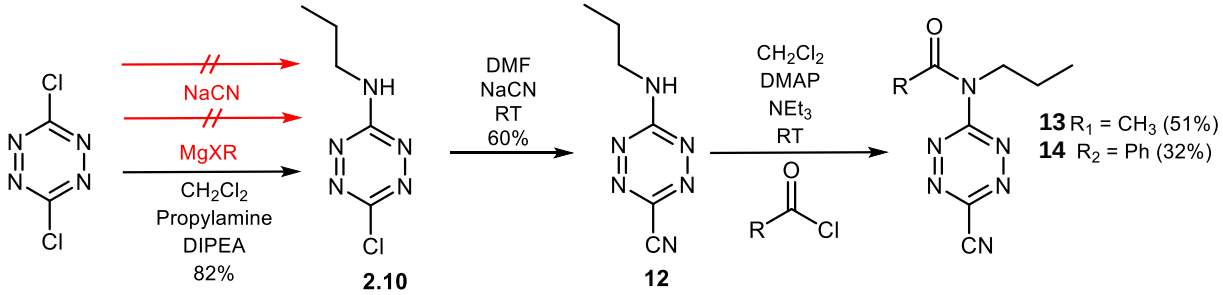


Scheme 33. Synthesis of 1,2,4-triazine nitrile **11.**

This was followed by a condensation reaction between the semicarbazide and diacetyl to form the core 3-(methylthio)-1,2,4-triazine (**2.8**). The sulphur atom was then oxidised to the +2 oxidation state with the use of *meta*-chloroperoxybenzoic acid (mCPBA) yielding a leaving group in the form of methanesulfinic acid (**2.9**). This group was substituted, as before, with cyanide to give the final 1,2,4 triazine nitrile, **11** (Scheme 33).

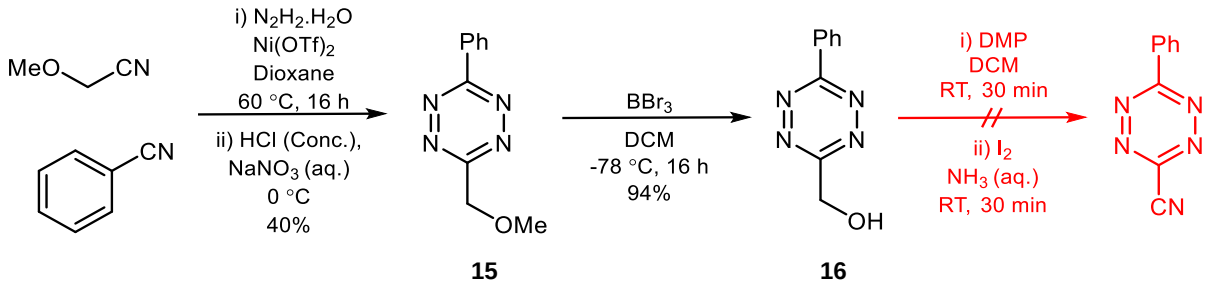
Next came the synthesis of tetrazine based nitriles from the readily available core; 3,6-dichloro-1,2,4,5-tetrazine. Initial attempts at substitution with cyanide or Grignard reagents, as before, failed to yield the product. Exchanging the nucleophile to simple amines such as propylamine proved to be effective in synthesising the 1st intermediate (**2.10**), this was successfully followed up with successful S_NAr utilising a cyanide nucleophile **12** (Scheme 34).

It was reasonable to assume the presence of such an electron-donating group in the *para* position would cause some issues regarding reactivity, the nitrogen of this group was therefore acylated to provide **13** and benzoylated to provide **14** in order to limit the lone pair delocalisation into the tetrazine ring (Scheme 34).



Scheme 34. Synthesis of 1,2,4,5-tetrazine nitriles **12, 13** and **14.**

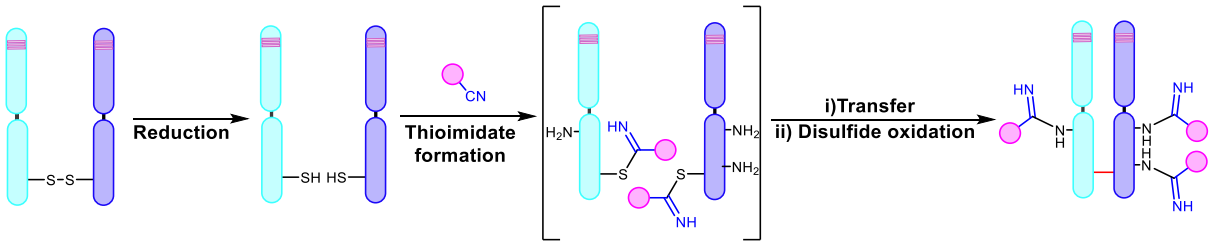
In order to fully exploit the features of the tetrazine ring, a carbon substituent became a key target. This is due to the lack of hindrance regarding the iEDDA functionalisation and electron-withdrawing capabilities of the tetrazine ring. Attempts to carry out S_NAr utilising cyanide-based nucleophiles resulted in no discernible product. It is thought that S_NAr on tetrazines with carbon nucleophiles is not easily achieved, and this is supported by the lack of literature data on such systems. The next plan of action was to synthesise the tetrazine ring and convert an incorporated group into the nitrile functionality. Togo *et al*. described a one-pot pathway in which various benzylic alcohols were efficiently converted into the corresponding nitriles in good yields *via* initial oxidation to the aldehyde followed by iodine-catalysed oxidative conversion to the nitrile moiety.[216] Tetrazine **15** was synthesised *via* a nickel triflate catalysed coupling of methoxyacetonitrile, phenylnitrile and hydrazine. This was followed by oxidation of the dihydrotetrazine with sodium nitrite and HCl (conc.) to the methoxy-substituted tetrazine **15**. The methoxy group was then demethylated using boron tribromide forming the tetrazine **16**, presenting a free alcohol group (Scheme 35). Phenylnitrile was selected as the use of acetonitrile generated a highly volatile tetrazine product. Attempts at oxidation of **16** to the aldehyde with agents such as TEMPO and Dess-Martin periodinane appeared to yield a new product however, the attempt to follow this with the oxidative conversion to the corresponding nitrile failed to yield any product. After multiple inconclusive attempts of synthesising a carbon-based tetrazine nitrile it was not further pursued. It is likely these carbon substituted tetrazine nitriles are highly reactive and thus may be unstable upon synthesis. Literature on such compounds is limited and requires special conditions e.g. the synthesis of a 1,4-dicyanotetrazine requires glovebox conditions with poor yields (10%) and harsh driving forces (160 °C; under 0.001 Torr), further limiting these regents as viable prospects.[217]



Scheme 35. Attempted synthesis of 1,2,4,5-tetrazine nitrile.

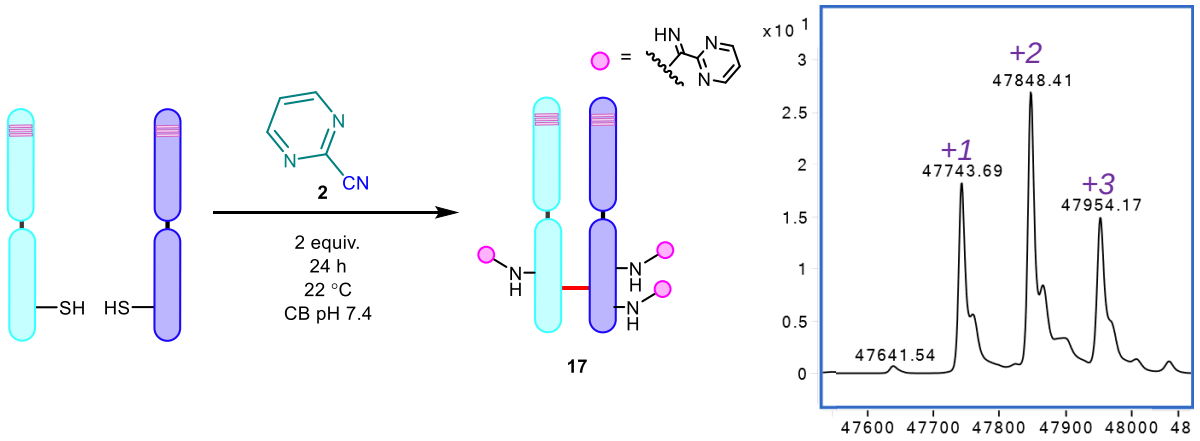
2.4. CLT bioconjugation studies

The CLT reagents were all studied on the trastuzumab Fab fragment. The fragment presents a single disulfide bond surrounded by proximal lysine residues. In order to access the disulfide bond, it must be reduced to the free cysteine thiols prior to the addition of nitrile reagents, allowing for the formation of thioimidate intermediates. These thioimidate species are able to spontaneously transfer to proximal lysine groups. Upon transfer, unlike thioesters, the disulfide bond does not reform however, when left for longer incubation times (e.g. 12-24 h) the disulfide bond was found to reform. Interestingly the purification of the conjugates with ultrafiltration using a Zebaspin resulted in oxidation of the disulfide bond, even during shorter reaction times, (e.g. 30 min), however, this proved to be inconsistent. Therefore, following the reaction, an oxidising agent such as 5,5-dithio-bis-(2-nitrobenzoic acid), also known as DTNB, was added for shorter timeframe reactions (0-4 h).



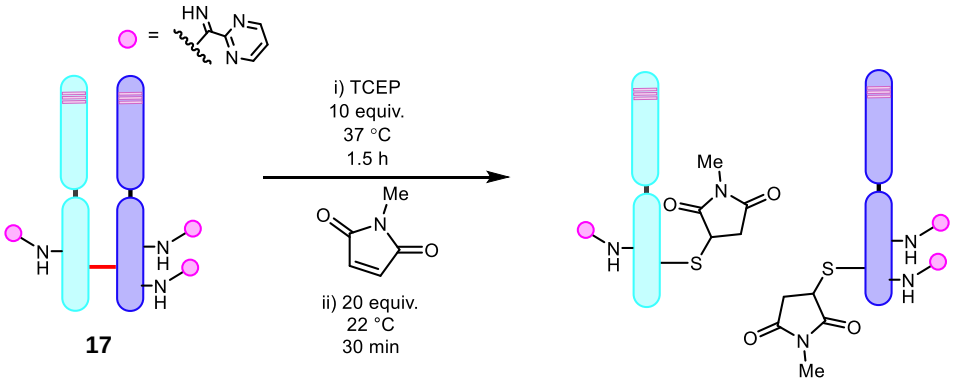
Scheme 36. Example of nitrile CLT protocol on trastuzumab Fab.

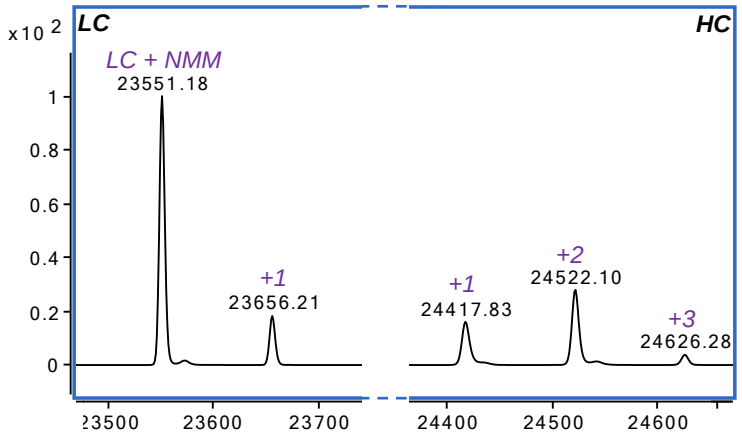
A major drawback to the use of thioesters is their vulnerability towards hydrolysis. In order to determine whether nitriles suffer from a similar drawback, reduced Fab was reacted with a stochiometric amount of 2-cyano pyrimidine (2 equiv.) over 24 h. Here, we can see that quantitative conversion is achieved, to afford an average loading of ~2 (1.92), indicating that the reagent is not hydrolytically labile (Scheme 37).



Scheme 37. LCMS of 2-cyano pyrimidine **2** reactions on Fab. 2-cyano pyrimidine **2**, 2 equiv., 24 h, 22 ºC, pH 7.4. Mass of native Fab = 47641 Da. Expected mass of addition = 105 Da. +1, +2, +3 refer to number of additions.

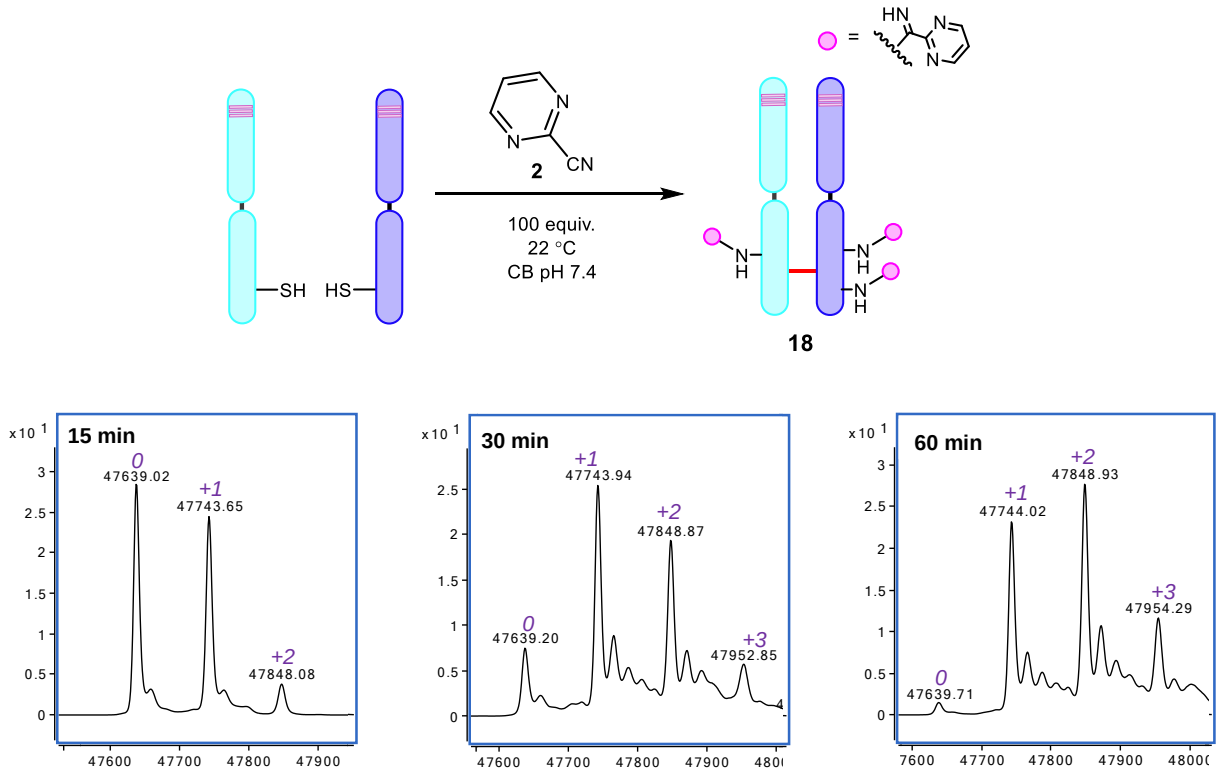
In order to further investigate the sites of modification, the Fab conjugate was reduced once again to yield the free thiols and capped with *N*-methyl maleimide (NMM), which would allow a comparison to the CLT sites of thioesters (Scheme 38). The resulting capped species seen were in line with that of thioester CLT, with the majority of the modification on the HC. This is thought to be modification on same lysine residues as with thioesters (K136, K221, K225). Modification was also seen on the LC (23544 Da) which aligns with possible minor modifications observed with thioester CLT reagents at either K126 or K190.[186]





Scheme 38. LCMS of reduced and capped Fab conjugate **17**. TCEP, 10 equiv., 1.5 h, 37 ºC, pH 7.4, NMM 20 equiv., 30 min, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Mass of native LC = 23440 Da. Mass of native HC = 24201 Da. Expected mass of addition = 105 Da. +1, +2, +3 refer to number of additions

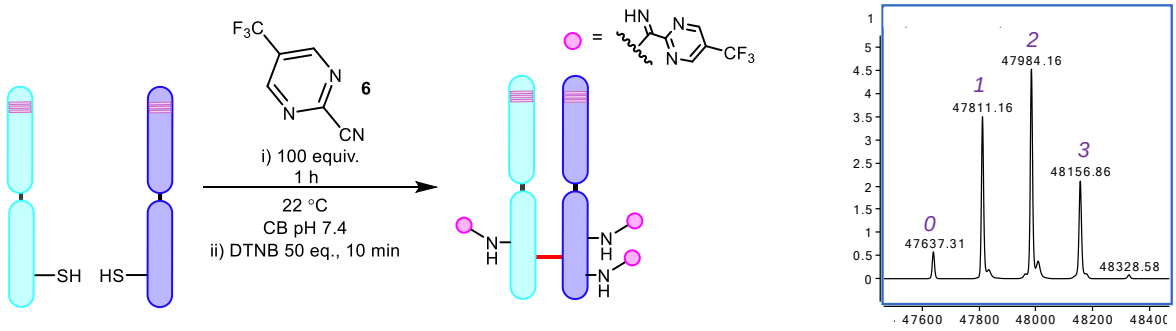
In order to increase the amount of CLT, the number of equivalents was increased drastically, from 2 equiv. to 100 equiv. A key feature for CLT reagents is the thiol selectivity; control experiments on non-reduced Fab (100 equiv., 1 h, 22 ºC, pH 7.4), indicated that 2-cyano pyrimidine **2** does not react unselectively with lysine residues. Upon reduction of Fab and addition of 100 equiv. of reagent, time dependant modification of the fragment is seen (Scheme 39). After 1 hour, 98% modification is seen with the major species being 2 additions (43%). This yields an average DAR of 1.8 using 2-cyanopyrimidine. These high equivalent reactions provide the same modification pattern as seen with low equivalent nitrile reactions subjected to a second reduction and NMM cap. The resulting conjugates were found to be stable under a range of conditions including exposure to external thiols such as GSH.



Scheme 39. LCMS of 2-cyano pyrimidine **2** reactions on Fab. 100 equiv., 22 ºC, pH 7.4. Mass of native Fab = 47639 Da. Expected mass of addition = 105 Da. +1, +2, +3 refer to number of additions.

2.4.1. CLT of 5-(trifluoromethyl)pyrimidine-2-carbonitrile (6)

The initial set of conditions for 5-(trifluoromethyl)pyrimidine-2-carbonitrile **6** were used as for **2**, using high equivalents for a short timeframe, and as expected the incorporation of an electron-withdrawing group resulting in greater loading in comparison to 2- cyanopyrimidine (Scheme 40). The higher loading of +3 additions was shown to form a greater proportion of the conjugate with **6** though the average DAR remained similar at 1.8, however this was due to there being more native Fab left over (5% vs 2%).



Scheme 40. LCMS of 5-(trifluoromethyl)pyrimidine-2-carbonitrile **6** reactions on Fab.100 equiv., 1h, 22 ºC, pH 7.4. Mass of native Fab = 47637 Da. Expected mass of addition = 173 Da. +1, +2, +3 refer to number of additions.

The increased rate of reactivity of this reagent was also studied with a longer timeframe, utilising lower equivalents (1.2 equiv., 12 h, 22 °C). It was hypothesised that the highly reactive **6** would load onto the free thiols of reduced Fab and transfer to nearby lysine residues before the reversal of the thioimidate. This would eliminate the need to remove excess reagents as with the high equivalent conditions. This however, failed to yield any advantages over the prior high equivalent conditions. The reaction was incomplete and provided a distribution of additions as before, whilst requiring significantly more time (supplementary information **S18**).

2.4.2. CLT of 5-(methylsulfonyl)pyrimidine-2-carbonitrile (8)

The next nitrile, **8**, provided an interesting design as the sulfonyl group is not only a strongly electron-withdrawing group but also an efficient leaving group. This combination of features could possibly provide rapid CLT *via* the nitrile followed by S_NAr mediated functionalisation. The reagent however demonstrated a major issue regarding an unusual addition of +26 Da to the final conjugate, though the exact identity of this species is yet to be confirmed. A common feature also seen with this reagent was the presence of a desulfonylated species. It was unknown whether the S_NAr was occurring at the cysteine thiol or another nucleophile such as the lysine amine. Work done by Bollong *et al.*[218] and Baud *et al*.[219] shows the use of sulfonyl heteroaromatic reagents for use in chemoselective cysteine arylation. This may support the idea of desulfonylation occurring *via* the attack of the cysteine thiol (Figure 26). The formation

of the desulfonylated species correlates with the number of equivalents added, where larger equivalents favour nitrile reaction over desulfonylation; 100, 75 and 50 equivalents result in 14%, 27% and 38% of the desulfonylated species, respectively. Due to the various shortcoming of the reagent, further bioconjugation was halted and alternative reagents were explored.

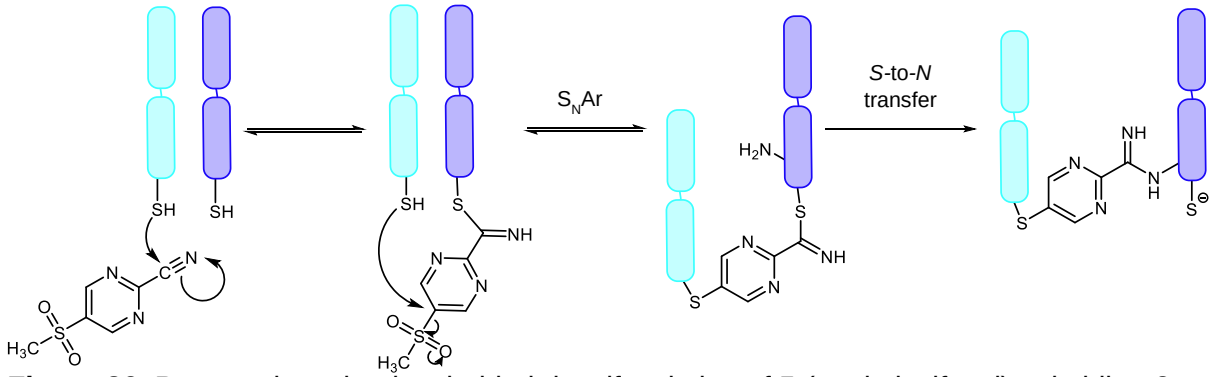
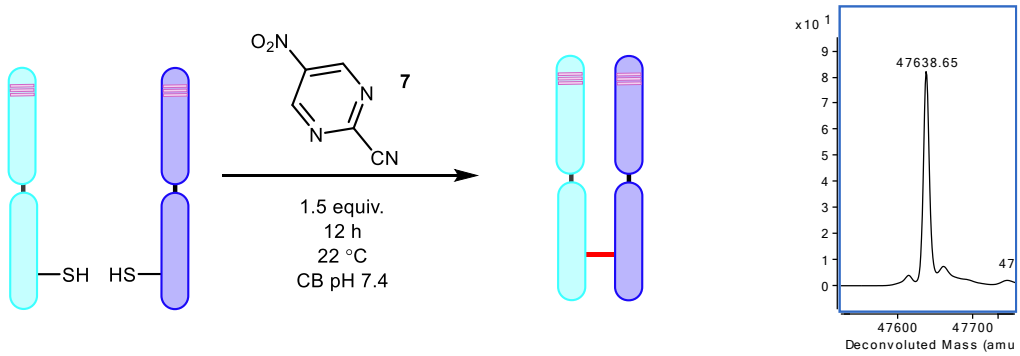


Figure 26. Proposed mechanism behind desulfonylation of 5-(methylsulfonyl)pyrimidine-2-carbonitrile **8**.

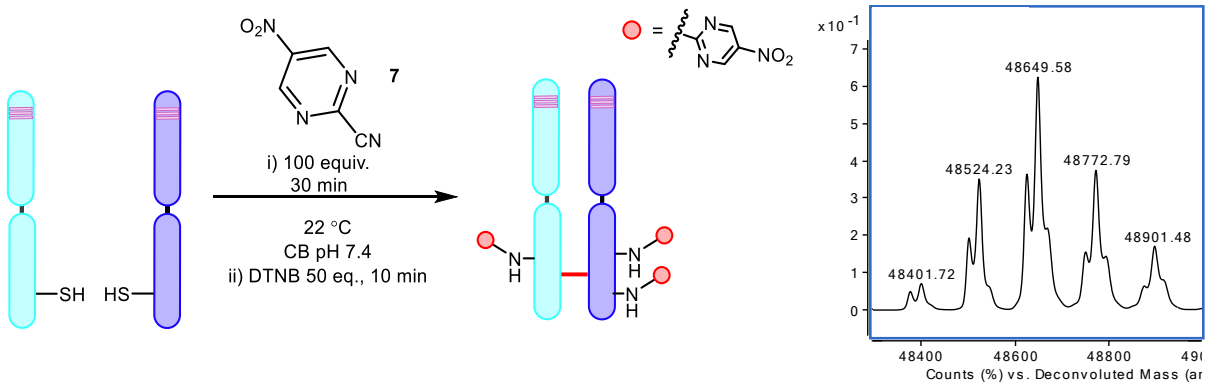
2.4.3. CLT of 5-nitropyrimidine-2-carbonitrile (7)

The next nitrile tested on Fab contained the strongest electron-withdrawing group in the form of the *p*-nitro pyrimidine. The *p*-nitro group has a Hammett constant of 0.78, which is 0.06 greater than the sulfonyl Hammett constant (0.72). Although this may not seem like a large increase in withdrawal, Hammett constants are in a logarithmic scale. Surprisingly, when subjected to Fab conjugation at low equivalents (1.5 equiv., 22 °C, 12 h), conjugation is not seen (Scheme 41). This may in fact be due to the very electron-deficient pyrimidine reagent degrading under the aqueous conditions, possibly due to hydrolysis.



Scheme 41. LCMS of 5-nitropyrimidine-2-carbonitrile **7** reactions on Fab. 1.5 equiv., 12 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 150 Da.

In order to investigate further and establish whether or not **7** undergoes CLT, a high equivalent reaction as before was conducted (100 equiv. 22 °C, 30 min). The result was found to show uncontrolled addition (up to +8) which correlated to the displacement of the nitrile group and not CLT (Scheme 42).



Scheme 42. LCMS of 5-nitropyrimidine-2-carbonitrile **7** reactions on Fab. 100 equiv., 30 min, 22 ºC, pH 7.4. Mass of native Fab = 47637 Da. Expected mass of addition = 150 Da.

The highly electron-deficient *p*-nitro pyrimidine ring combined with the ability of the nitro group to resonance-stabilise the Meisenheimer-complex support the formation of this species. The leaving group is that of the cyanide group (pK_a =9.21).

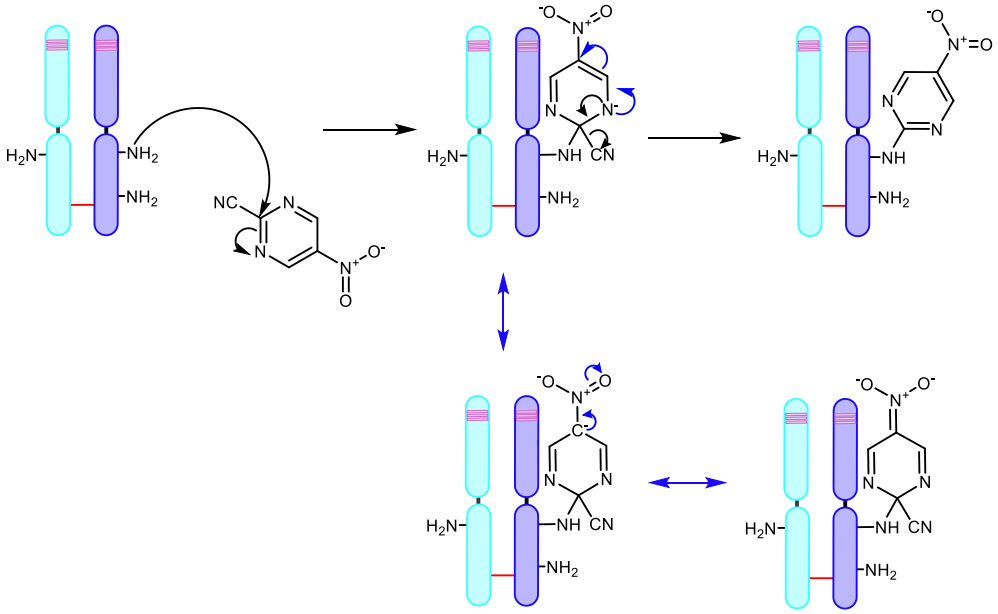
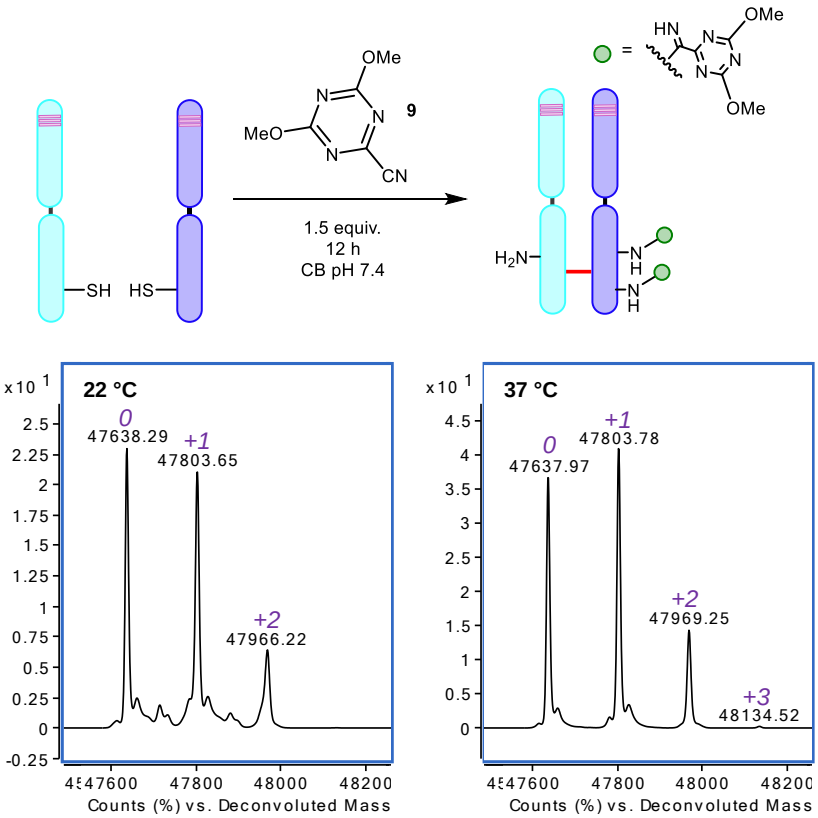


Figure 27. Proposed mechanism of S_NAr and resonance stabilisation resulting in loss of the cyanide group.

2.4.4. CLT of 4,6-dimethoxy-1,3,5-triazine-2-carbonitrile (9)

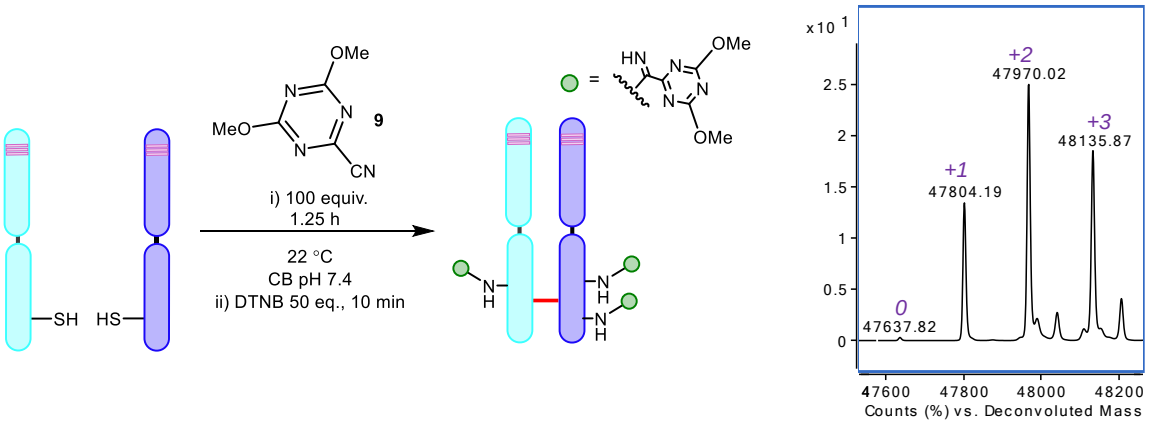
4,6-Dimethoxy-1,3,5-triazine-2-carbonitrile **9** was the first triazine nitrile tested on the Fab fragment. As seen before, the choice of group attached can have a large impact on the reactivity, therefore the next step in understanding nitrile CLT was to investigate alternate ring systems. As with prior reagents the use of stoichiometric equivalents resulted in the incomplete conversion to the CLT product (1.5 equiv., 12 h, pH 7.4). An increase in temperature from 22 ºC to 37 ºC was found to have no significant impact on the outcome (Scheme 43). At 22 ºC, 55% conversion was seen with a DAR of 0.67, when the temperature

was increased to 37 ºC only a 5% increase in conversion (60%) was seen with an increase of 0.09 for the DAR (0.76).



Scheme 43. LCMS of 4,6-dimethoxy-1,3,5-triazine-2-carbonitrile **9** reactions on Fab. 1.5 equiv., 12 h, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 166 Da. +1, +2, +3 refers to number of additions.

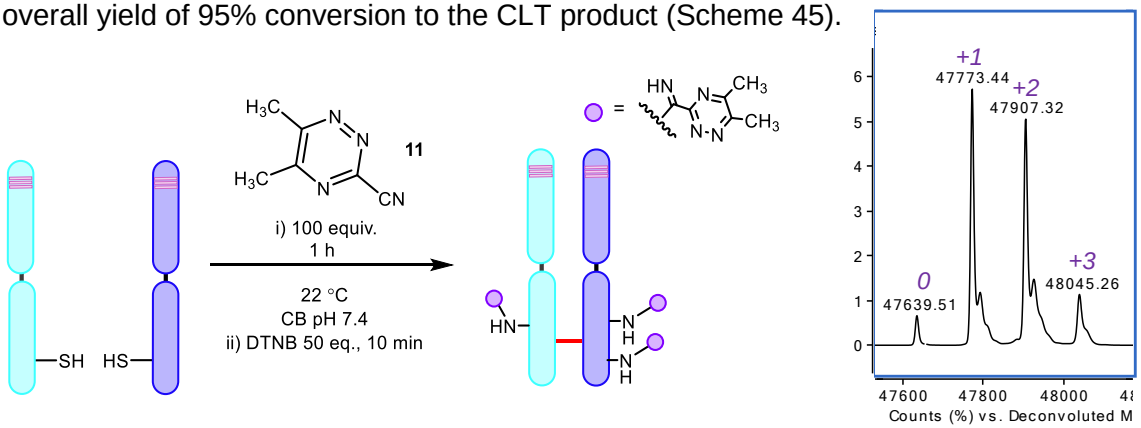
Therefore, in order to push the CLT towards completion, higher equivalent experiments were conducted. Control experiments (100 equiv., 1.25 h, 22 ºC, pH 7.4) showed no unselective lysine modification. Under these conditions (100 equiv., 1.25 h, 22 ºC, pH 7.4) with reduced Fab, a conversion of above 99% to the CLT product, with a DAR of 2.1 was seen. These conditions as before produced a distribution of up to +3 (Scheme 44). Small additional peaks can be seen in the LCMS trace however, this is thought to be due to impurities present in the LCMS as they do not correspond with a logical species related to the reaction. A second set of high equivalent conditions were also applied to investigate the ability to 'load' the reagent onto the cysteine thiols, followed by removal of the excess regent prior to transfer (125 equiv., 5 min, 22 ºC, pH 7.4 followed by 2 h 22 ºC, pH 7.4 for transfer). Impressively the reagent was able to obtain an 85% conversion to the CLT product with a DAR of 1.4 however no additional selectivity was obtained as a distribution of up to 3 additions was seen.



Scheme 44. LCMS of 4,6-dimethoxy-1,3,5-triazine-2-carbonitrile **9** reactions on Fab. 100 equiv., 1.25 h, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 166 Da. +1, +2, +3 refers to number of additions.

2.4.5. CLT of 5,6-dimethyl-1,2,4-triazine-3-carbonitrile (11)

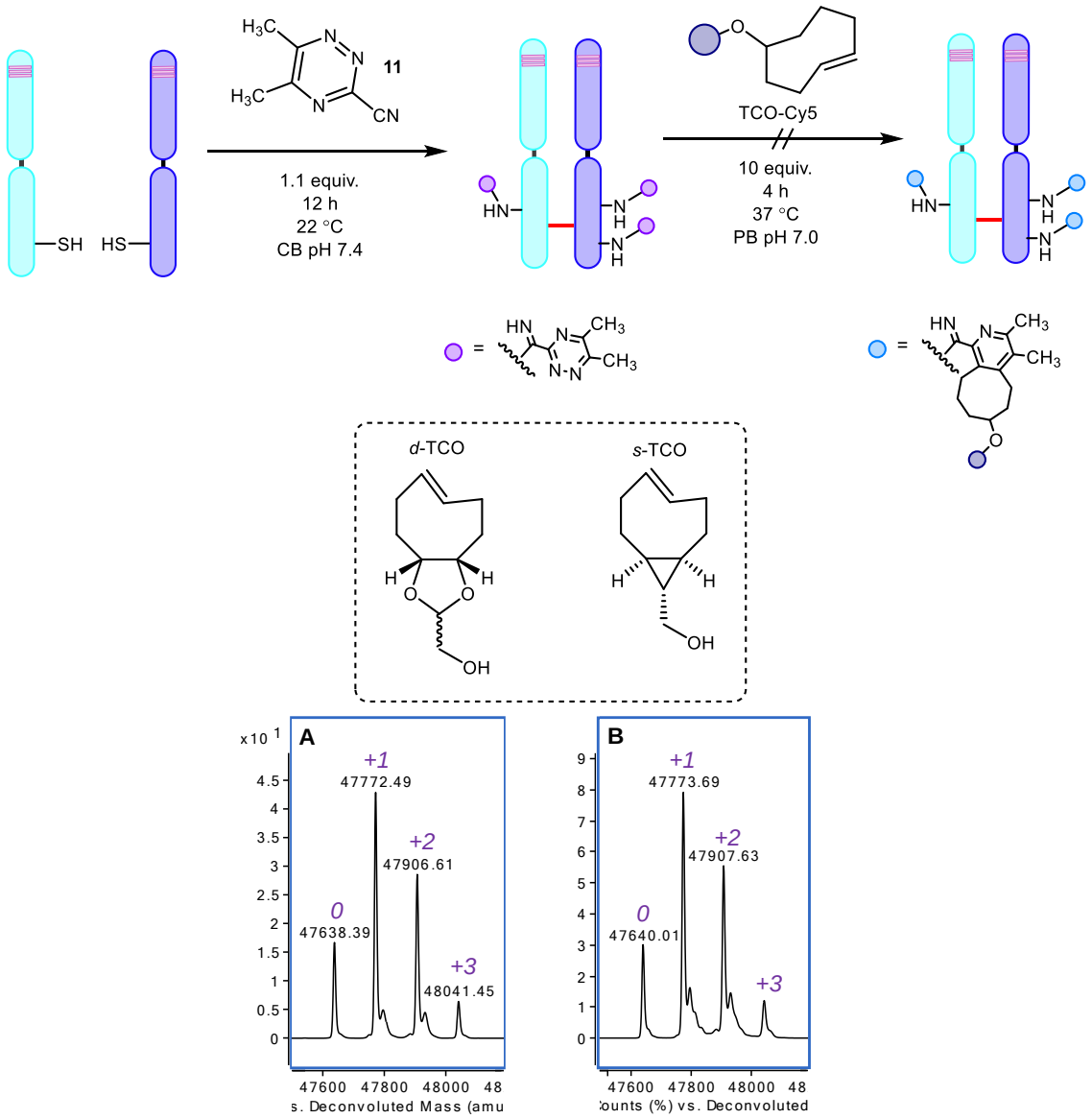
The next triazine nitrile tested on the Fab was the 1,2,4-triazine; **11**. The first set of conditions trialled on **11** were as follows: 100 equiv., 1 h, 22 ºC, pH 7.4. With these conditions, a distribution of up to +3 additions was seen. The conjugation resulted in a DAR of 1.5 and an overall yield of 95% conversion to the CLT product (Scheme 45).



Scheme 45. LCMS of 5,6-dimethyl-1,2,4-triazine-3-carbonitrile **11** reactions on Fab. 100 equiv., 1 h, pH 7.4. Mass of native Fab = 47639 Da. Expected mass of addition = 134 Da. +1, +2, +3 refers to number of additions.

As with prior nitrile CLT reagents, low equivalent reactions were also attempted with **11** (1.1 equiv., 12 h, 22 ºC, pH 7.4), and with these conditions, an 82% conversion was obtained with a DAR of 1.3. Attempts were then made to functionalise the conjugate *via* an iEDDA click reaction utilising a *trans*-cyclooctene fluorophore (Cy5). The resulting conjugate showed no signs of functionalisation *via* LCMS, this was thought to be due to the slow rate of iEDDA with 1,2,4 triazines and traditional TCO rings. In comparison to tetrazine TCO reactions (2000 $M^{-1}s^{-1}$), triazine TCO reactions are approximately 266-fold slower (7.5 $M^{-1}s^{-1}$).[214] In order to obtain

greater rates of reactivity (up to 76-fold greater), further strained systems such as dioxolane-fused or trans-bicyclononene derivative TCOs must be employed (Scheme 46).[220]

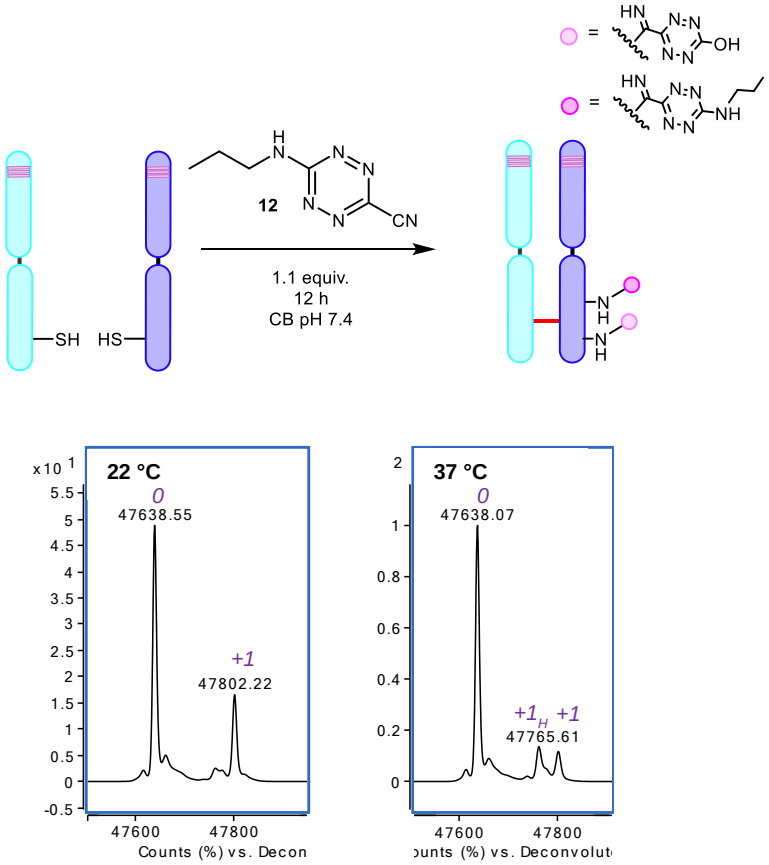


Scheme 46. LCMS of 5,6-dimethyl-1,2,4-triazine-3-carbonitrile **11** reactions on Fab. (A) 5,6-dimethyl-1,2,4-triazine-3-carbonitrile **11**, 1.1 equiv., 12 h, 22 °C, pH 7.4. (B) TCO-Cy5, 10 equiv., 4 h, 37 °C, pH 7.0. Mass of native Fab = 47638 Da. Expected mass of nitrile addition = 134 Da. Expected mass of TCO-Cy5 addition = 959 Da. +1, +2, +3 refers to number of nitrile additions.

2.4.6. CLT of 1,2,4,5-tetrazine-3-carbonitriles (12, 13 and 14)

As mentioned before, investigating alternative ring systems is key in uncovering new and potentially useful information regarding nitrile CLT. The next ring system was the 1,2,4,5-tetrazine. 6-(Propylamino)-1,2,4,5-tetrazine-3-carbonitrile **12** was the first tetrazine nitrile to be

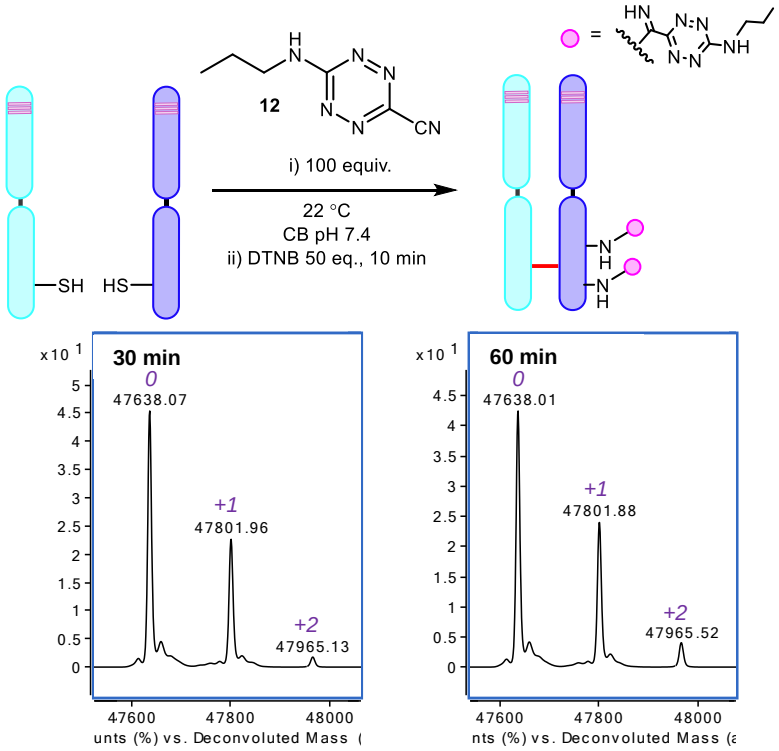
investigated, however, due to the directly attached *p*-amino group, concerns were that the lone pair would donate into the system and drastically impact the nitrile reactivity. Low equivalent reactions resulted in minimal modification over long durations (1.1 equiv., 22 °C, 12 h, pH 7.4), the resulting conversion was just 25% with a single addition. Surprisingly when the temperature was increased to 37 °C to push CLT, the conjugate appeared to undergo hydrolysis over time, resulting in the loss of propylamine. The conjugate was formed in 20% yield with two distinct peaks correlating to the unhydrolysed and hydrolysed products in 9% and 11% yields respectively (Scheme 47). Tetrazine rings are known to be susceptible to reduction, however, removal of TCEP was found to have no impact on the formation of the conjugate.



Scheme 47. LCMS of 6-(propylamino)-1,2,4,5-tetrazine-3-carbonitrile **12** reactions on Fab. 1.1 equiv., 12 h, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 164 Da. +1 refers to number of additions. $+1_H$ refers to the hydrolysed species.

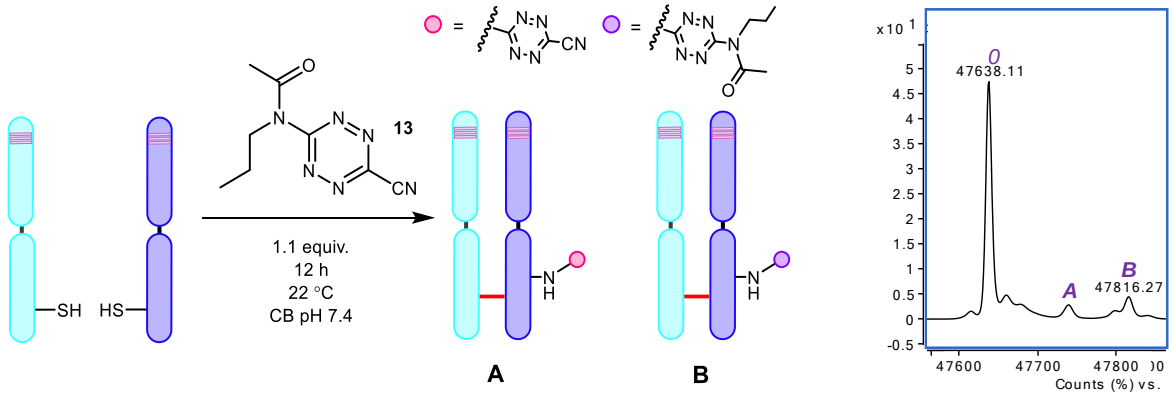
High equivalent reactions also failed to yield quantitative conversion to the CLT conjugates. 6-(Propylamino)-1,2,4,5-tetrazine-3-carbonitrile **12** also appeared to be less time-dependant as an increase of reaction time from 30 min to 1 h only resulted in a 1% increase in conversion to a +1 species. This could suggest reagent instability or possibly oxidation of the disulfide, limiting the progress of the reaction. The shorter timeframes however did not present any hydrolysed species indicating the time dependence of this secondary reaction (Scheme 48).

These results reveal that the amino group was too electron-donating and therefore would not be an advantage over the prior reagents designed and tested.



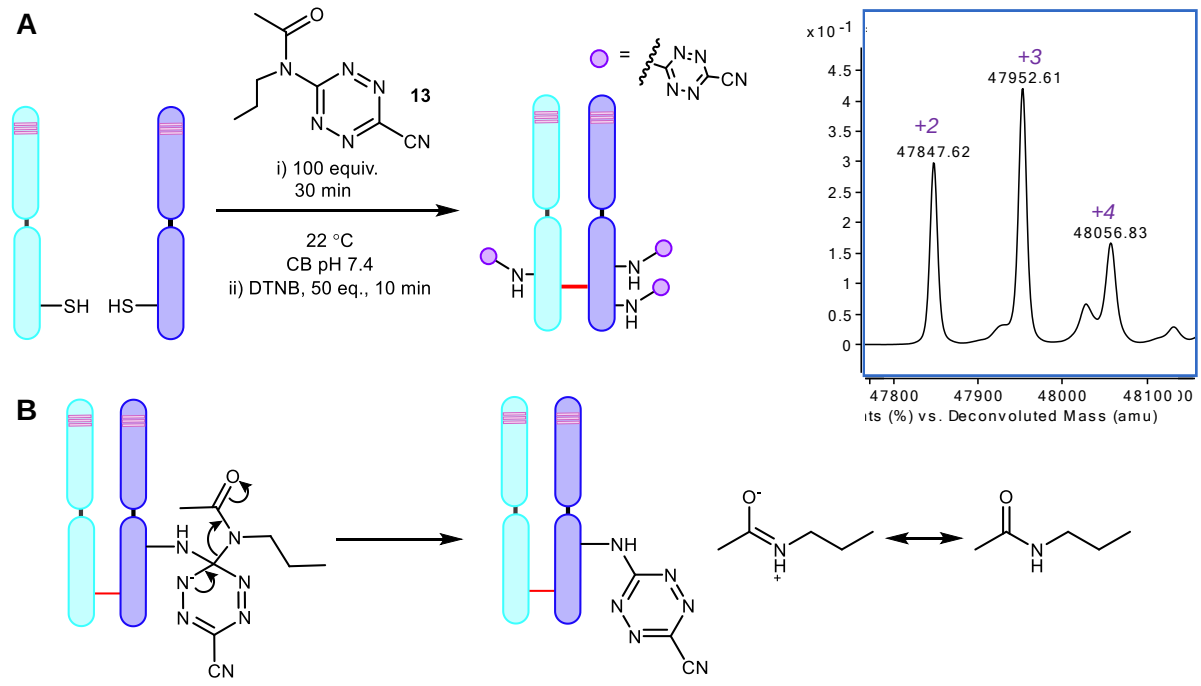
Scheme 48. LCMS of 6-(propylamino)-1,2,4,5-tetrazine-3-carbonitrile **12** reactions on Fab. 100 equiv., 22 °C, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 164 Da. +1, +2 refers to number of nitrile additions.

13 was the next tetrazine nitrile to be investigated, and the acylation of the nitrogen was thought to occupy the lone pair *via* resonance and limit the donation into the ring system in comparison to the previous design. The low equivalent reactions (1.1 equiv., 22 °C, 12 h, pH 7.4) showed a similar result to before, where there appeared to be minimal modification (Scheme 49). Upon closer inspection, the two minor species present corresponded to two forms of S_NAr; displacement of the amide moiety (**A**) or displacement of the nitrile moiety (**B**).



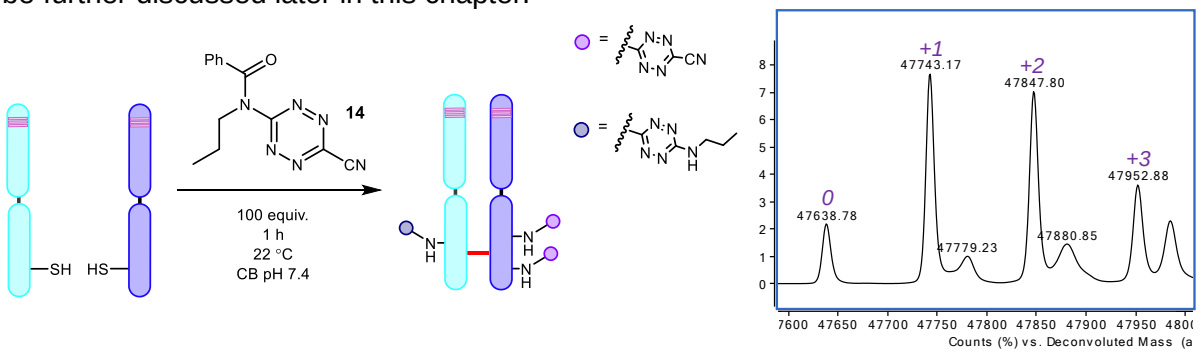
Scheme 49. LCMS of *N*-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylacetamide **13** reactions on Fab. 1.1 equiv., 12 h, 22 °C, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 206.

In order to push the reaction further and investigate the characteristics of this reaction a high equivalent reaction was carried out (100 equiv., 22 °C, 30 min, pH 7.4), which revealed raising the equivalents did not contribute to CLT (Scheme 50 A). Higher equivalents resulted in S_NAr at the 4-position of the tetrazine ring, leading to the loss of the amide moiety. The LCMS trace shows a species with the mass 47847 Da, this could correspond to two potential conjugates, the first being the CLT product (1 addition) and the second being the S_NAr product (2 additions). With the high equivalent reaction of this molecule, it is likely to be the S_NAr product, as even at low equivalents the S_NAr product was the only species seen of the two hypothesised. The mass for either conjugate both differ from the measured value by ±3 Da, which provides no further indication as to which species is present. The control experiment for this reaction with native Fab revealed multiple additional species of uncontrolled addition, when compared to the addition to reduced Fab there is a clear degree of selectivity in the presence of the free thiols. This S_NAr was unexpected as the leaving group, *N*-propylacetamide, is not generally considered as a good leaving group. The S_NAr can possibly be validated through the incredibly electron-deficient tetrazine ring paired with the *p*-cyano group, contributing to the labile nature of the *N*-propylacetamide group (Scheme 50 B).



Scheme 50. LCMS of *N*-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylacetamide **13** reactions on Fab. (**A**) *N*-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylacetamide **13**, 100 equiv., 30 min, 22 °C, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 206 Da. +2, +3, +4 refers to number of S_NAr additions. (**B**) Proposed mechanism of S_NAr.

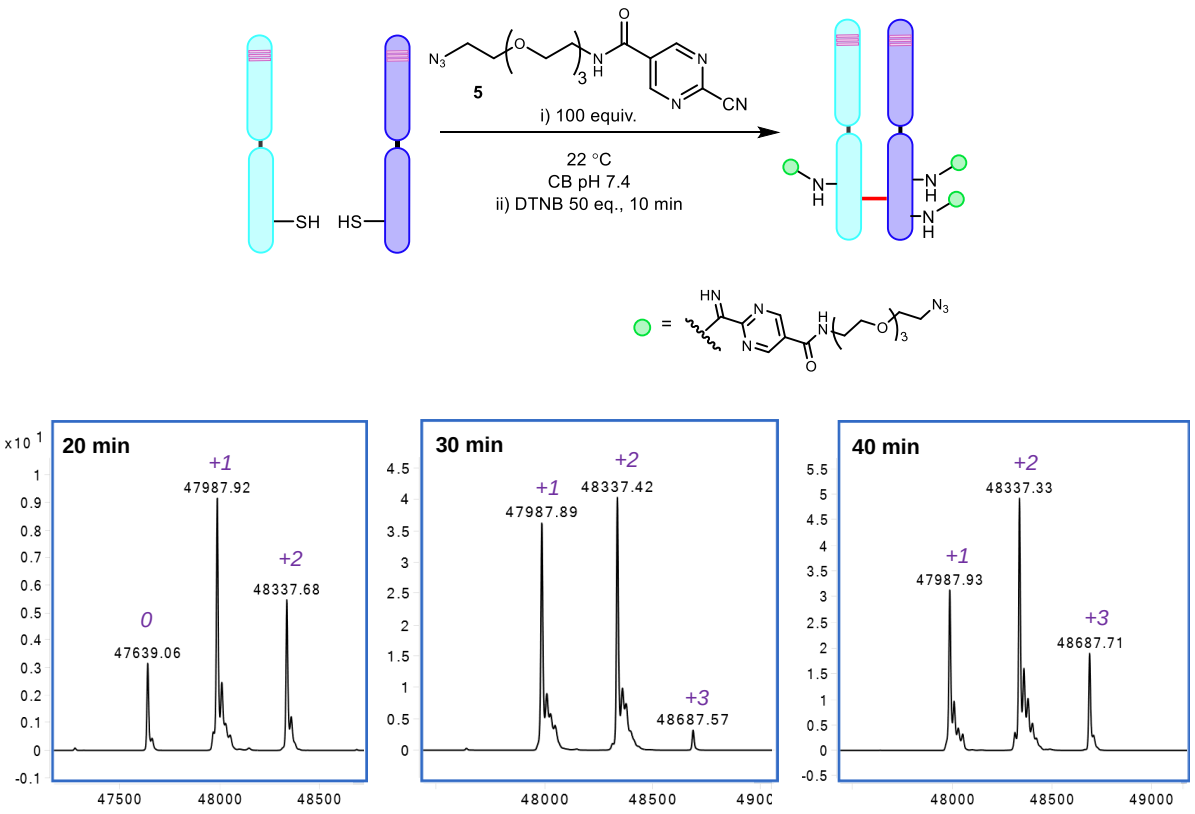
The final tetrazine design tested on Fab employed a benzoyl group **14**. This reagent was designed in an attempt to sterically hinder the rate of S_NAr in order to allow CLT *via* the nitrile group. The LCMS trace indicates the slight hindrance of S_NAr; however it is apparent that this was still the dominant reaction taking place (Scheme 51). The LCMS trace also shows the presence of new species (47779 Da), which does not correspond to the CLT product or the S_NAr of the amide as seen in the major species, it instead appears to be the combination of S_NAr at the nitrile position followed by amide hydrolysis. In general, aliphatic amides are known as hydrolytically stable at neutral pH and room temperature,[221] however, it is possible due to the aromatic nature of the tetrazine ring and benzoic acid leaving group, the conjugate is able to undergo slight hydrolysis.[222,223] This reagent was not further investigated due to the undesired outcomes, however, the reactivity of these reagent designs led to the idea of using S_NAr reagents as thiol selective modifiers and possibly a new class of CLT reagent. This will be further discussed later in this chapter.



Scheme 51. LCMS of *N*-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylbenzamide **14** reactions on Fab. 100 equiv., 1 h, 22 °C, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 268 Da. +1, +2, +3 refers to number of S_NAr additions.

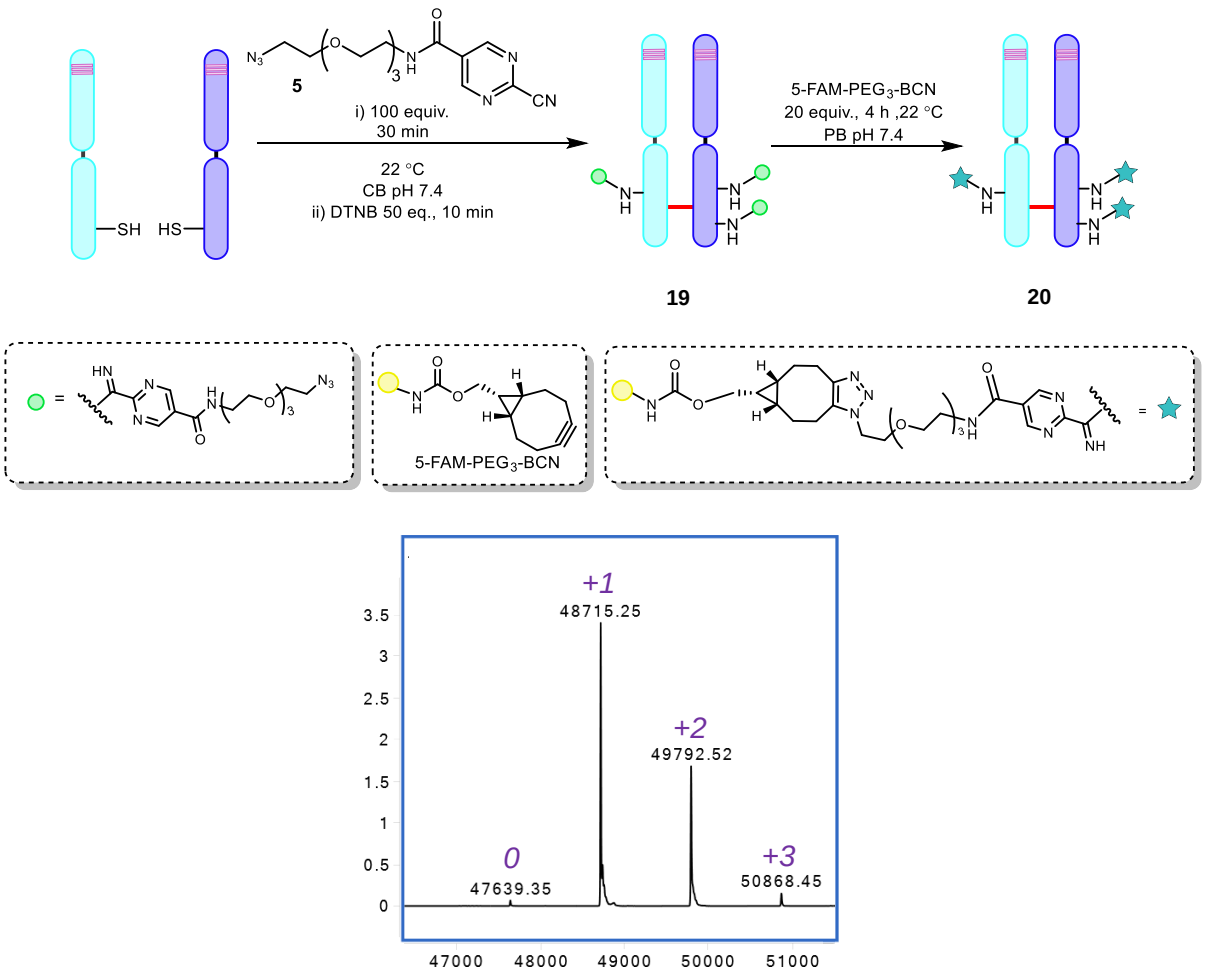
2.4.7. CLT of *N*-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide (5)

The final nitrile tested for CLT on Fab was the *p*-amide substituted pyrimidine nitrile **5**. Under high equivalent conditions (100 equiv. 30 min, 22 °C, pH 7.4), **5** provided a quantitative conversion to the CLT conjugate with up to +3 additions seen. The major species present are that of +1 and +2 additions resulting in a DAR of 1.6, the formed conjugate provides functional handles in the form of an azide moiety (Scheme 52). High equivalent conditions were optimised due to the prior data highlighting the limitations of low equivalent reaction such as timeframe and incomplete conversion.



Scheme 52. LCMS of N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide **5** reactions on Fab. 100 equiv., 22 °C, pH 7.4. Mass of native Fab = 47639 Da. Expected mass of addition = 350 Da. +1, +2, +3 refers to number of additions.

The CLT conjugate was functionalised with a 5-carboxyfluorescein fluorophore *via* a bicyclo[6.1.0]non-4-yne (BCN) click reaction (SPAAC), this functionalisation is a cycloaddition promoted by the bond angles of sp-hybridized carbons (160°) and does not require the use of a catalyst or ligand (e.g. Cu(I)).[224] The clicked conjugate provides a conversion of 99% with a DAR of 1.4 (Scheme 53). This slight change from the prior step may be due to the ionisation differences between smaller and larger proteins and subsequent ion detection. The smaller proteins are detected to a greater degree, this becomes more apparent as the mass difference increase.



Scheme 53. LCMS of SPAAC reaction on Fab conjugate **19**. 5-FAM-PEG$_3$-BCN, 20 equiv., 22 °C, 4 h, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 1077 Da. +1, +2, +3 refers to number of additions.

A FAR was determined photometrically by obtaining the UV/Vis absorption reading of the conjugate **20** and using the following formula, in which *Cf* is the correction factor for the absorbance of the carboxyfluorescein at 280 nm. The FAR value of 1.5 is close to that of the CLT loading calculated from the LCMS spectrum of conjugate **19** (1.6), this supports the idea that ionisation differences result in the unexpected distribution seen in the LCMS trace for conjugate **20**.

$$FAR = \frac{\frac{Abs_{505}}{\varepsilon_{505}}}{\frac{Abs_{280} - (Cf\ x\ Abs_{505})}{\varepsilon_{280}}} = \frac{\frac{0.434}{83000}}{\frac{0.320 - (0.178\ x\ 0.434)}{68590}} = 1.5$$

2.5. CLT Stability studies

2.5.1. ELISA

An enzyme-linked immunosorbent assay (ELISA) was carried out on native Fab and the Fab conjugate **18**, containing the CLT product of 2-cyanopyrimidine, to test affinity towards the

HER2 antigen. A 96-well polystyrene plate was coated first with HER2 antigen followed by the application of native Fab and Fab conjugate **18**. Upon completion of the incubation time with a secondary antibody linked to HRP, *o*-phenylenediamine.HCl was added to quench the reaction. This was converted to 2,3-diaminophenazine by the enzyme HRP, which shows strong absorbance at 490 nm, allowing for spectrophotometric quantification of the antigen binding activity of the Fab conjugate **18** and native Fab. The Fab conjugate **18** demonstrated analogous antigen binding activity to that of native Fab. This supports the idea of modification of proximal lysine residues proximal to the disulfide bond, distant from the CDRs.

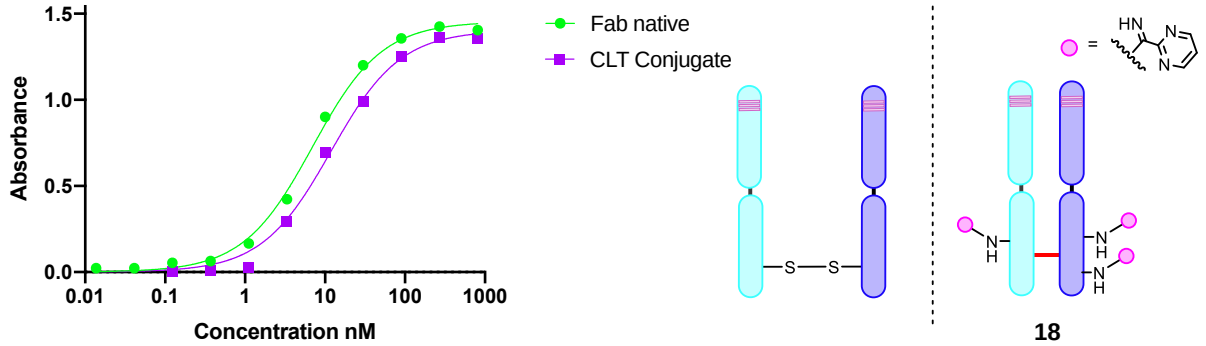


Figure 28. Binding activity to HER2 of trastuzumab native Fab and Fab conjugate **18** over concentration ranges from 810-0.0137 nM, the absorbance was measured at 450 nm.

2.5.1.1. CLT aqueous stability

Due to the amidine nature of the CLT product an investigation into its stability was required to gauge the viability of nitrile CLT for use within bioconjugation. A key aspect is the aqueous stability, to investigate this the CLT conjugate **18** was subjected to incubation at 22 ºC, PBS pH 7.4. Under these conditions, no cleavage or degradation of the conjugate was seen, supporting the idea that the CLT products are stable in aqueous solutions for prolonged periods of time (5 days).

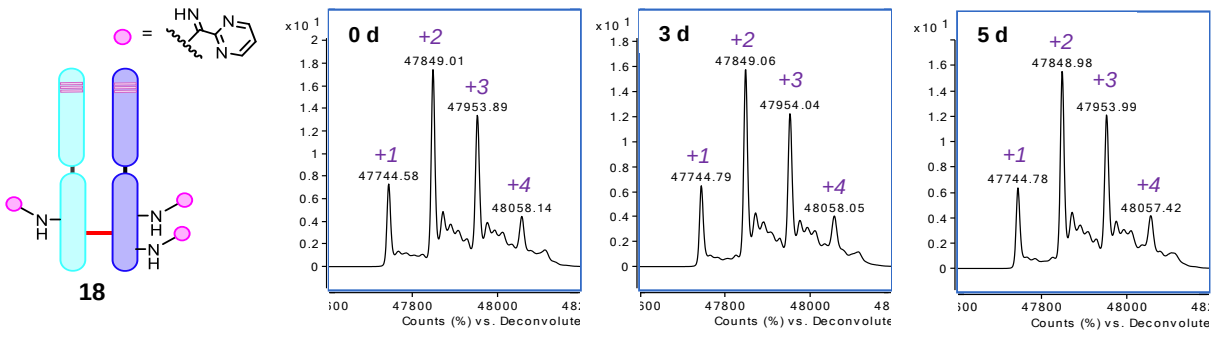


Figure 29. Aqueous stability of the 2-cyano pyrimidine CLT conjugate **18** over 5 days in pH 7.4 conditions at 22 ºC.

The conjugate was also subjected to mimic conditions for relevant physiological environments, namely early endosomal and serum conditions. The specific nucleophile in question was the

abundant peptidic monothiol; glutathione (GSH). The early endosome has high concentrations of GSH (5 mM) and a slightly more acidic pH (pH = 6.5) whereas the blood serum presents lower concentrations of GSH (5 µM) and a higher physiological pH (pH = 7.4). The Fab conjugate **18** was incubated at 37 °C to mimic physiological temperatures for 24 h. Longer timeframes than this are not dependable due to the oxidation of GSH to its respective disulfide over time in buffered conditions. This oxidation by ambient molecular oxygen is stimulated by traces of metals such as copper, iron and nickel and can be limited *via* the inclusion of EDTA into the reaction mixture. EDTA is a well-known chelating agent which chelates divalent cations such as copper, magnesium, and nickel ions, which attenuates the oxidation of thiols.[225,226] The CLT conjugate showed no sign of degradation in both early endosomal or serum mimicking conditions after incubation at 37 °C for 24 h (Figure 30).

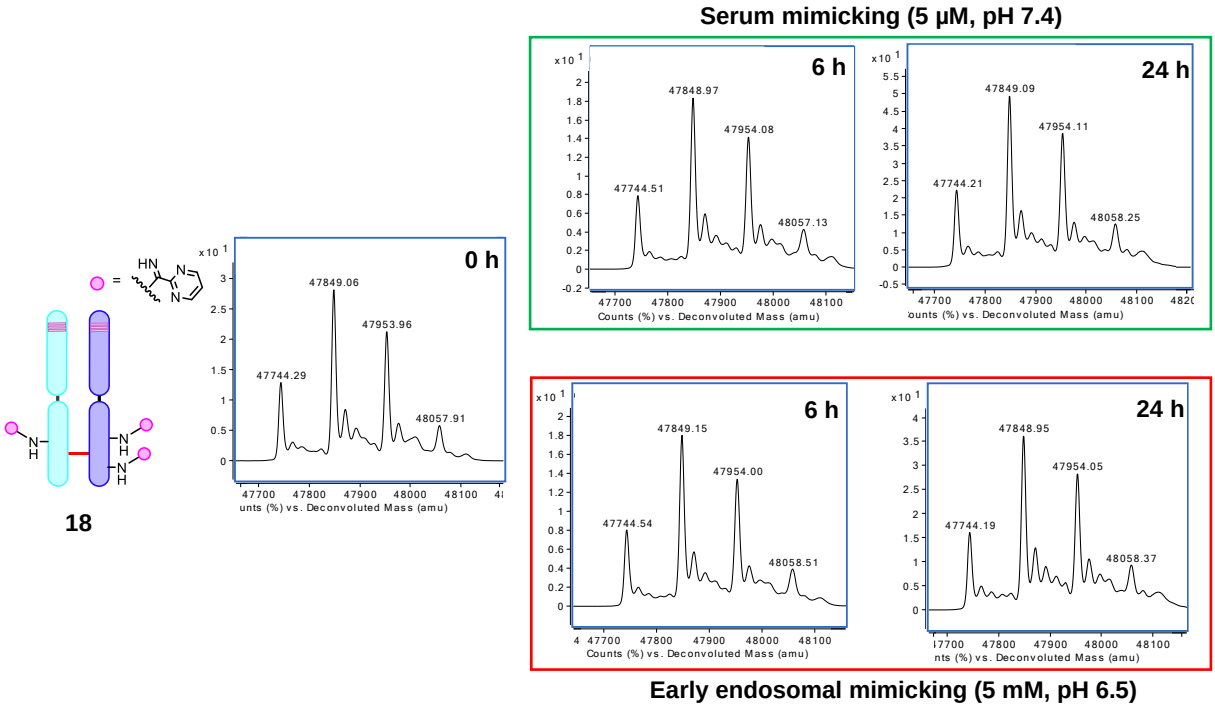


Figure 30. Stability investigation of the 2-cyano pyrimidine CLT conjugate **18**. Serum (5 μM GSH, pH 7.4, 37 ºC) - highlighted with a green box and endosomal (5 mM GSH, pH 6.5, 37 ºC) – highlighted with a red box.

Next the stability of the nitrile CLT conjugate in human serum was considered. A serum stability study was conducted to establish whether the conjugated Fab fragment would participate in thiol exchange with the highly abundant blood thiol HSA. HSA is the most abundant protein in the vascular system with concentrations of 35 - 50 mg/mL.[227] This study required the use of the fluorescent Fab conjugate in order to track the potential transfer to HSA, therefore, the fluorescein functionalised conjugate **20** was employed. The conjugate **20**

(7 µM) was incubated in human serum for 5 days and analysed by SDS-PAGE. The two concentrations of serum tested were that of 50 and 100% in order to acquire a better understanding with regards to the serum stability. The conjugate was found to be stable as no transfer of fluorescence was detected in the HSA region (Figure 31).

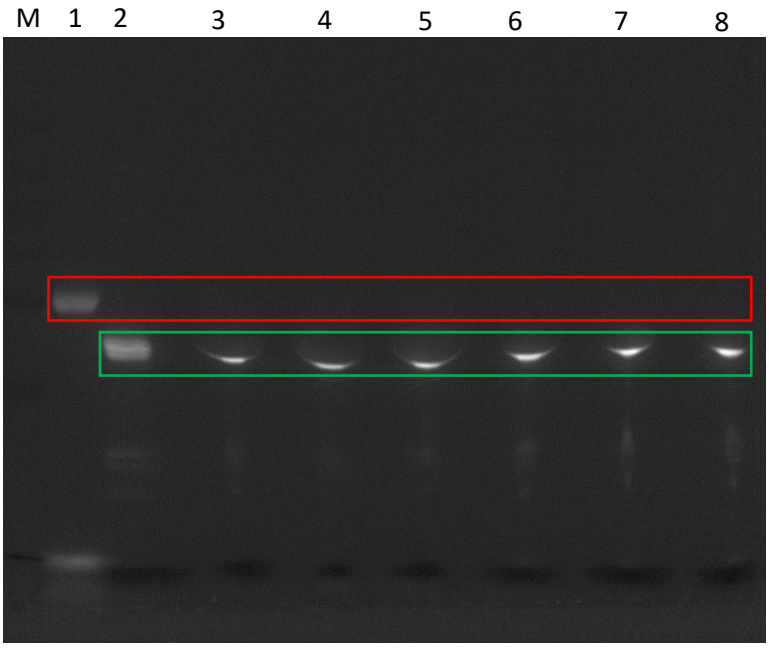
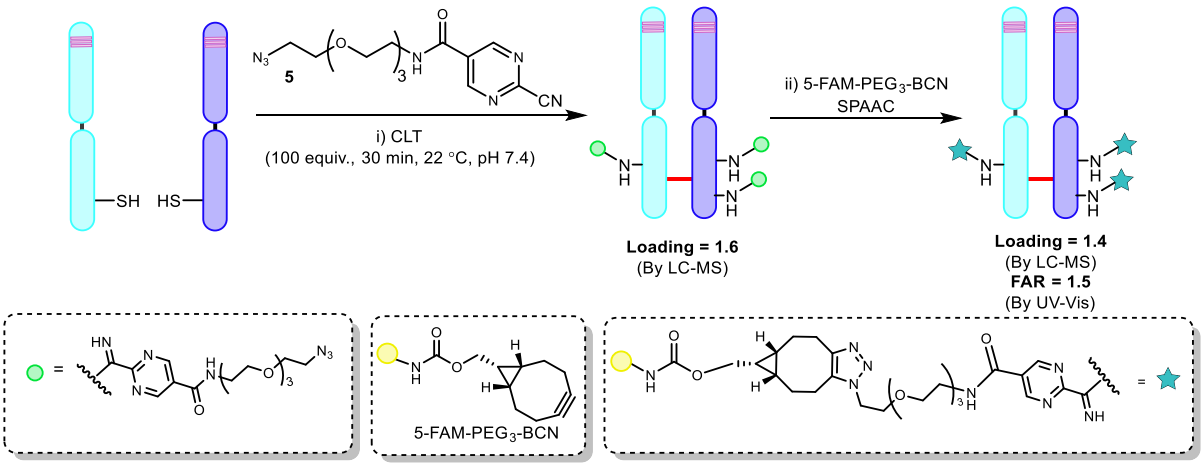


Figure 31. SDS-PAGE of serum stability study; M – molecular marker, 1 – fluorescent HSA reference, 2 – fluorescent bridged Fab reference (conjugate **20**), 3 – Day one (100%), 4 – Day three (100%), 5 – Day five (100%), 6 – Day one (50%), 7 – Day three (50%), 8 – Day five (50%). The red box highlights the HSA band, the green box highlights the Fab band.

2.6. Summary and conclusion of nitrile CLT

In conclusion, these results describe the reactivity, stability, and functionalisation of a 2nd generation of CLT reagents. These reagents provide marked advantages over the 1st generation reagents (thioesters) including hydrolytic stability and short reaction times. Nitrile CLT, unlike with thioesters, occurs spontaneously in a one-pot conjugation. This means that the thioimidate portion of the process is not observed due to the combination of the reverse reaction and transfer to lysine. The *p*-amide pyrimidine nitrile **4** provides quantitative conversion (>95%) to the CLT conjugate with an average DAR of 1.6. A statistical mixture containing +1, +2 and +3 additions was obtained, with the major product being the +1 and +2 species. The total conjugation time from native Fab to the product requires less than 2.2 h of total reaction time whilst in comparison, thioester CLT requires 10 h. The CLT conjugate was shown to be successfully functionalised with a SPAAC reaction with 5-carboxyfluorescein-PEG_3-BCN (Scheme 54).



Scheme 54. Overview for optimised nitrile CLT on trastuzumab Fab.

These reagents were found to be highly tuneable *via* the modulation conditions and also the core heterocyclic ring employed. The introduction of EWGs was found to have a lesser impact on the overall reactivity. The difference in reactivity between the electron withdrawing trifluoromethyl pyrimidine nitrile **6** and the plain pyrimidine nitrile **2** was shown to be not as impactful as hypothesised. A significant increase in reactivity was seen between the pyridine and pyrimidine core however, further development to 1,2,4 and 1,3,5 triazine rings provided a negligible increase in reactivity. The increase in reactivity was hypothesised to aid the 'loading' of the free thiols *via* shift of the equilibrium in flavour of thioimidate formation, however the data indicated this increase failed to provide the desired effect. Achieving a single conjugate species from nitrile CLT is unlikely however due to the reversible nature of nitrile-thiol reactivity however, use of CLT to modify proteins still provides a valuable tool due to their customisability with regards to degree of modification. Variation of reagent design i.e. core heterocycle identity and reaction conditions (e.g. time, temperature, and number of equivalents) provides access to the diversity of nitrile CLT.

Future work primarily includes analysis regarding the specificity of nitrile CLT. Preliminary LCMS/MS analysis has been carried out *via* a tryptic digest; the initial results suggest the modification of the predicted lysine residues proximal to the disulfide bond (HC: K136, K221, K225 and LC: K126, K190). The sample however, lacks compete coverage of the Fab sequence, and more importantly lysine residues. Further LCMS/MS digests are therefore required to confidently assign sites of modification. The library of nitrile reagents developed can also be subjected to a kinetic study between the nitrile in question and cysteine methyl ester. This would provide quantitative rate values, allowing for more informed decisions regarding future use. This system is also yet to be applied and optimised on full antibody

systems alongside other classes of protein, such as Protein disulfide isomerase (PDI) which contains two sets of disulfide bonds. Rat PDI contains a lysine residue adjacent of one of these disulfide bonds (W**C**GH**CK**)[228] and with modification of this lysine, a potential allosteric-type inhibitor could be developed for enzyme activity assays (Figure 32).

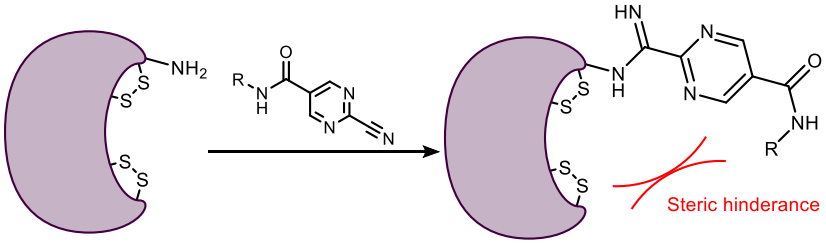
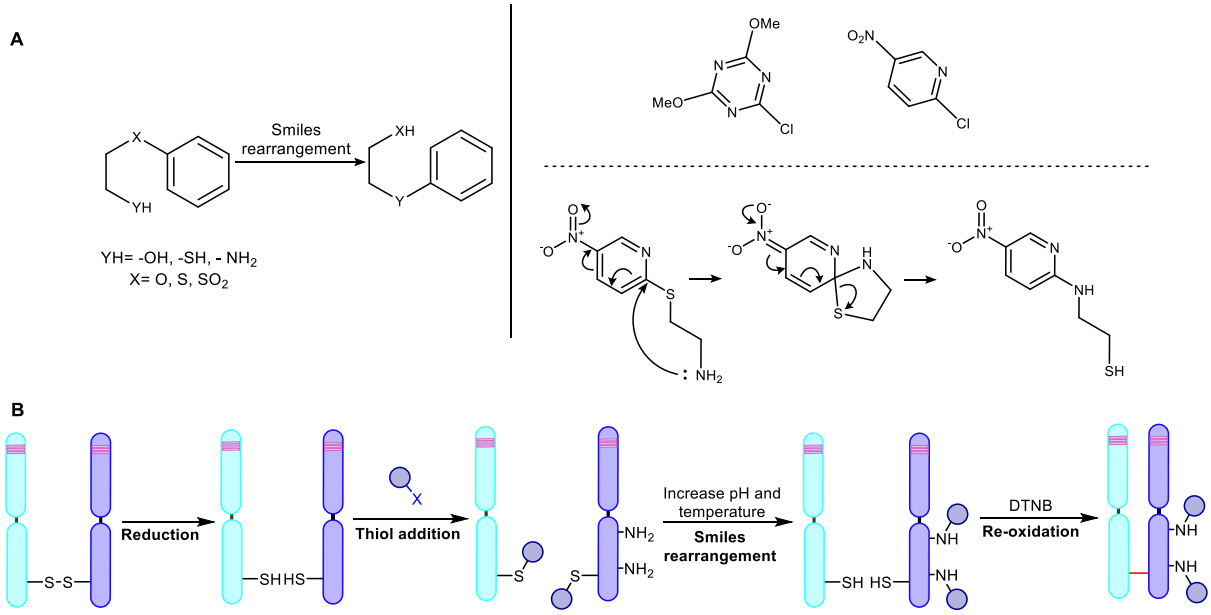


Figure 32. Proposed CLT based allosteric-type inhibitor.

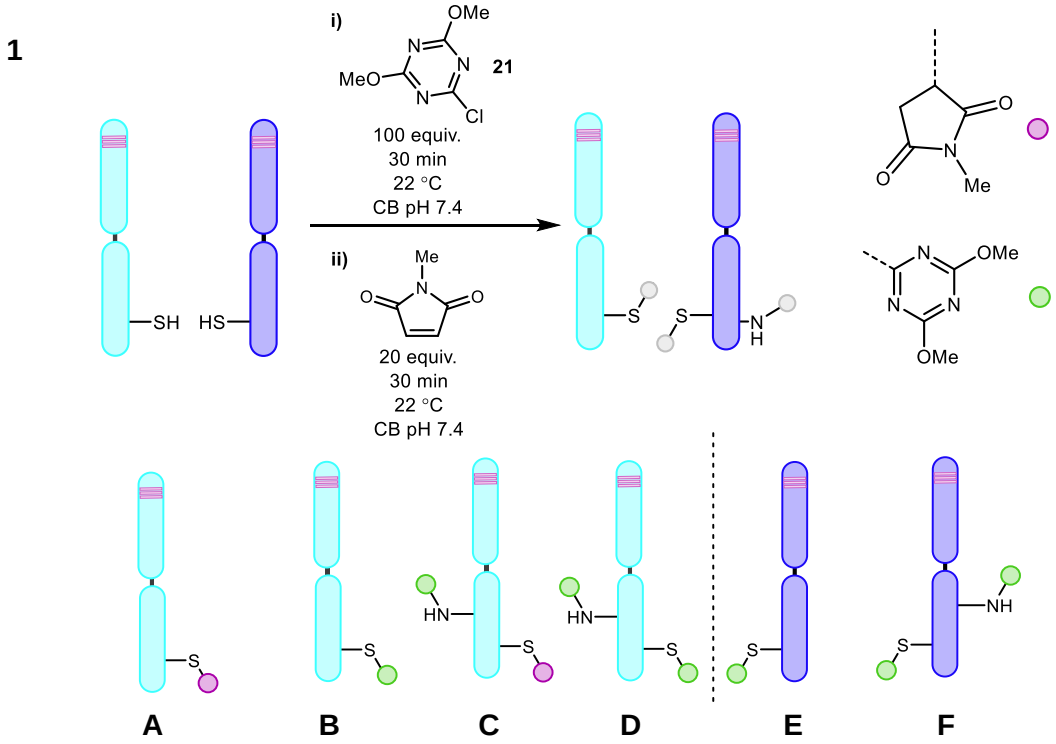
2.7. S_NAr for CLT

As seen with the prior reagents employed for nitrile CLT, S_NAr is a feasible reaction under bioconjugation conditions. This has been supported *via* previous peptide stapling and labelling work.[229,230,231] The hypothesis behind the use of S_NAr in CLT stems from Smiles rearrangements, which involve base catalysed heteroatom migration. These rearrangements have been shown to proceed under acid and heat catalysis. The relevant form of rearrangement, exemplified by Hearn *et al*., is the *S*-to-*N* type Smiles rearrangement. The proposed methodology is similar to that of thioester CLT; 1) reduction of the disulfide bond 2) loading of the free thiols 3) conditions change to promote lysine transfer 4) reoxidation of the disulfide bond. The aryl sulfanyl intermediate will ideally prove greater aqueous stability allowing for greater conversion with similar degrees of control.

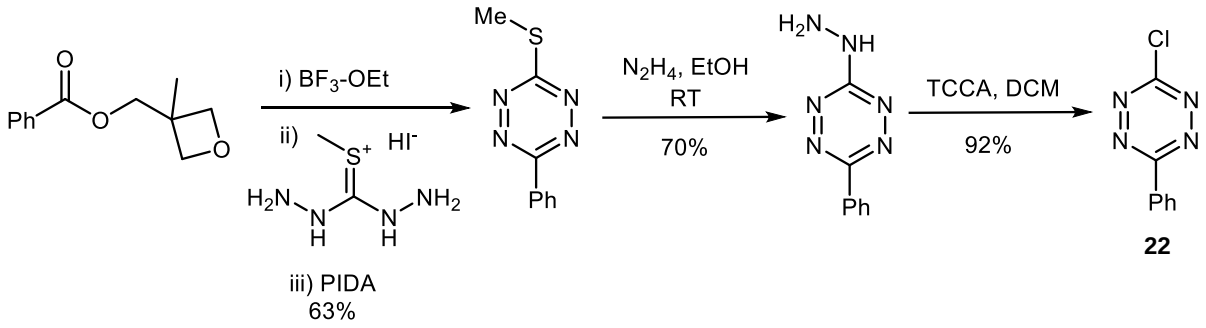


Scheme 55. (**A**) General scheme for *S*-to-*N* type Smiles rearrangement demonstrated by Hearn *et al*. (**B**) Example of S_NAr CLT protocol on trastuzumab Fab.

The first S_NAr reagent explored was the commercially available 2-chloro-4,6-dimethoxy-1,3,5-triazine (**21**) employed by Hearn *et al*. In order to investigate the thiol selectivity of these reagents addition to reduced Fab were carried out. NMM was then added to check the availability of the thiols and therefore the location of S_NAr on Fab (cysteine or lysine). The MS revealed a mixture of cysteine and lysine reactivity with a preference towards cysteine, these additional reactions may have stemmed from the large equivalents of **21** (100 equiv.) added, leading to unselective addition (Scheme 56).

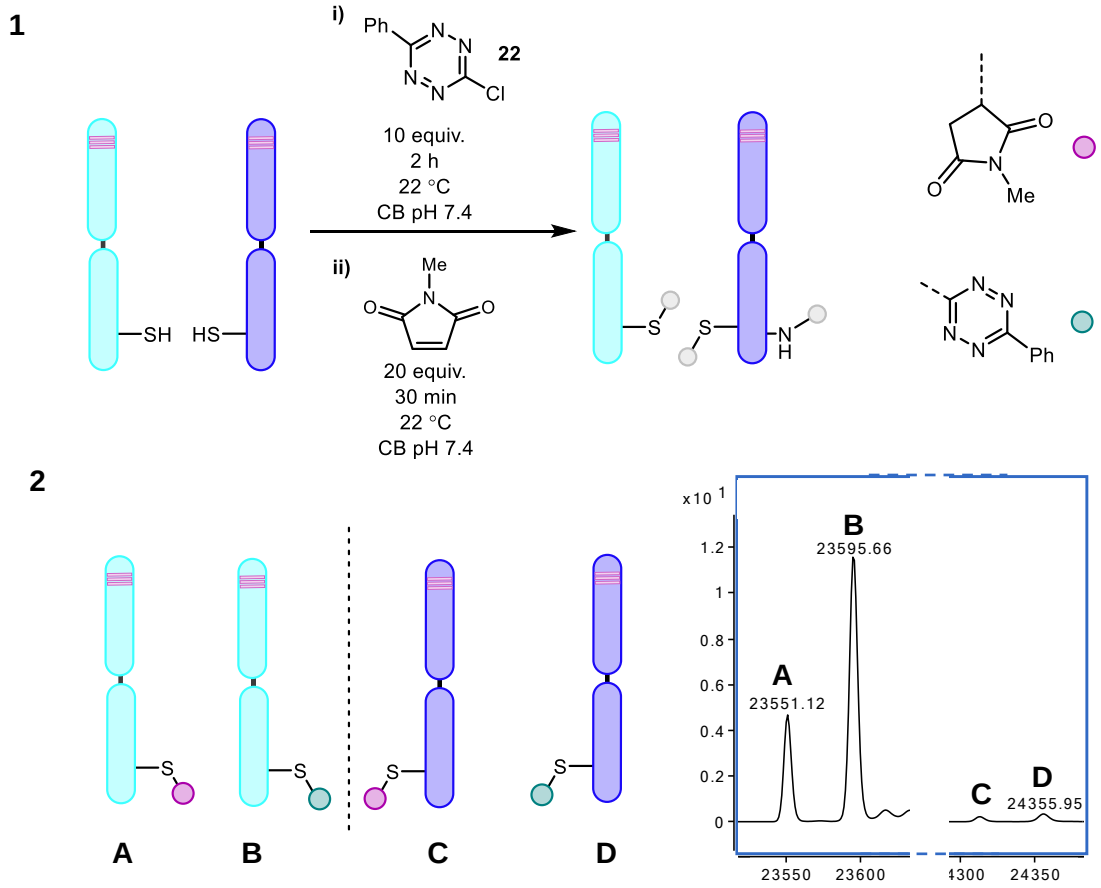


Scheme 56. (1) Reaction scheme for the thiol selectivity of 2-chloro-4,6-dimethoxy-1,3,5-triazine (**21**). Pink sphere represents maleimide, green sphere represents the heteroaromatic ring. **(2)** LCMS of 2-chloro-4,6-dimethoxy-1,3,5-triazine **21** reactions on Fab. 2-chloro-4,6-dimethoxy-1,3,5-triazine, 100 equiv. 30 min, 22 ºC, pH 7.4. NMM 20 equiv., 30 min, 22 ºC, pH 7.4. Mass of native LC = 23440 Da. Mass of native HC = 24201 Da. Expected mass of 2-chloro-4,6-dimethoxy-1,3,5-triazine modification = 139 Da.



Scheme 57. Synthetic overview of **22**.

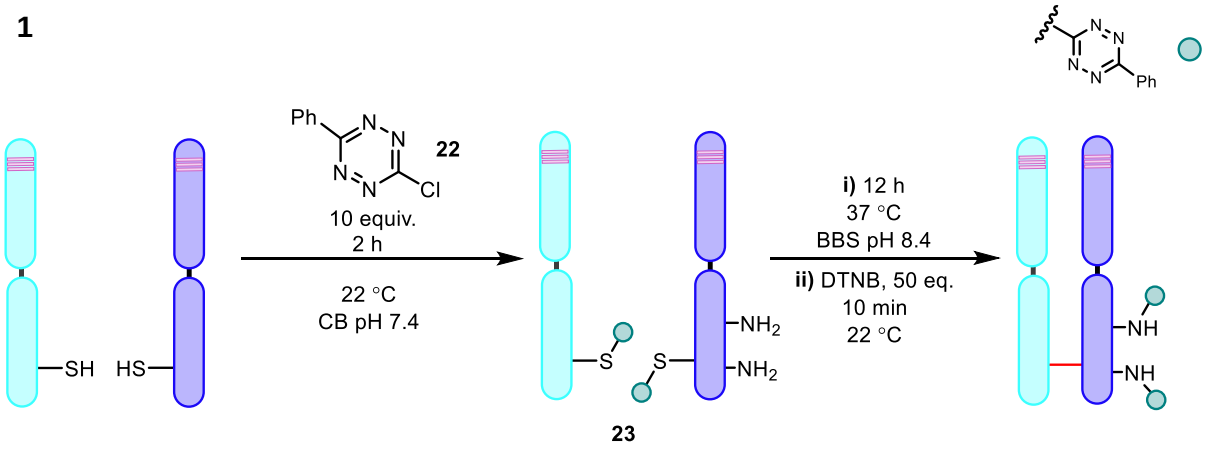
With the assistance of Dr. Strachan a second reagent was synthesised for S_NAr CLT in the form of a carbon substituted tetrazine ring, 3-chloro-6-phenyl-1,2,4,5-tetrazine (**22**). As seen with previous reagents, tetrazine rings are particularly prone to the S_NAr pathway and therefore provide an exciting prospect. When subjecting the tetrazine reagent **22** to analogous conditions to the protocol above, (100 equiv., 30 min, 22 °C, pH 7.4) it was apparent that Fab was intolerant of this as degradation was seen *via* MS. The conditions therefore had to be modulated in order to cope with the new tetrazine reagent, of the few trialled conditions; 10 equiv. for 2 h at 22 °C provided clean addition with preference for the free cysteine thiols (Scheme 58). The reaction is not fully complete as some re-oxidised Fab is seen in the MS trace, therefore equivalents or time will need to be increased.



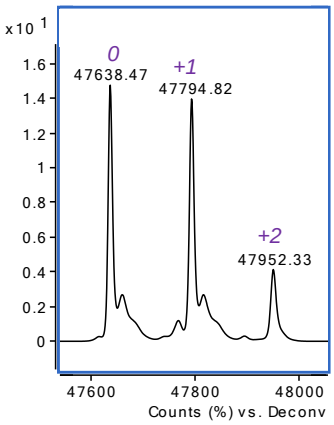
Scheme 58. (1) Reaction scheme for the thiol selectivity of 3-chloro-6-phenyl-1,2,4,5-tetrazine (**22**). The pink sphere represents maleimide, green sphere represents the heteroaromatic ring. **(2)** LCMS of 3-chloro-6-phenyl-1,2,4,5-tetrazine **22** reactions on Fab. 3-chloro-6-phenyl-1,2,4,5-tetrazine, 10 equiv. 2 h, 22 ºC, pH 7.4. NMM 20 equiv., 30 min, 22 ºC, pH 7.4. Mass of native LC = 23440 Da. Mass of native HC = 24201 Da. Expected mass of 3-chloro-6-phenyl-1,2,4,5-tetrazine modification = 156 Da.

Once the modification of the free chain was obtained, the next step was to establish the Smiles rearrangement to attempt to provide the lysine modified conjugate and reform the disulfide bond. In order to promote the rearrangement, the conjugate **23** was buffer-swapped into a higher pH environment (pH 8.4) and left overnight at 37 °C.

1



2



Scheme 59. (1) Reaction scheme for the CLT of 3-chloro-6-phenyl-1,2,4,5-tetrazine (**22**). **(2)** LCMS of 3-chloro-6-phenyl-1,2,4,5-tetrazine **22** reactions on Fab. 10 equiv. 2 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 156 Da. +1, +2 refers to number of additions.

The LCMS trace shows successful re-formation of the disulfide bond with up to +2 additions on native Fab. This result provides the proof of concept for the use of S_NAr as a CLT reagent class (Scheme 59). The conversion, however, is not ideal (57%) with a DAR of 0.71, this method therefore requires further investigation and optimisation in order to obtain the desired conjugates with quantitative modification. A limitation behind the formation of an amnio tetrazine conjugate is the dampening effect the lone pair donation has on a potential iEDDA functionalisation.

2.8. Summary and conclusion of S_NAr CLT

S_NAr provides a new 3rd generation of CLT reagents presenting novel bioconjugation chemistry. This work has provided a basis and proof of concept regarding the viability of S_NAr as a bioconjugation tool. 3-chloro-6-phenyl-1,2,4,5-tetrazine was identified as the first reagent for successful transfer *via* this chemistry however, the conditions for this reaction require development in order to improve the overall conversion.

Future reagent designs still present an immense opportunity for optimisation, as other heterocycles and functional moieties can provide access to useful bioconjugation tools. 2-Chloro-5-nitropyrimidine provides an enticing reagent design due to the presence of the *p*-nitro moiety which as describes earlier is able to stabilise the key Meisenheimer complex formed during S_NAr. The electron-withdrawing ability of the *p*-nitro moiety also provides a more reactive carbon centre, which in turn, is more prone to thiol-mediated nucleophilic attack. As well as differing ring systems, the leaving group itself can also be further investigated. A well-documented leaving group is the methylsulfonyl group, which when compared to the halogen equivalent provided superior thiol reactivity across a pyrimidine series put together by Baud *et al*.[219] Novel designs which incorporate the highly attractive tetrazine ring can also be investigated, which allow the use of iEDDA chemistry without the dampening effect of amino substituents.

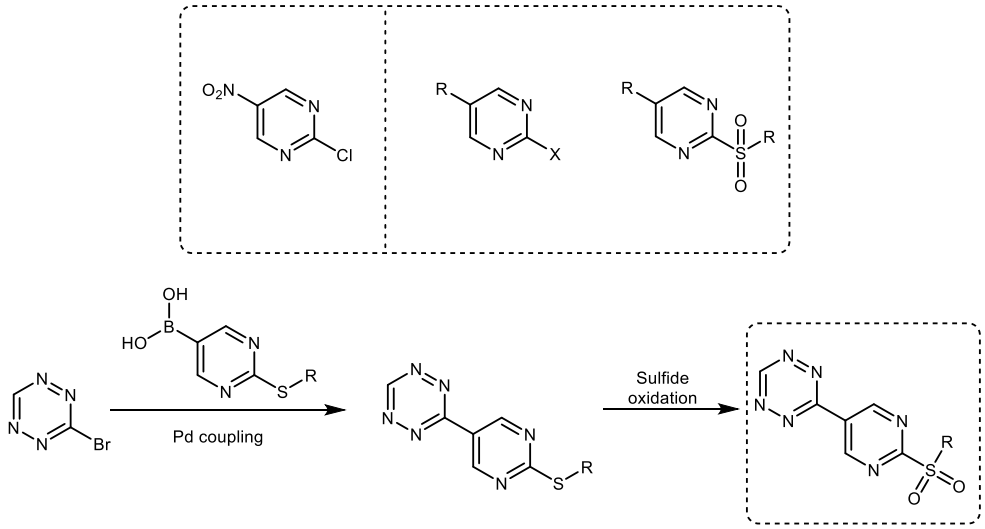


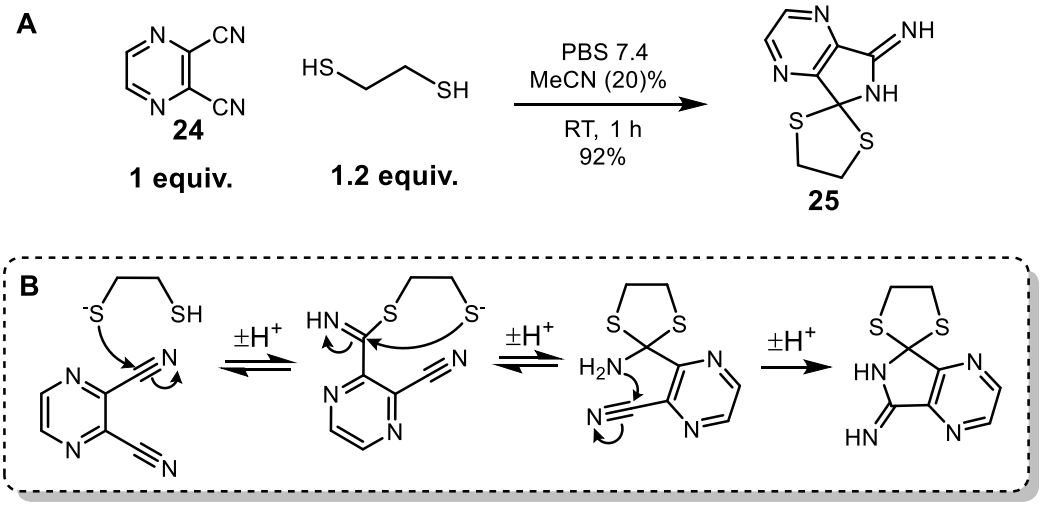
Figure 33. Future S_NAr reagent designs.

3. Chapter 3: Nitrile Disulfide Bridging Reagents

This chapter aims to develop a successful disulfide bridging reagent class. These reagents will aim to rapidly and efficiently re-bridge two proximal thiols, the resulting bridge will look to be stable under physiological conditions. As with the CLT reagents, this new class of disulfide bridging reagent will also look to incorporate functional handles for the addition of relevant payloads. The application to multiple protein systems is a key aim to expand the use of these reagent in widespread protein modification.

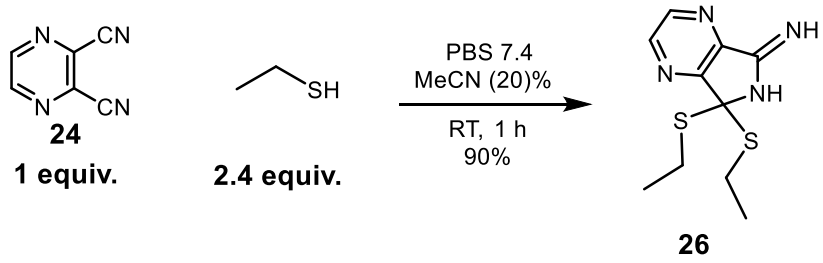
3.1. Bis Nitrile Bridging

To advance the chemistry behind the nitrile bis thiol (NBT) reaction for wider applications, a second class of nitrile-based reagents were envisioned. This new class of reagents incorporate an electrophilic trap for the amino group formed in the amino dithioacetal (ADTA) intermediate. Pyrazine-2,3-dicarbonitrile **24** was identified as a core to the design due to its symmetry and commercial availability. As seen in chapter 2, a key part of the project revolves around the understanding of how these reagents react on the protein systems. Pyrazine bis-nitrile **24** was therefore also reacted with ethanedithiol under buffered conditions to yield the nitrile bis-thiol (NBT) reaction followed by the hypothesised 'trapping' of the ADTA intermediate. As with the mono-nitrile reaction, this reaction provided an impressive yield of the respective ADTA, **25** (Scheme 60 A). In general, amino-nitrile reactivity required for the trapping of the ADTA intermediate, requires long reaction times or harsh conditions. However, the proximity and positioning of the reactive groups provides the appropriate driving force for the final cyclisation (Scheme 60 B). This chapter looks to successfully develop a new class of disulphide bridging reagents which can be applied to a variety of systems.



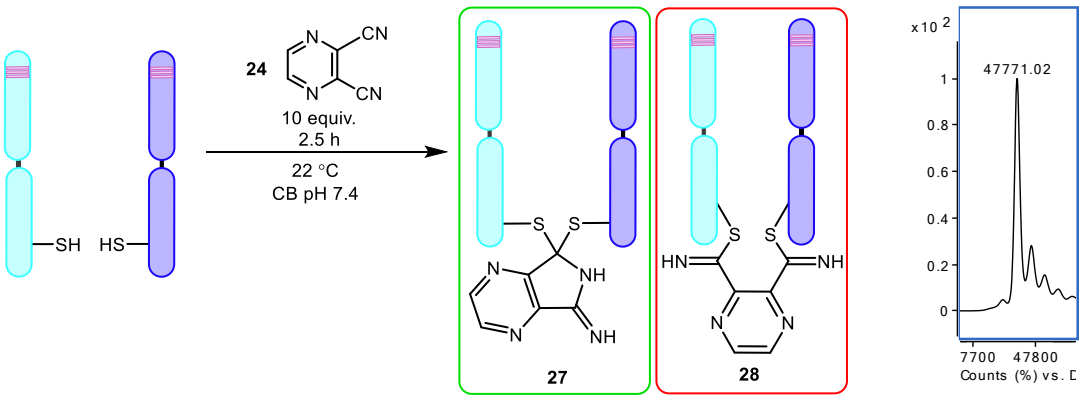
Scheme 60. (**A**) Pyrazine-2,3-dicarbonitrile-EDT NBT reaction for the synthesis of ADTA **25**. (**B**) Proposed mechanism of disulfide ADTA trapping.

The NBT reaction with the incorporation of an electrophilic trap was also able to undergo an analogous trapping pathway with mono-thiols, as seen in the reaction with EDT (Scheme 61). This is an exciting novel reaction class in which mono-thiols can be trapped by nitriles, which is not possible normally, due to the reversible nature of thioimidate formation. The new design promoted the proposed hypothesis in which nitriles in the presence of a second electrophilic group could form stable NBT conjugates, and this would represent a new 1-carbon disulfide bridging strategy. In comparison to other bridging reagents such as maleimides or pyridazinediones which have multiple carbon atoms in the bridge, the one carbon strategy may provide the prospect of less disruption to the tertiary structure. This is due to the shorter total bond distance, which is more comparable to the native disulfide bonds (~2.04 Å).[232]



Scheme 61. Pyrazine-2,3-dicarbonitrile-ET NBT reaction, synthesising ADTA **26**.

Pyrazine-2,3-dicarbonitrile **24** was then applied to the Fab fragment in order to determine if the small molecule studies would translate on to a protein model. Control experiments on non-reduced Fab (10 equiv., 2.5 h, 22 ºC, pH 7.4) demonstrated no unselective reactivity. Reaction with reduced Fab resulted in the full clean bridging of the disulfide bond in just 2.5 h (Scheme 62).

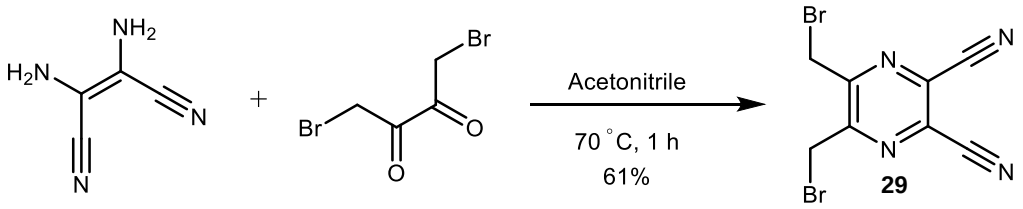


Scheme 62. (A) Bioconjugation conditions employed for pyrazine-2,3-dicarbonitrile on reduced Fab. The green box represents the trapped species akin to **26**. The red box represents a bis-thioimidate product (B) LCMS of pyrazine-2,3-dicarbonitrile reactions on Fab. pyrazine-2,3-dicarbonitrile, 10 equiv. 2.5 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 130 Da.

The bridged conjugate **27** demonstrated excellent stability towards the addition of BME (100 equiv., 2 h, 37 °C), supporting the ADTA structure as the bis-thioimidate (**28**) is likely to cleave *via* thiol exchange under these conditions. Addition of further equivalents of TCEP (10 equiv., 1 h, 37 °C, pH 7.4) and then NMM (10 equiv., 1 h, 22 °C, pH 7.4) also showed a lack of change to the conjugate indicating a lack of lysine transfer and the confirmation of a disulfide bridge. Conjugate **27** showed no change upon incubation in buffered conditions for 7 days (22 °C, pH 7.4) showing hydrolytic stability, the disulfide bridge was found to remain intact after 7 days as no transfer had occurred confirmed *via* an attempt to reduce and cap the disulfide bond.

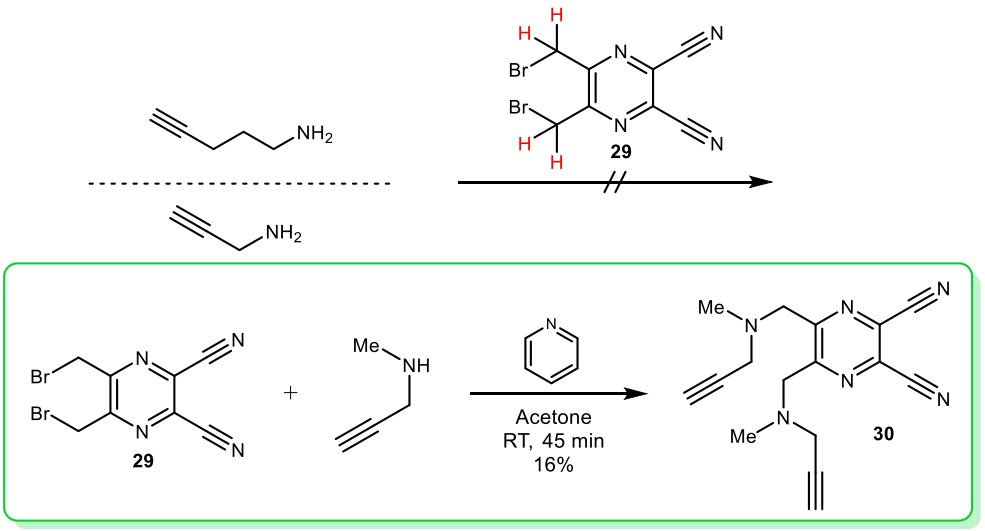
3.1.1. Pyrazine-based Disulfide Bridging reagents

3.1.1.1. Chemical synthesis



Scheme 63. Synthesis of 5,6-bis(bromomethyl)pyrazine-2,3-dicarbonitrile **29**.

As with the CLT reagents, the pyrazine core was then advanced *via* the introduction of functional handles. **29** was synthesised using a literature protocol in moderate yield, from commercially available starting materials diamino malononitrile (DAMN) and 1,4-dibromobutane-2,3-dione, which undergo a condensation reaction to produce the pyrazine ring (Scheme 63). This provided a starting scaffold for modification *via* S_N2 at the 5 and 6 position however, upon testing of commercially available alkyne amines at varying equivalents, temperatures, and timeframes, a clear intolerance towards basic conditions was identified. The benzylic hydrogens were thought to be acidic, and use of primary amines led to full degradation of starting material. *N*-methyl propargylamine provided the final species in poor yield (16%) generating the functional bis-nitrile bridging reagent, **30**. The presence of the dual alkyne handles presents the opportunity for additional loadings of either fluorophore or drug molecules which can be an exciting prospect for dual-functionalised conjugates.

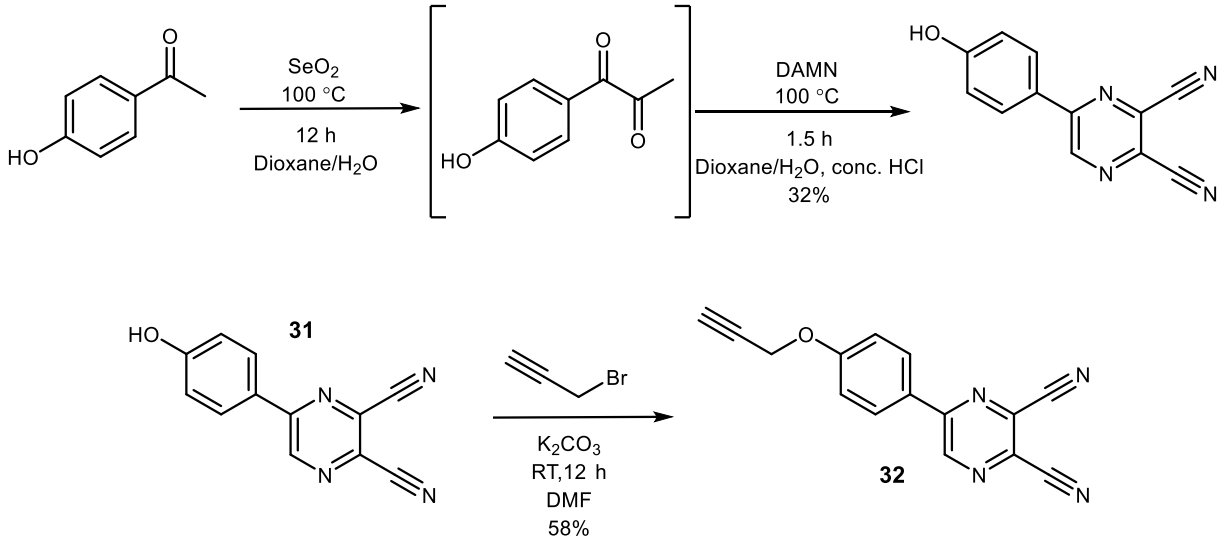


Scheme 64. Synthesis of 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile.

Nucleophile (Nu^-)	Time (min)	Base	T(°C)	Nu^- equiv.	Yield (%)
Propargyl amine	45	NEt_3	RT	2.2	0
Propargyl amine	60	NEt_3	RT	1.0	0
Propargyl amine	60	Pyridine	0	2.2	0
Pent-4-yn-1-amine	60	Pyridine	0	2.2	0
N-methyl propargylamine	45	Pyridine	RT	2.2	16

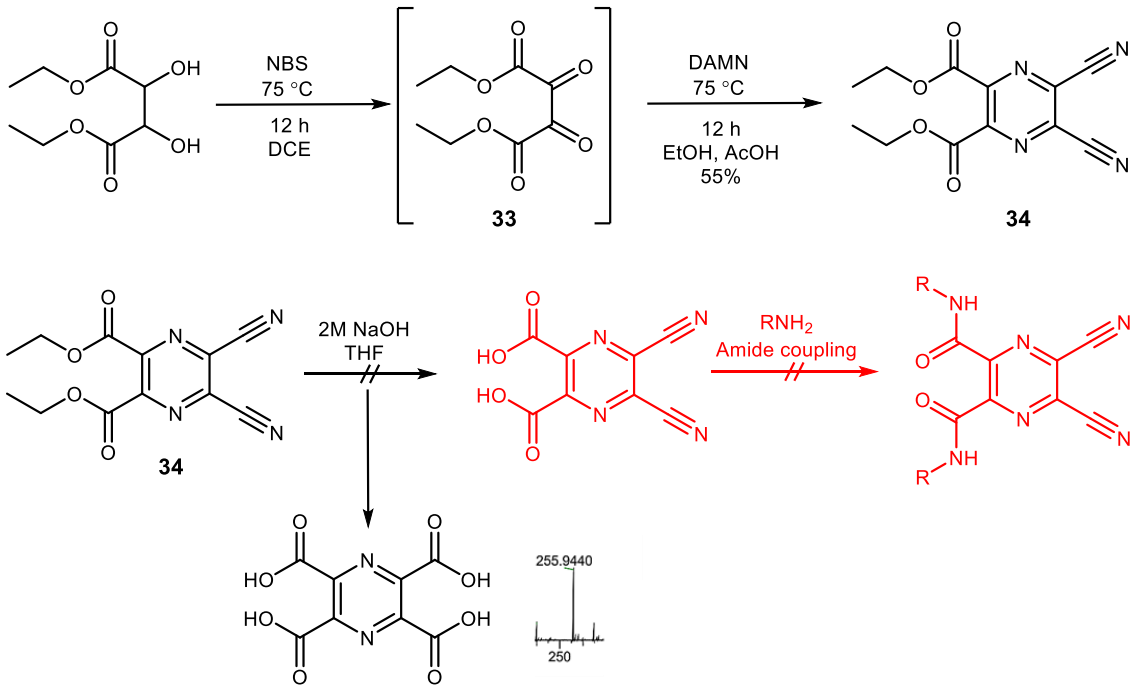
Table 1. Investigated conditions for the synthesis of functional bis-nitrile bridging reagent **30**.

Subsequent designs of pyrazine based bridging reagents looked to avoid the incorporation of the labile benzylic hydrogens as in **30**. One of these designs utilised a phenol moiety as a site for further attachment. **32** was synthesised in a one-pot manner starting with the Riley oxidation of 4'-hydroxyacetophenone to synthesise the 1,2-dione intermediate, **31** which was converted to the pyrazine scaffold, as before, using a DAMN self-condensation. The hydroxyl group was then functionalised through S_N2 utilising 3-bromoprop-1-yne as the chosen electrophile (Scheme 65).



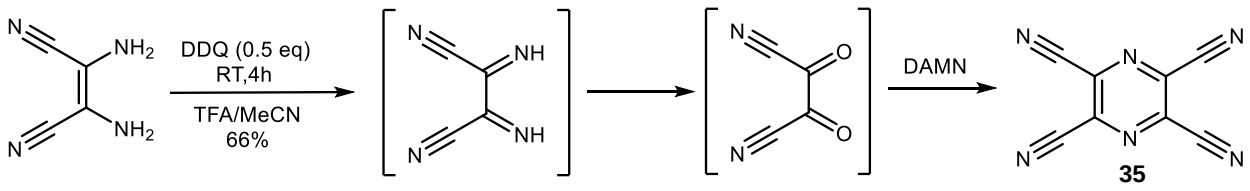
Scheme 65. Synthesis of 5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile **32**.

The next design looked to incorporate both a functional handle and exert an electronic effect on the core pyrazine ring. **34** incorporates two ester groups in the 5 and 6 positions of the pyrazine ring, each with a Hammett constant (σ_p) of 0.45. The synthesis of **34** started with the *N*-bromosuccinimide (NBS) mediated oxidation of diethyl 2,3-dihydroxysuccinate to generate the intermediate **33** which underwent a self-condensation with DAMN to provide **34**.[233] Due to the hydrolytic nature of ester moieties, they are not desirable for bioconjugation reactions. Current attempts of amide synthesis have failed to yield any product, due to the likely hydrolysis of both ester and nitrile groups under strongly basic conditions (Scheme 66). Despite the lack of functionality, **34** is still a molecule of interest for investigation into the impact of the electronic effects provided by the two ester moieties.



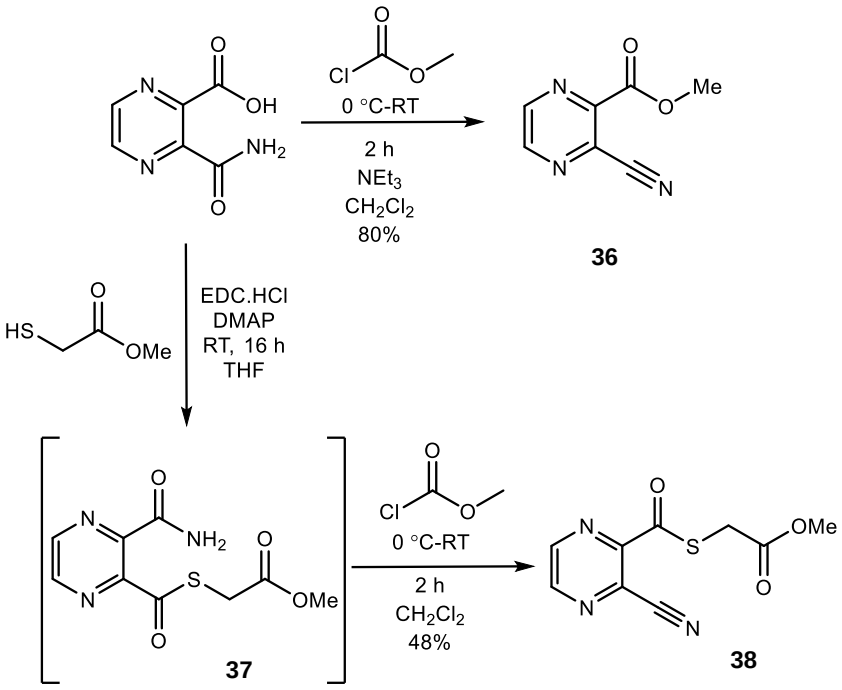
Scheme 66. (A) Synthesis of diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate **34**. (B) Attempted functionalisation of diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate. Structures in red highlight an unsuccessful functionalisation pathway.

A tetra-nitrile reagent was also synthesised, and this provided an even greater electronic effect on the pyrazine ring due to the two highly electron-withdrawing, nitrile groups with a Hammett constant ($\sigma_p = 0.66$) ~1.5-fold that of the ester moieties present in **34**. Unlike in **34**, the nitrile handles provide potential points of direct functionalisation *via* a nitrile-aminothiol (NAT),[234] or 2-cyanobenzothiazole (CBT),[235] click reaction. **35** was synthesised in a one-pot manner, in which DAMN was oxidised to diiminosuccinonitrile (DISN) with 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ). Controlled hydrolysis of DISN produced oxalyl dicyanide *in situ* which underwent a self-condensation reaction with the remaining DAMN to produce **35** (Scheme 67).[236]



Scheme 67. Synthesis of pyrazine-2,3,5,6-tetracarbonitrile.

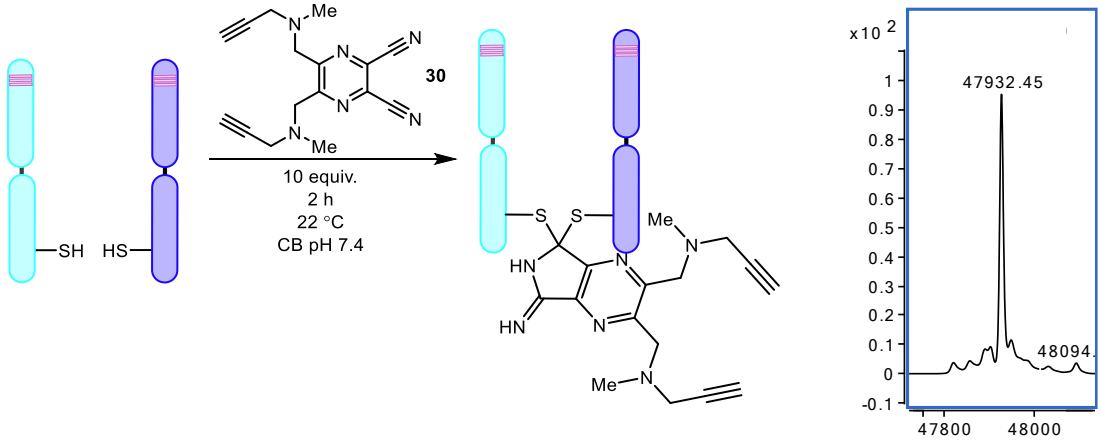
In addition to introducing functionality, additional pyrazine-based reagents were synthesised in order to determine the impact of the design on reactivity. Due to the dynamic nature of the NBT reaction, it was hypothesised that introduction of alternate trapping groups could provide additional degrees of control over the bridge formation. Two examples were synthesised with a nitrile-ester (**36**) and nitrile-thioester (**38**) combination respectively. **36** was synthesised *via* the simultaneous dehydration and esterification of pyrazine-2,3-dicarboxylic acid monoamide.[237] **38** was synthesised *via* an EDC coupling reaction, forming an activated intermediate which was successfully substituted with methyl thioglycolate. This intermediate (**37**) was then telescoped to the dehydration of the amide moiety to yield the final nitrile-thioester, **38** (Scheme 68).



Scheme 68. Synthesis of methyl 3-cyanopyrazine-2-carboxylate **36** and methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate **38**.

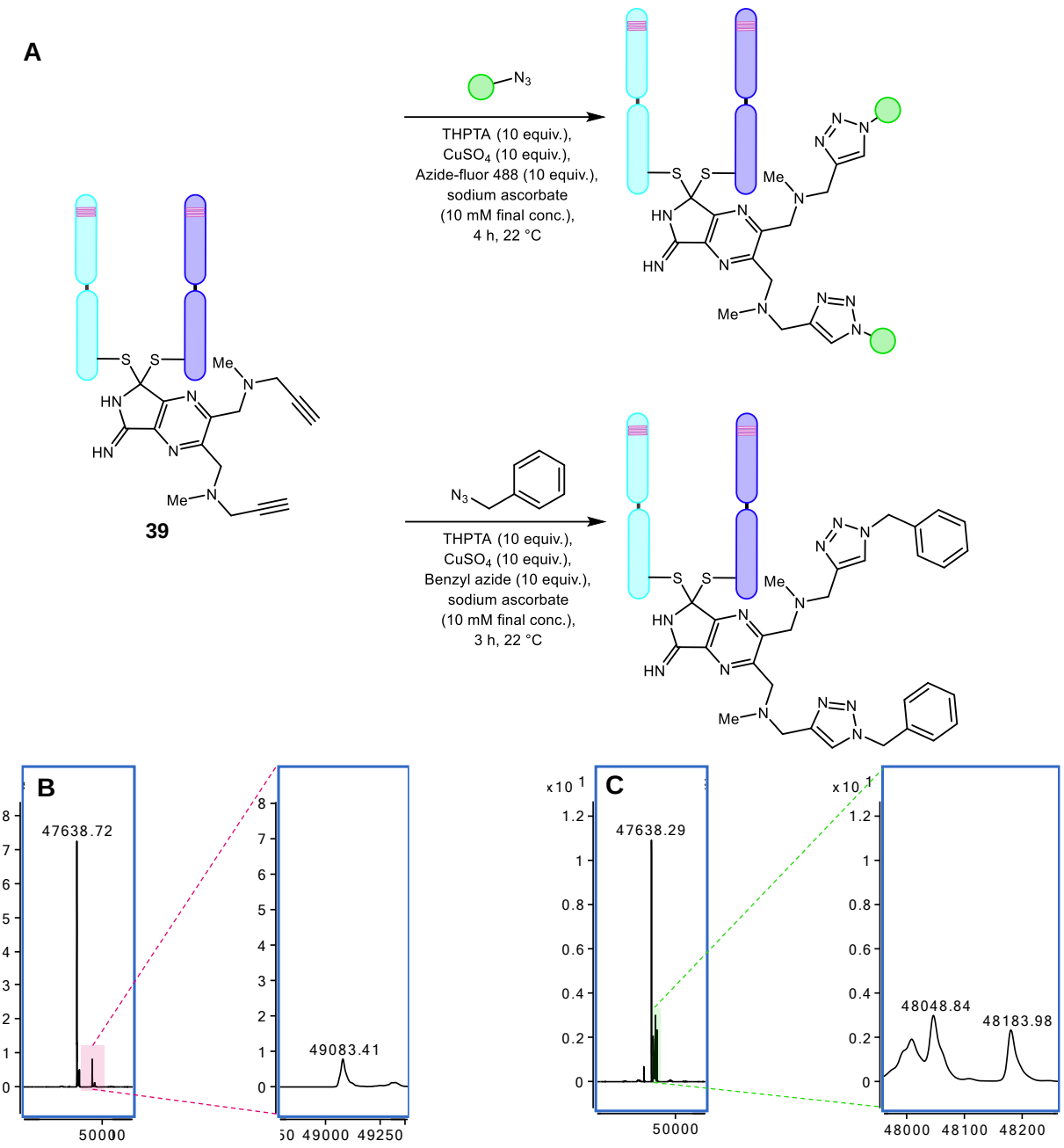
3.1.1.2. Bioconjugation

As with pyrazine-2,3-dicarbonitrile, the bioconjugation of **30** was investigated on reduced Fab, the conditions utilised (10 equiv., 2 h, 22 ºC, pH 7.4) resulted in a clean bridged Fab conjugate. The bridged species was also found to form in the timespan of 1 h, however, to ensure complete conversion to the final species the 2 h reaction was preferred (Scheme 69). A control reaction (10 equiv., 2 h, 22 ºC, pH 7.4) confirmed the lack of undesired additions when added to native Fab.



Scheme 69. (A) Bioconjugation conditions employed for 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile **30** on reduced Fab. (B) LCMS of 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile on Fab. 10 equiv. 2 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 292 Da.

The introduction of alkyne handles allows for the use of click chemistry, specifically Copper-Catalysed Azide–Alkyne Click Chemistry (CuAAC). CuAAC is a stepwise cycloaddition between a terminal alkyne and azide forming a 1,2,3-triazole ring, and these reactions can be rapidly accelerated *via* the use of a Cu(I) catalysis.[238] These reactions are also compatible with aqueous conditions and can therefore be applied to bioconjugation with great success.



Scheme 70. (A) CuAAC conditions employed for Fab conjugate **39**. (B) LCMS of Azide-fluor 488 CuAAC on Fab conjugate **39**. Mass of native Fab = 47638 Da. Expected mass of addition = 292 Da + (2 X 574 Da). (C) LCMS of benzyl azide CuAAC on Fab conjugate **39.** Expected mass of addition = 292 Da + (2 X 133 Da).

The functionality was initially investigated using Azide-fluor 488 (AF488). When the conjugate **39** was placed under CuAAC conditions, the conjugate appeared to cleave in a very clean manner yielding native Fab and a small quantity of the desired dual-clicked species (10%) with no clearly visible LC or HC. This was then expanded with the use of the simpler benzyl azide click partner, allowing for the clarification of the prior result. A similar profile was seen as with

the use of Azide-fluor 488, the major species seen was native Fab, alongside the presence of mono and dual-click products (48049 and 48184 Da respectively). Additional peaks can be seen in the range of 47800 - 48000 Da however, these do not correspond to any logical species and are assumed to be due to copper disruption during sample ionisation. This led to the proposed that the bridge was unstable under CuAAC conditions (Scheme 70), and further small molecule studies with **30** established an intolerance to copper/ascorbate conditions. This mechanism of release may be akin to that of copper mediated deprotection of thiazolidine derivatives, the rate of which appears to be greater than that of the CuAAC.[239] This cleavage provides the enticing prospect of high concentration metal ion triggered release for the controlled triggered cleavage of pyrazine bis-nitrile NBT conjugates. The sulfur atom is thought to coordinate to the copper ions as a Lewis acid, followed by a cascaded removal of the bridge. Copper mediated disulfide reformation inhibits the reformation of the bridged species via the reformation of native Fab (Figure 34).

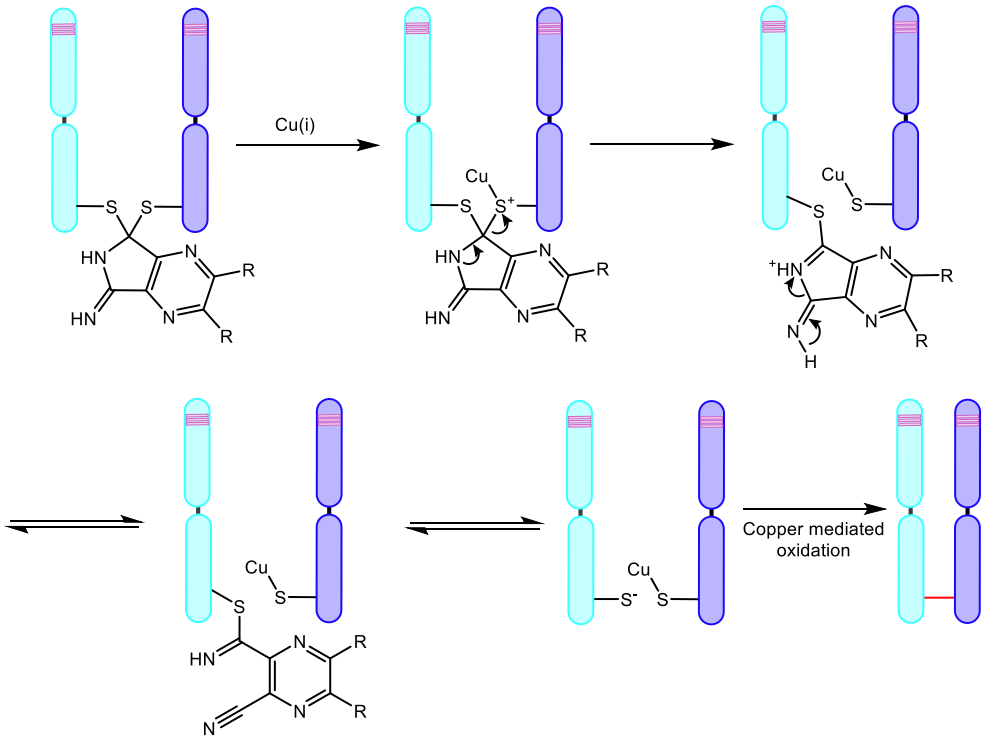
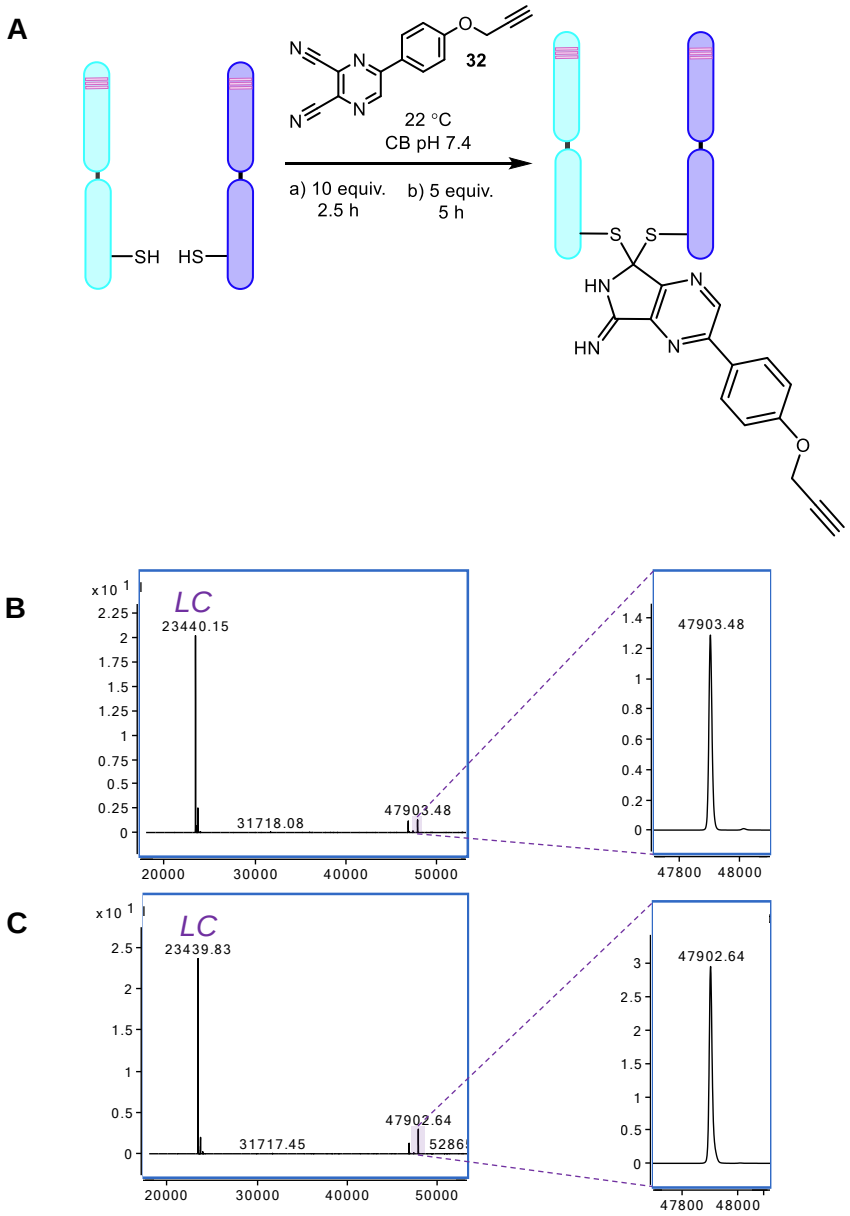


Figure 34. Proposed mechanism of copper mediated NBT cleavage on conjugate **39**. Co-ordination of the sulfur atom in the ADTA ring to copper ions, as a Lewis acid, results in the opening of the ADTA ring releasing the bis-nitrile reagent.

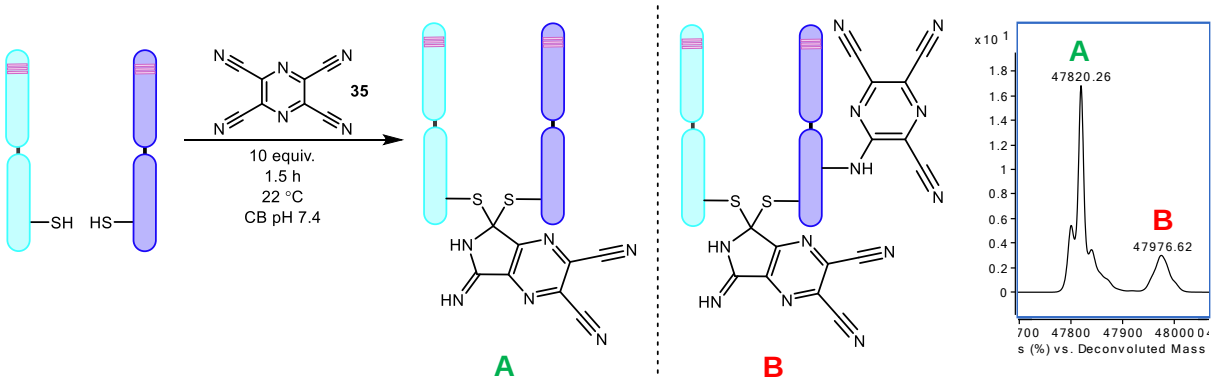
Conjugation attempts with **32** yielded minor modification under conditions similar to previous (10 equiv., 2.5 h, 22 ºC, pH 7.4). The reagent was found to be poorly soluble in aqueous conditions leading to incomplete conversion to the desired bridged product. Lower equivalents for a longer timeframe provided marginally improved conversion (5 equiv., 5 h, 22 ºC, pH 7.4).

Both LCMS traces show significant quantities of native LC and due to solubility limitations, the phenyl-pyrazine design was not pursued further (Scheme 71).



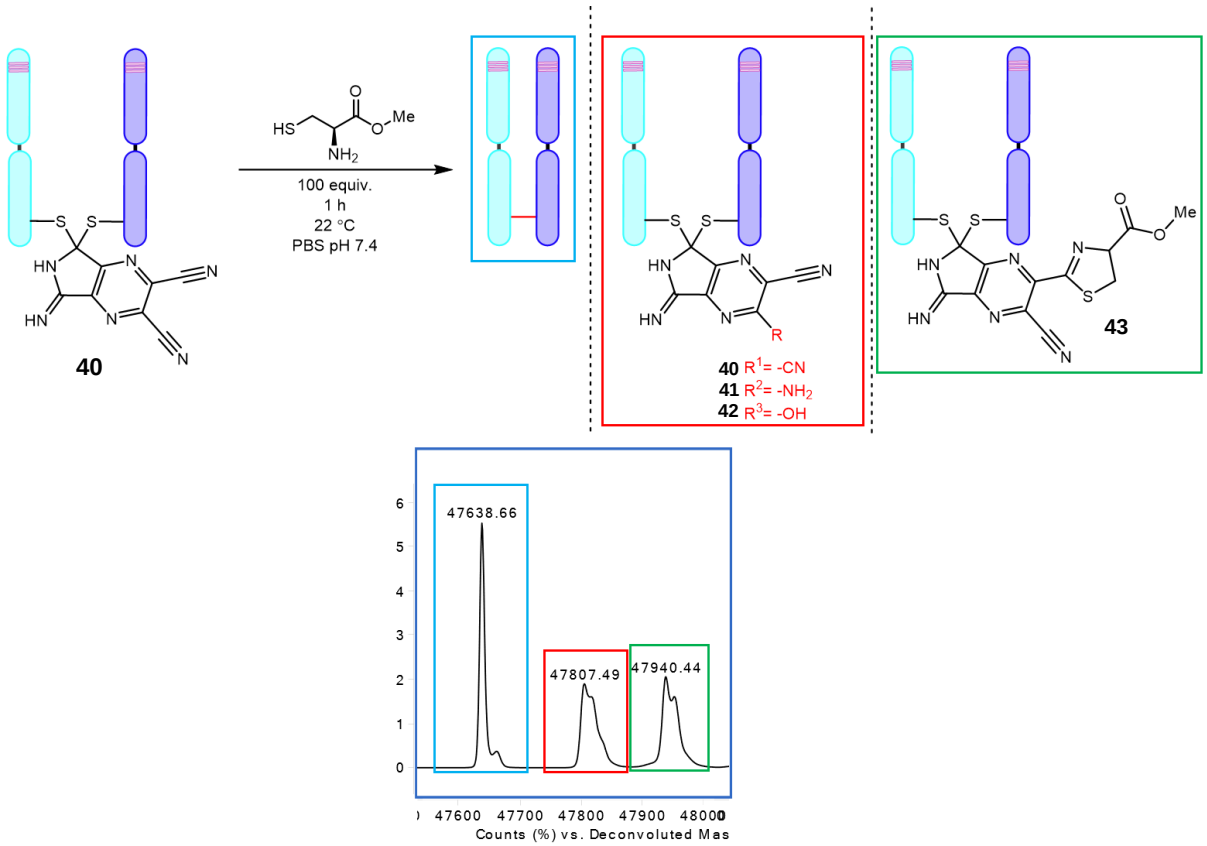
Scheme 71. (A) Bioconjugation conditions employed for **32** on reduced Fab. (B) 10 equiv., 2.5 h, 22 ºC, pH 7.4. (C) 5 equiv., 5 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 260 Da.

As mentioned earlier the bridging reagents, **34** and **35** both provide a unique insight into the effect of EWGs on the pyrazine system. Both reagents provided clean bridges under the conditions employed (10 equiv., 22 °C, pH 7.4). **34** provided slightly faster reaction times (1 h) over the tetra nitrile species **35** (1.5 h).



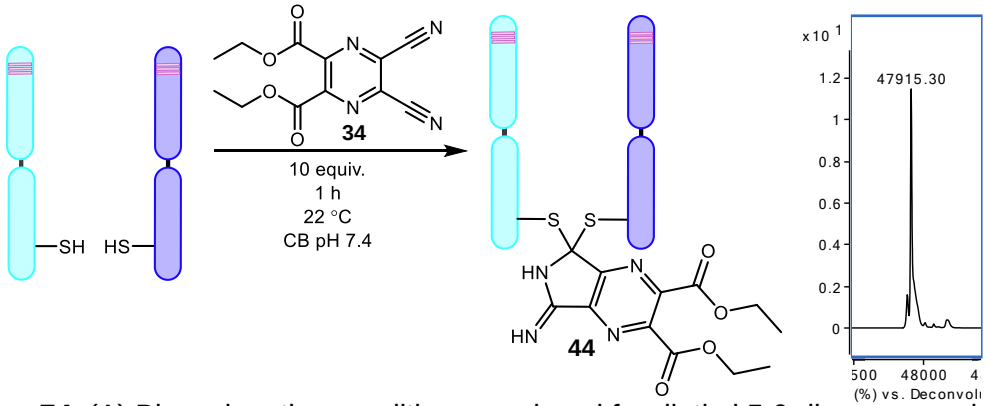
Scheme 72. (**1**) Bioconjugation conditions employed for pyrazine-2,3,5,6-tetracarbonitrile on reduced Fab. **A** represents the bridged species, **B** represents the bridged species with an S_NAr addtion. (**2**) LCMS of pyrazine-2,3,5,6-tetracarbonitrile on Fab. 10 equiv., 1.5 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 180 Da.

35 underwent S_NAr of a cyano group in addition to the formation of the bridge system, resulting in unselective addition, as seen by the mass 47977 Da in the LCMS trace (Scheme 72). Despite seeing signs of unselective addtion, the viability for this reagent as a functional tool was investigated. The conjugated system was reacted with high equivalents of cysteine methyl ester (100 equiv.) in order to push the potential NAT reaction. The resulting reaction generated a mixture of species, the major species was found to be native Fab, with no signs of LC or HC present in the LCMS. The LCMS also indicates the retention of the unmodified bridge **40** (47819 Da = R^1) however, this is buried under the peak at 47807 Da, this new species is 13 Da smaller than that of **40**. This could be the result of S_NAr by the ammonia molecule released upon the competion of the NAT reaction (**41**), or S_NAr mediated hydrolysis (**42**). It is not possible to fully conclude the exact idntity of this species. However due to the aqueous nature of bioconjugation, it is lilkely that hydrolysis is the outome of this reaction. The final species (**43**) at 47940 Da is that of one addition of cysteine methyl ester, while this is the desired species, the conversion is poor (22%) and will require further optimisation (Scheme 73). A key outcome of this reaction is the relativity clean cleavage of the bridge system to yield native Fab with reasonable conversion (59%), this method of cleavage will be further explored in this chapter.

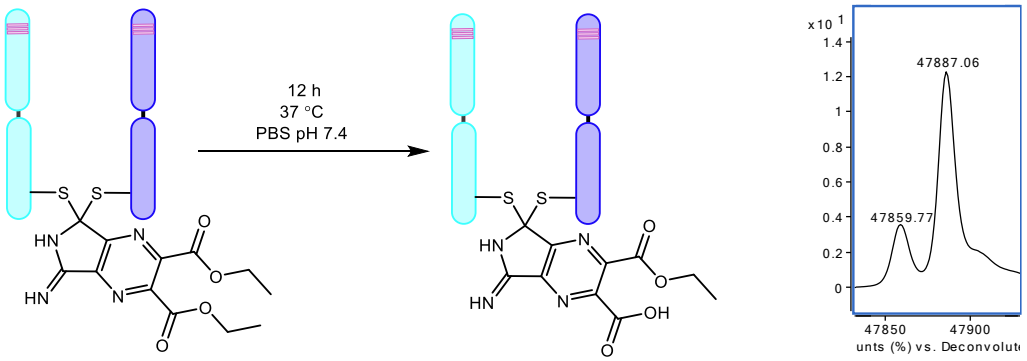


Scheme 73. (A) Conditions employed for NAT modification on Fab conjugate **40**. (B) 100 equiv., 1 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of Cys-OMe addition = 135 Da.

34 provided a clean bridged conjugate **44** in 1 h (Scheme 74), when incubated overnight at 37 °C the ester groups underwent partial hydrolysis as predicted (Scheme 75). This demonstrates the incompatibility of ester moieties for stable protein modification; however, this could be pursued as an approach for the slow release of cargo. The ADTA bridge remained intact, and no signs of cleavage were seen over the allocated reaction time (12 h).

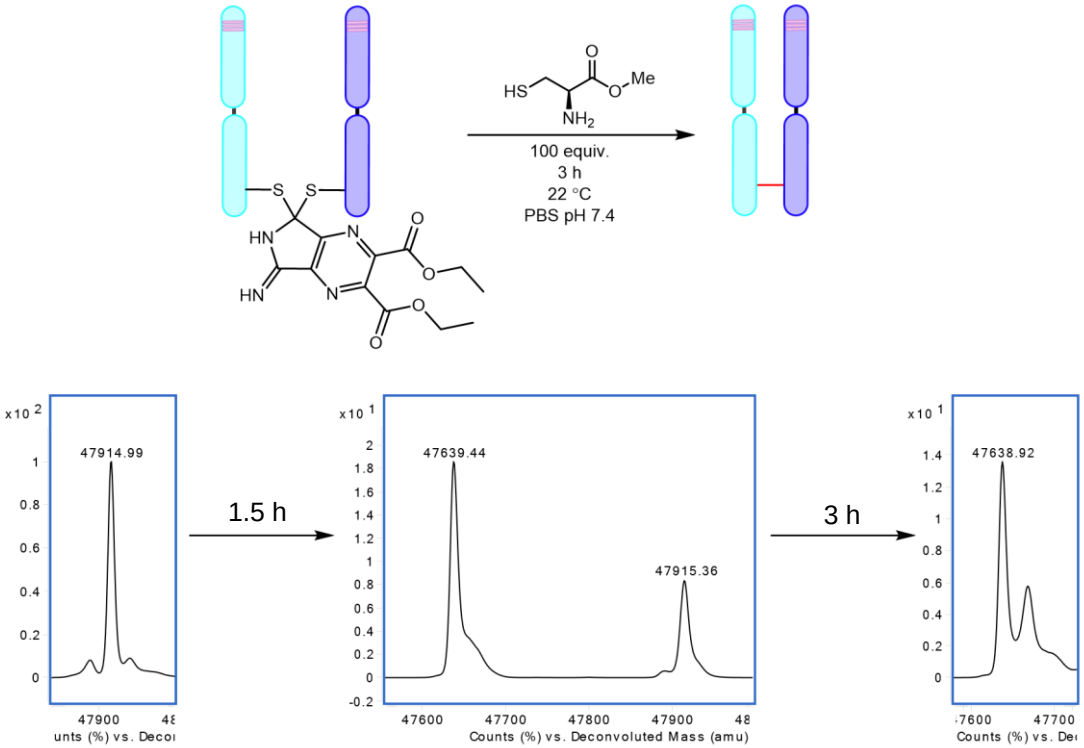


Scheme 74. (A) Bioconjugation conditions employed for diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate on reduced Fab. (B) LCMS of diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate on Fab. 10 equiv., 1 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 274 Da.



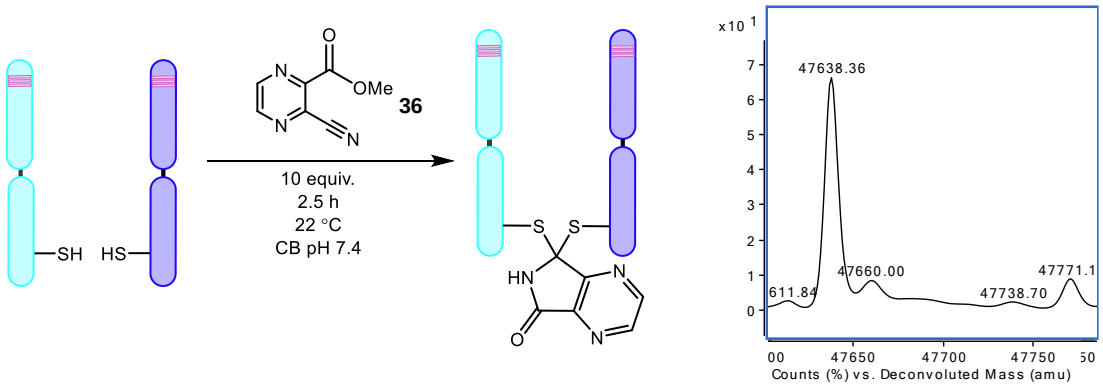
Scheme 75. LCMS trace of incubation of Fab conjugate **44** in aqueous conditions (12 h, 37 °C, pH 7.4).

As seen with **35**, the electron-deficient pyrazine ring underwent cleavage and reformation of the native disulfide bond. The bridge formed with **34** underwent a similar cleavage and proved to yield native Fab after 3 h of incubation with cysteine methyl ester (100 equiv., 22° C, pH 7.4). The LCMS after the allocated reaction time showed an additional peak at 47669 Da, alongside the major native Fab peak (Scheme 76). This is likely due to the presence of methanol which is formed from the large excess of cysteine methyl ester remaining in the reaction mixture. The hydrolysis of the methyl ester over time leads to the release of methanol which could appear as adducts under ionisation[240] (+30 Da) caused by incomplete removal during purification.



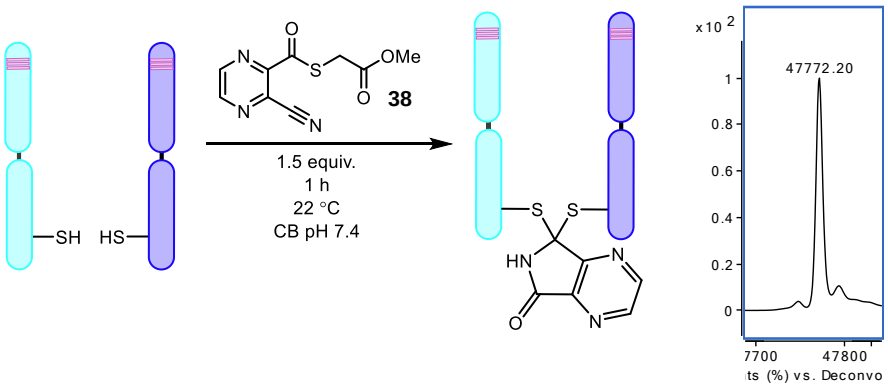
Scheme 76. LCMS trace of incubation of Fab conjugate **44** with cysteine methyl ester (100 equiv. 3 h, 22 °C, pH 7.4). Cleavage of the bridge system is seen over 3 h at 22 °C.

The final set of pyrazine-based bridging reagents explore the use of alternate trapping partners, such as the ester and thioester moieties present in **36** and **38**. **36** provided minor conversion to the bridged conjugate (12%) over 2.5 h, though this was as expected due to the poor nature of the leaving group (-OMe) (Scheme 77). Despite the limited conversion, the ester provided a proof of concept regarding the use alternate trapping partners.

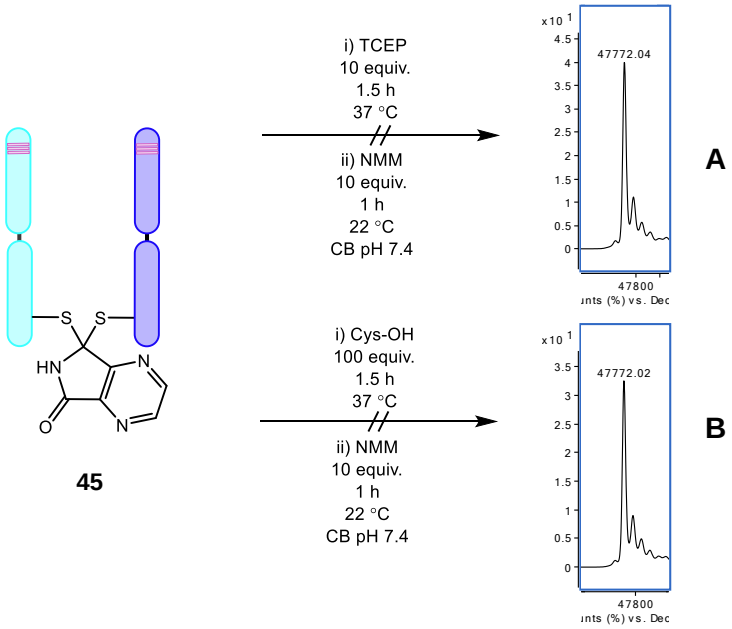


Scheme 77. (A) Bioconjugation conditions employed for methyl 3-cyanopyrazine-2-carboxylate **36** on reduced Fab. (B) LCMS of methyl 3-cyanopyrazine-2-carboxylate on Fab. 10 equiv. 2.5 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 133 Da.

Switching the trapping partner to a thioester, as in **38**, provides a much larger degree of reactivity, which can be >100-fold more reactive than corresponding oxoesters to nucleophiles such as amines. This is partly due to the significantly greater leaving group ability of thiols over alcohols. This can be quantified by the larger relative pK_a value for alcohols, for example, MeOH pK_a = 15.2, whilst MeSH pK_a = 10.3. Oxoesters also present greater thermodynamic stability over thioesters, due to the greater degree of electron delocalisation provided by the oxygen neighbouring the carbonyl group, relative to an analogous interaction in thioesters.[187] This reactivity is demonstrated by the efficient bridging of the disulfide bond using **38** (1.5 equiv., 1 h, 22 °C, pH 7.4) (Scheme 78). Attempts to reduce and cap the cysteine thiols with NMM led to no change in LCMS, this indicates the bridge has indeed formed as opposed to CLT or other reaction modes. To further discern the identity of the bridge, a large equivalent of free cysteine (100 equiv., 1.5 h, 37 °C, pH 7.4) was added followed by NMM upon which no change was identified (Scheme 79). This limits the possibility of thioester formation as addition of cysteine would trigger cleavage.



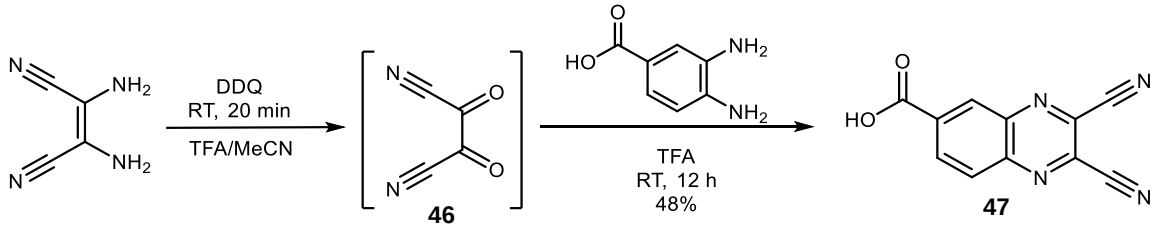
Scheme 78. (A) Bioconjugation conditions employed for methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate on reduced Fab. (B) LCMS of methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate on Fab. 1.5 equiv. 1 h, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 133 Da.



Scheme 79. Incubation of Fab conjugate **45** (**A**) TCEP, 10 equiv., 1.5 h, 37 °C pH 7.4. NMM, 10 equiv., 1h, 22 °C pH 7.4. (**B**) Cysteine, 100 equiv., 1.5 h, 37 °C, pH 7.4. NMM, 10 equiv., 1h, 22 °C pH 7.4. Smaller unlabelled peaks correspond to sodium adducts (M+Na, +23 Da).

3.1.2. Quinoxaline-based Disulfide Bridging reagents

3.1.2.1. Chemical synthesis



Scheme 80. Synthesis of 2,3-dicyanoquinoxaline-6-carboxylic acid (**47**).

The quinoxaline core was synthesised *via* a literature protocol utilising DAMN and 3,4-diaminobenzoic acid. As with the tetra-nitrile pyrazine reagent **35**, DAMN was oxidised to form DISN, after which controlled hydrolysis generates the reactive 'dione' intermediate, **46**. The dione species then undergoes a condensation to yield the quinoxaline bis-nitrile, **47** presenting a carboxylic acid group which can be functionalised *via* an amide coupling reaction (Scheme 70).

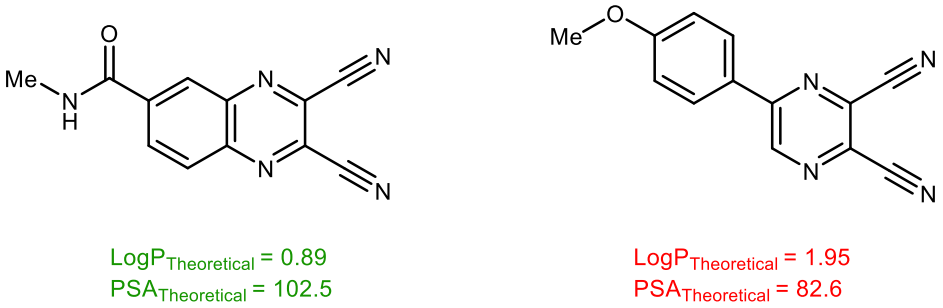
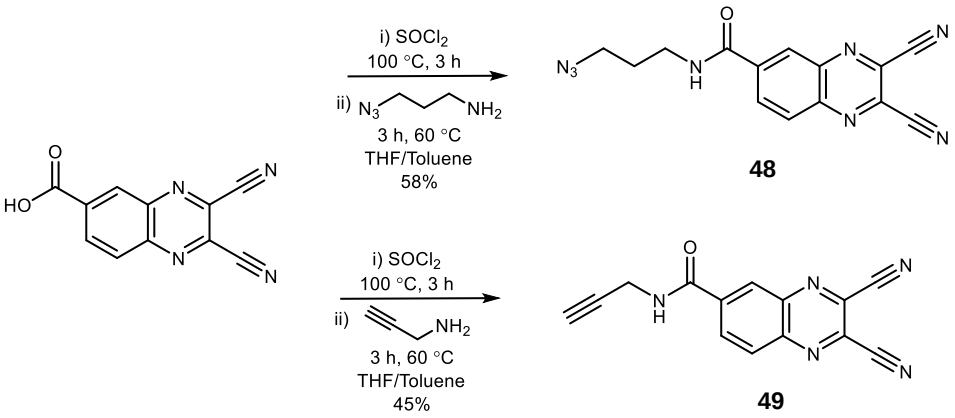


Figure 35. Physicochemical comparison between two designs.

The quinoxaline design eliminates the benzylic hydrogens which limited the functionalisation of the pyrazine bis-nitrile (**29**), and the quinoxaline ring also presents more desirable physical properties, notably the water solubility. The quinoxaline bis-nitrile possesses a 2.2-fold lower theoretical LogP value in comparison to the phenyl-pyrazine bis-nitrile (0.89 Vs 1.95). LogP values depict the partition coefficient of a molecule between aqueous and lipophilic phases, typically octanol and water.[241] The greater the value, the greater the quantity present in the lipophilic phase. Another measure of hydrophilicity is the Polar Surface Area (PSA) defined as the combined surface area, over all polar atoms, primarily oxygen and nitrogen and including their attached hydrogen atoms (Figure 35).[242] The quinoxaline analogue possess a value over 100, 1.24-fold that of the phenyl-pyrazine analogue (82.6). Quinoxaline cores are also prevalent in a wide variety of established drug molecules, providing a second layer of interest to the use of these cores in nitrile-based protein modification.[243]

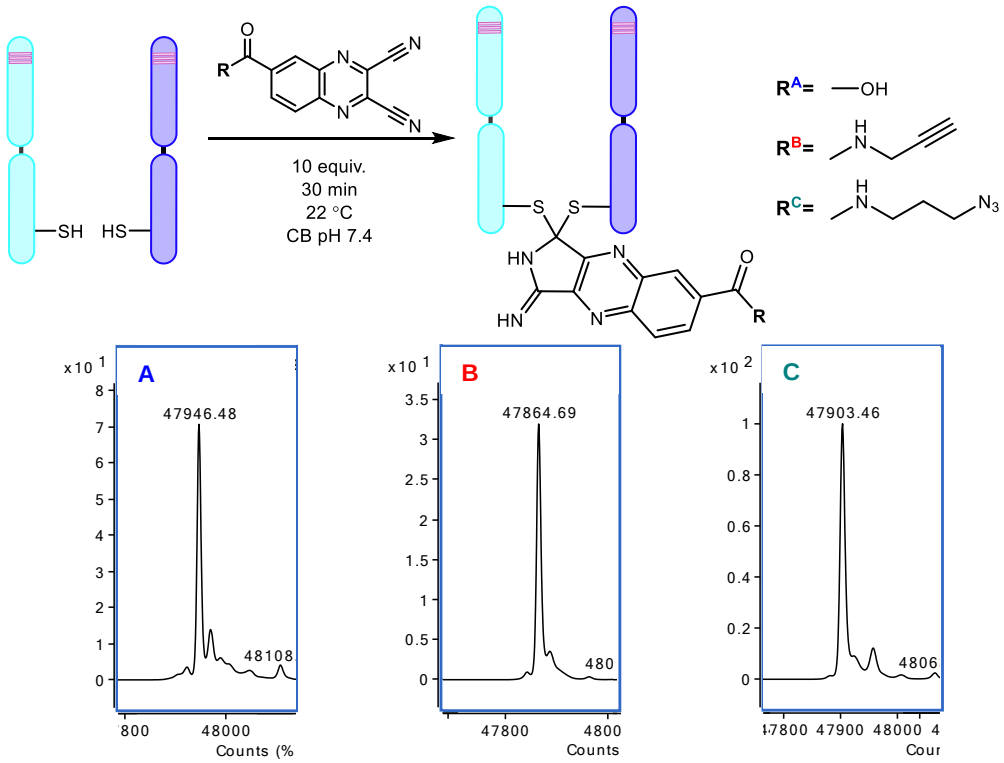


Scheme 81. Synthesis of 2,3-dicyanoquinoxaline derivatives **48** and **49**.

The quinoxaline bis-nitrile **47** was then directly functionalised, *via* an activated acyl chloride intermediate to yield two functional variants: **48** and **49,** capable of SPAAC and CuAAC reactions respectively (Scheme 81).

3.1.2.2. Bioconjugation

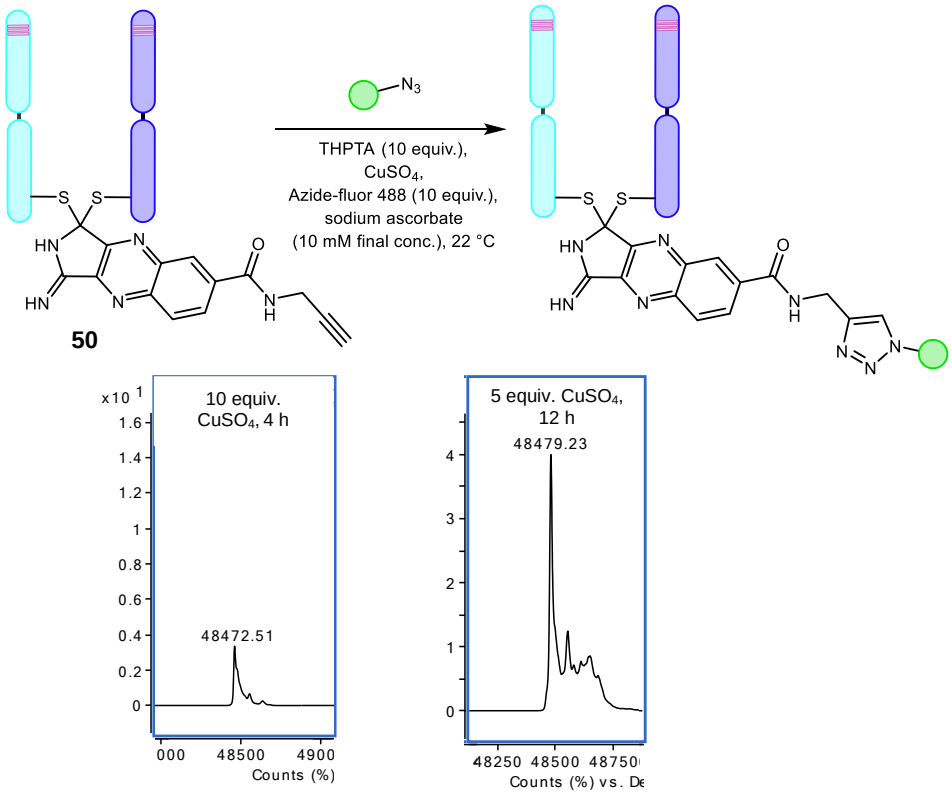
Initial bioconjugation studies carried out on the Fab fragment (10 equiv., 2.5 h, 22 ºC, pH 7.4) with **47** successfully yielded the bridged conjugate in a clean manner. This was also found to be the case with both **48** and **49**. Further investigation into the reaction timescale, with **49**, revealed the true time for the bridging cascade to be just 30 minutes, 5-fold shorter than the timeframe previously employed (Scheme 82).



Scheme 82. Bioconjugation conditions employed for 2,3-dicyanoquinoxaline derivatives on reduced Fab. (**A**) LCMS of 2,3-dicyanoquinoxaline-6-carboxylic acid **47** on Fab. 10 equiv. 15 min, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 224 Da.

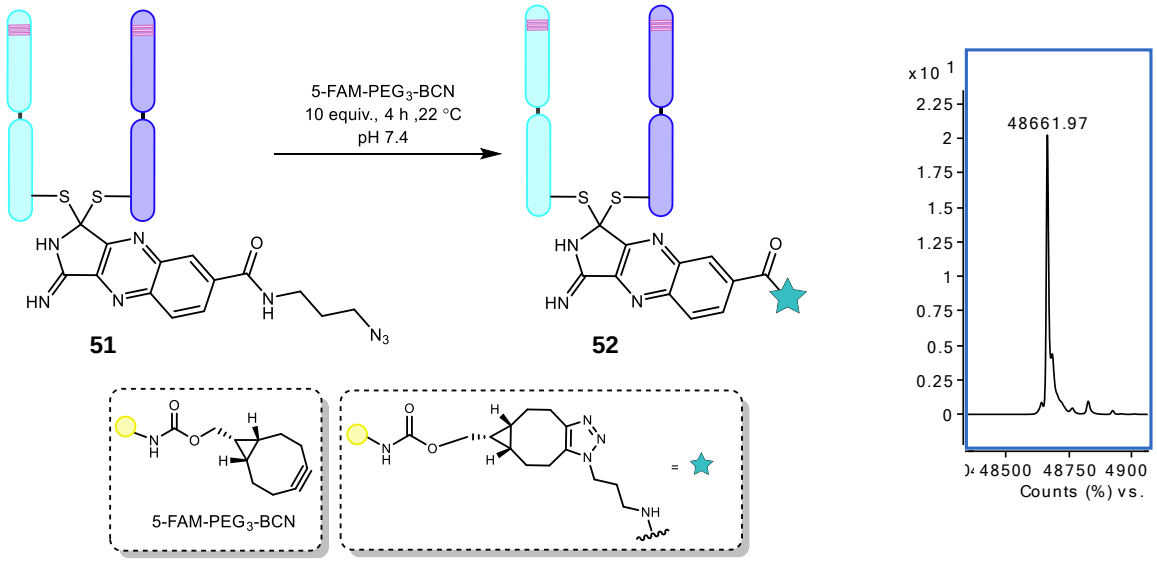
(**B**) LCMS of 2,3-dicyano-*N*-(prop-2-yn-1-yl)quinoxaline-6-carboxamide **49** on Fab. 10 equiv. 15 min, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 261 Da. (**C**) LCMS of *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide **48** on Fab. 10 equiv. 15 min, 22 ºC, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 306 Da.

As mentioned, **49** contains a pendant alkyne and could therefore be explored for CuAAC functionalisation with Azide-fluor 488. The pyrazine equivalent showed cleavage when exposed to the Cu(I) conditions however it was yet to be seen if these conditions would impact the quinoxaline derivative in a similar manner. The quinoxaline derivative excitingly yielded a greater degree of clicked conjugate; however, a large quantity of native LC was also seen, indicating cleavage. With this positive result, the copper was reduced from the prior 10 equivalents down to 5 equivalents and left for a longer duration of time (12 h). The optimised conditions yielded the clicked conjugate with quantitative conversion with no visible cleavage of the bridge. This shows there is a threshold regarding Cu(I) tolerability, depending on the reagent design utilised. Noise can be seen in the in the LCMS trace however this is due to the presence of copper ions interfering with the ionisation of the sample.



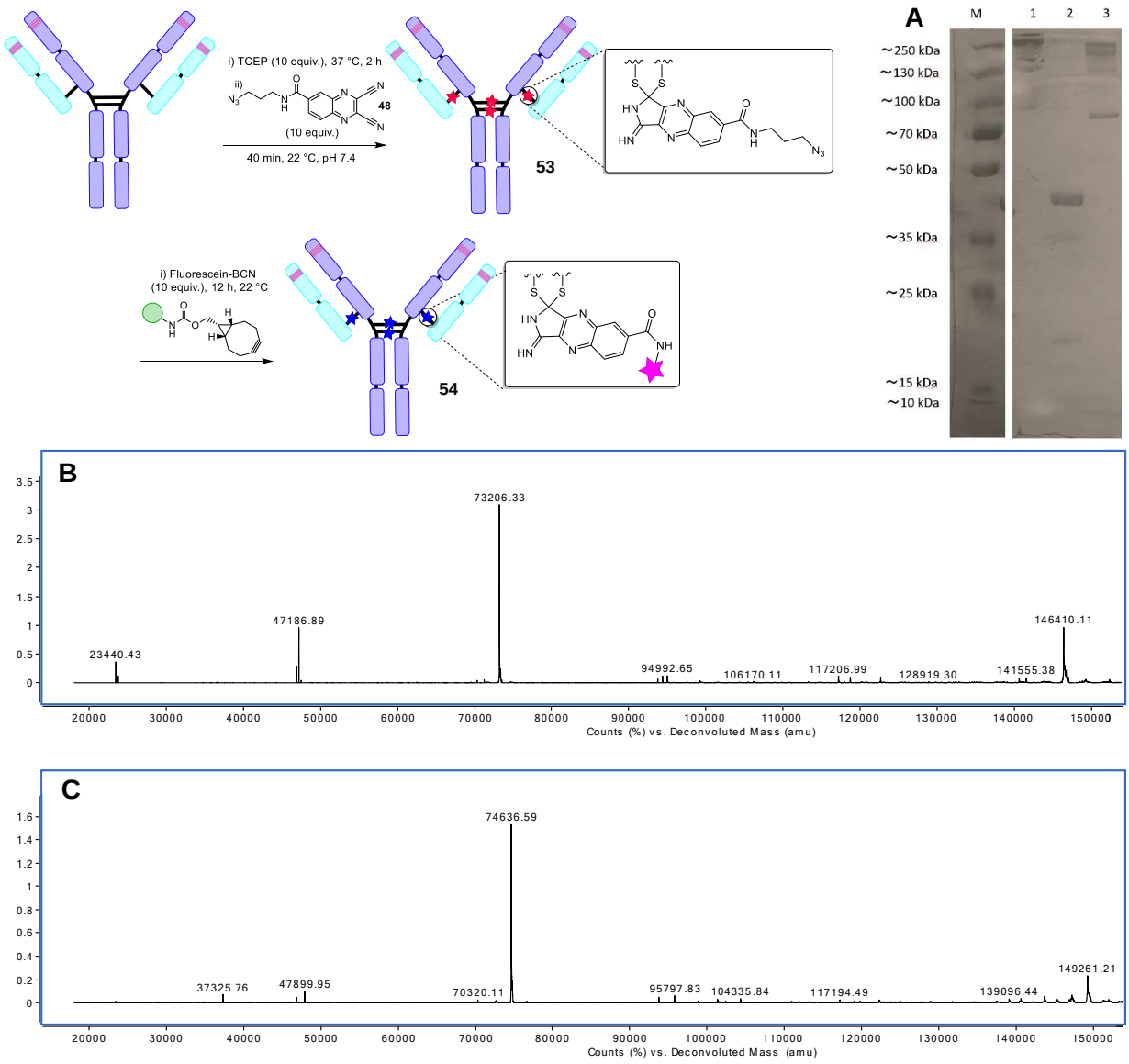
Scheme 83. (A) CuAAC conditions employed for Fab conjugate **50**. (B) LCMS traces of Azide-fluor 488 CuAAC on Fab conjugate. Mass of native Fab = 47638 Da. Expected mass of addition = 261 Da + 574 Da.

The azide derivative **48** was synthesised with the aim to undergo SPAAC functionalisation. Successful bridging of Fab was then followed by SPAAC attachment of fluorescein, conveniently affording the final functionalised conjugate **52** (Scheme 84).



Scheme 84. LCMS of SPAAC reaction on Fab conjugate **51**. 5-FAM-PEG$_3$-BCN, 10 equiv., 22 °C, 4 h, pH 7.4. Mass of native Fab = 47638 Da. Expected mass of addition = 306 Da + 727 Da.

Due to the efficient bridging and subsequent functionalisation of the Fab fragment with **48**, the next step was to apply this to a more complex system in the form of the full antibody, trastuzumab. 10 equiv. for 40 min at 22 °C, pH 7.4 yielded the final bridged full antibody. Within the LCMS trace native LC can be seen however, this was deemed to be due to the excellent ionisation of the LC. A second peak at 47187 Da accounts for double LC with the addition of one molecule of **48**. This is an artifact of the mass spectrometer and upon analysis of the final conjugate by SDS-PAGE no signs of LC or HC could be seen, further indicating the completion of bridging. A dual addition can be seen in the HL region, 73206 Da (native HL 72588), which is referred to as half-antibody which forms as a result of disulfide scrambling.[244,148] The full bridged antibody can be seen with a mass of 146410 Da showing 4 additions of **48**. The bridge was then functionalised as before with BCN-fluorescein to yield the final functional conjugate **54**. The LCMS trace shows limited to no signs of native LC or the MS artifact around 47187 Da which further reduces the likelihood of these peaks in the LCMS trace for the initial bridging (Scheme 85).



Scheme 85. Bioconjugation conditions employed for *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide on reduced trastuzumab. (**A**) SDS-PAGE of bridged species; ; M – molecular marker, 1 – Full antibody, 2 – Reduced full antibody, 3 – Full antibody bridged with **48**. (**B**) LCMS of N-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide on trastuzumab. 10 equiv. 40 min, 22 ºC, pH 7.4. Mass of native trastuzumab = 145177 Da. Expected mass of addition = 4 X 306 Da. (**C**) LCMS of SPAAC reaction on conjugate **53**. 5-FAM-PEG$_3$-BCN, 10 equiv., 22 °C, 12 h, pH 7.4. Mass of native trastuzumab = 145177 Da. Expected mass of addition = (4 X 306 Da) + (4 X 727 Da).

3.2. ELISA

An enzyme-linked immunosorbent assay (ELISA) was carried out on native Fab and the Fab conjugate **27**, containing the bridged product of pyrazine-2,3-dicarbonitrile, to test affinity towards the HER2 antigen. A 96-well polystyrene plate was coated first with HER2 antigen followed by the application of native Fab and Fab conjugate **27**. Upon completion of the incubation time with a secondary antibody linked to HRP, *o*-phenylenediamine.HCl was added to quench the reaction. This was converted to 2,3-diaminophenazine by the enzyme HRP, which shows strong absorbance at 490 nm, allowing for spectrophotometric quantification of the antigen binding activity of the Fab conjugate **27** and native Fab. The Fab conjugate **27** demonstrated little deviation in binding activity to that of native Fab.

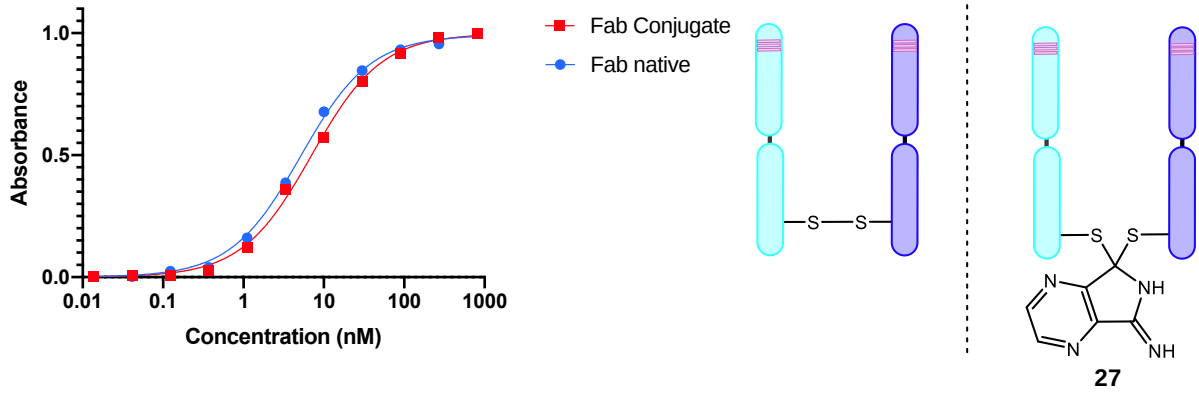


Figure 36. Binding activity to HER2 of trastuzumab native Fab and Fab conjugate **27** over concentration ranges from 810-0.0137 nM, the absorbance was measured at 450 nm.

3.3. Bis nitrile bridge stability

As with the CLT conjugates the hydrolytic stability is a key piece of information regarding the viability of such reagents in protein modification. To investigate this, the bridged conjugate **55** was subjected to incubation at 22 ºC, pH 7.4 for 5 days. under these conditions, no cleavage or degradation of the conjugate was seen and therefore it was concluded that the bridge is stable under aqueous conditions.

The conjugate was also subjected to 2-mercaptoethanol (BME); 100 equiv., 2 h, 37 °C, and showed no signs of cleavage indicating great stability towards external thiols. As with the CLT conjugates, the bridged conjugate **55** was also subjected to mimic conditions for relevant physiological environments, namely early endosomal and serum conditions. The conjugate was found to be stable over the course of 24 h at 37 °C under the corresponding conditions (Figure 37).

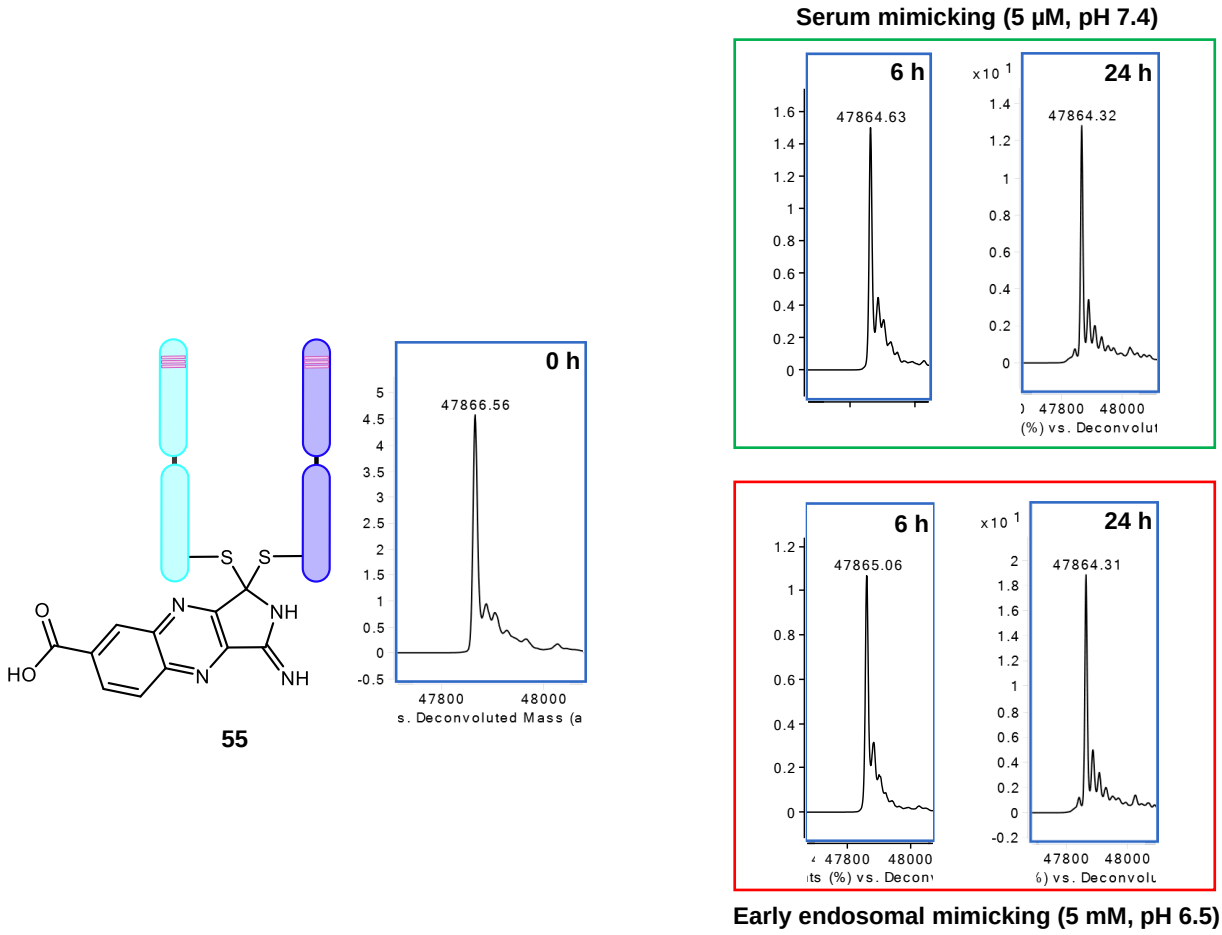


Figure 37. Stability investigation of the 2,3-dicyanoquinoxaline-6-carboxylic acid bridged conjugate **55**. Serum (5 μM GSH, pH 7.4, 37 ºC) - highlighted with a green box and endosomal (5 mM GSH, pH 6.5, 37 ºC) – highlighted with a red box.

Next the stability of the bridged conjugate in human serum was considered. This study required the use of the fluorescent Fab conjugate in order to track the potential transfer to HSA, therefore, the fluorescein functionalised conjugate **52** was employed. The conjugate **52** (7 µM) was incubated in human serum for 5 days and analysed by SDS-PAGE. The two concentrations of serum tested were that of 50% and 100% in order to acquire a better understanding with regards to the serum stability. The conjugate was found to be stable as no transfer of fluorescence was detected in the HSA region (Figure 38).

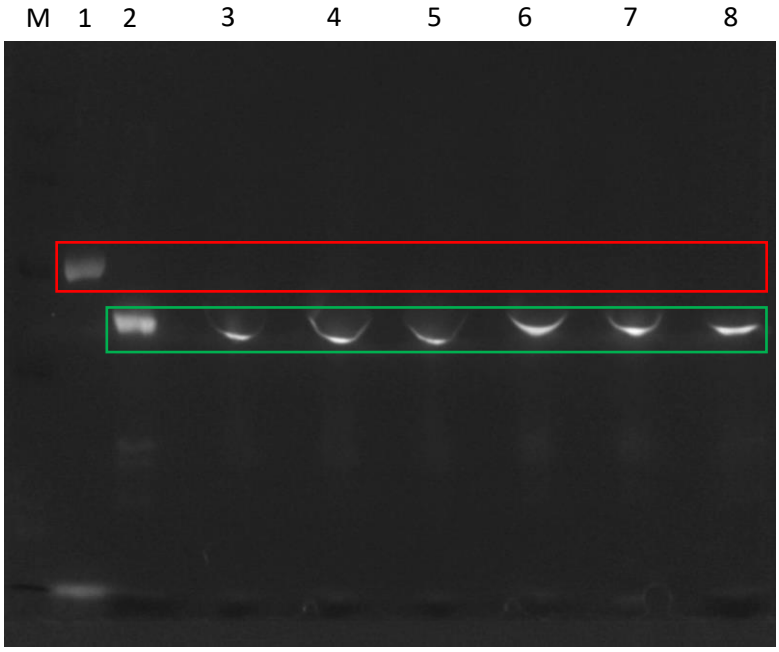


Figure 38. SDS-PAGE of serum stability study; M – molecular marker, 1 – fluorescent HSA reference, 2 – fluorescent bridged Fab reference (conjugate **52**), 3 – Day one (100%), 4 – Day three (100%), 5 – Day five (100%), 6 – Day one (50%), 7 – Day three (50%), 8 – Day five (50%). The red box highlights the HSA band, the green box highlights the Fab band.

3.4. Selectivity reactions

Due to the hypothesised dynamic nature of the bis-nitrile bridging cascade, a degree of selectivity between mono-thiol and di-thiol systems was hypothesised. The bridging cascade presents a level of reversibility till the final 'trapping' cyclisation, the proximity induced acceleration of the cyclisation step for the di-thiol systems was thought to be the foundation for the selectivity (Figure 39).

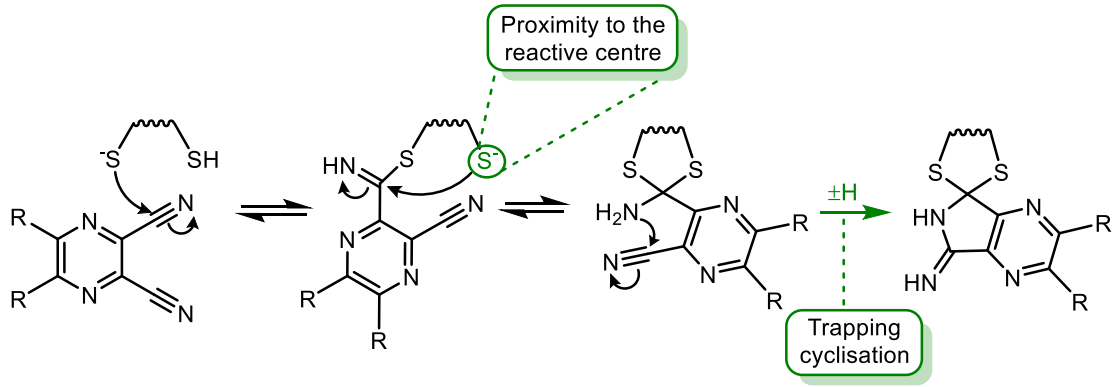
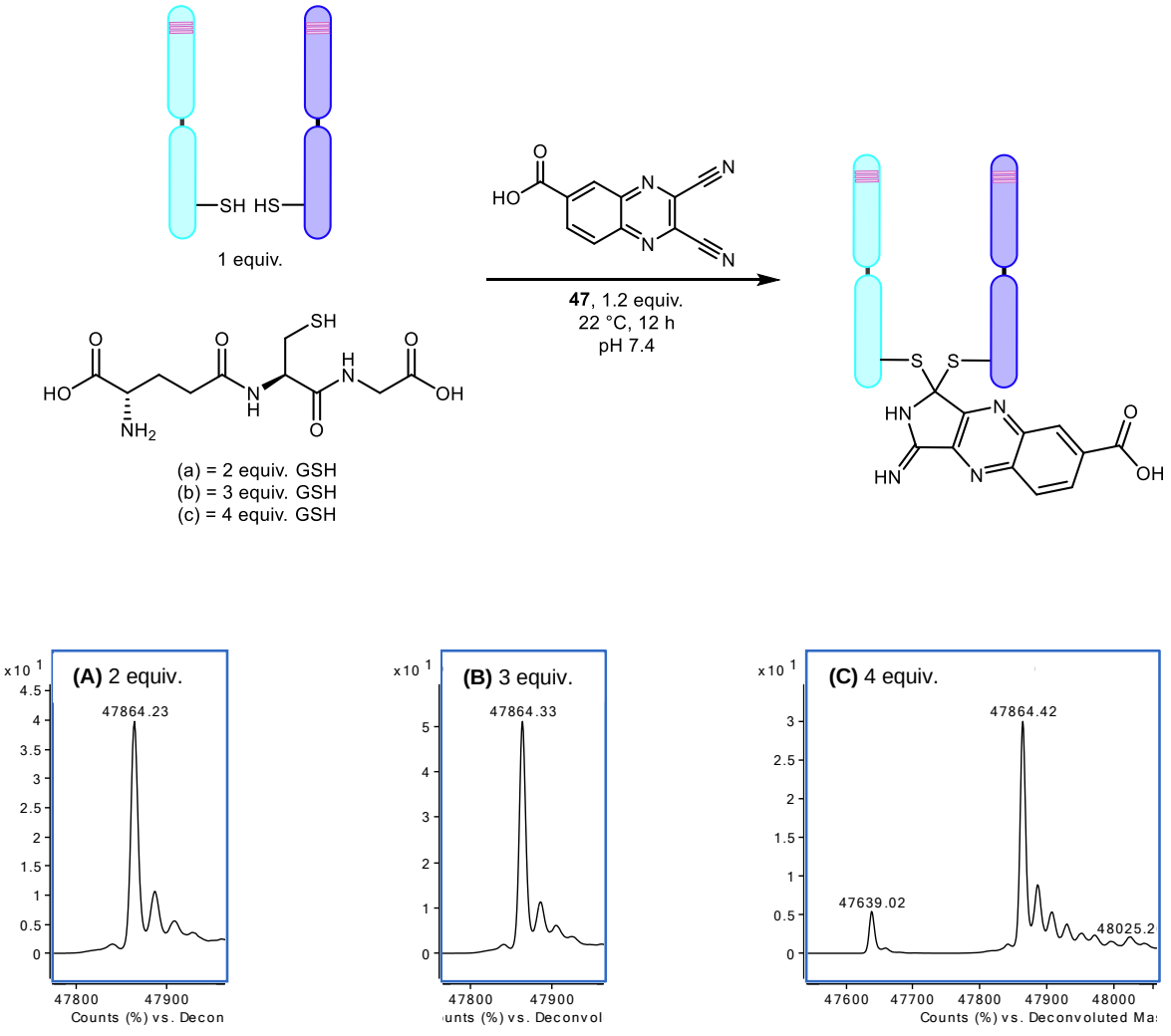


Figure 39. Schematic representation of proximity induced ADTA trapping.

In order to test this selectivity, a competition experiment was setup between Fab with a disulfide bond, as the di-thiol system, and GSH as the bio-compatible mono-thiol system. The quinoxaline reagent **47** was selected as the bis-nitrile bridging reagent. The study revealed an inherent selectivity for di-thiol systems, yielding the bridged product in quantitative yield (47864 Da). Up to 3 competing equivalents provided the full bridge, and at 4 equivalents signs of inhibition is seen, represented by the presence of native Fab. NMM was employed as the standard for comparison, due to the well-known sulfide capping ability. When compared to NMM, the selectivity becomes more apparent as the maleimide is unable to differentiate between the systems to the same degree and leads to minimal conversion to the capped Fab conjugate (supplementary information **S78**).



Scheme 86. ADTA trapping competition experiment. Mass of native Fab = 47639 Da. Expected mass of addition = 224 Da. (**A**) 2 equiv. GSH, (**B**) 3 equiv. GSH, (**C**) 4 equiv. GSH

3.5. Thioredoxin

Thioredoxins (Trx) are a small class of redox proteins with a size of approximately 12 kDa. The Trx system is responsible for a large part of cellular redox balance and plays an influential role in neurodegenerative diseases. The system consists of Trx, thioredoxin reductase (TrxR), NADPH coenzyme, and thioredoxin-interacting protein (Txip).[245] The active site contains a surface accessible disulfide bond in the form of a CXXC motif, and in the variant Trx1, this is specified as CGPC. The mechanism of action is carried out by the reduced form and starts with a nucleophilic attack of the *N*-terminal thiol on the target disulfide. This nucleophilic attack of the thiolate is the result of the low pK_a value of the *N*-terminal cysteine of the CXXC motif (pK_a ~7). By contrast, the sulfur of the *C*-terminal cysteine has a high pK_a (~9) and is therefore present as a thiol, which can assist in stabilising the *N*-terminal thiolate.[246] Oxidised Trx is then reduced back to its active form by TrxR (Figure 40).

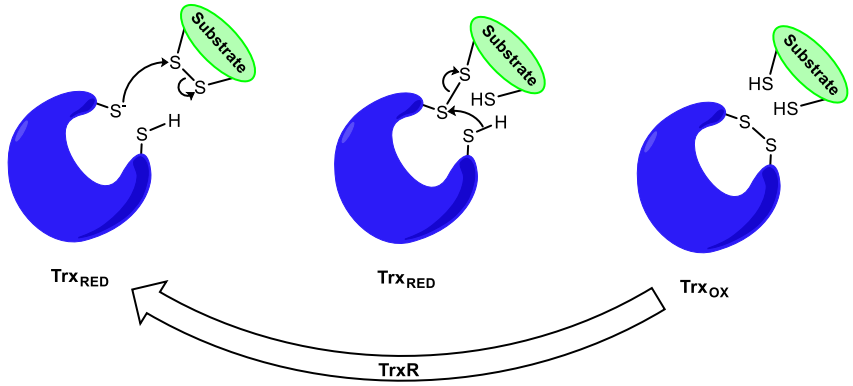


Figure 40. Thioredoxin mechanism of action.

A crystal structure of recombinant human Trx relevels a distance of ~4.3 Å between the two cysteine residues involved in the formation of the disulfide bond (Cys32-Cys35) in solution (Figure 41). This provides a great platform to further explore and expand upon the use of bis-nitrile reagents in other protein systems. The enzymatic nature also enables the investigation of these reagents as potential scaffolds for inhibition of disulfide-bearing enzymes.

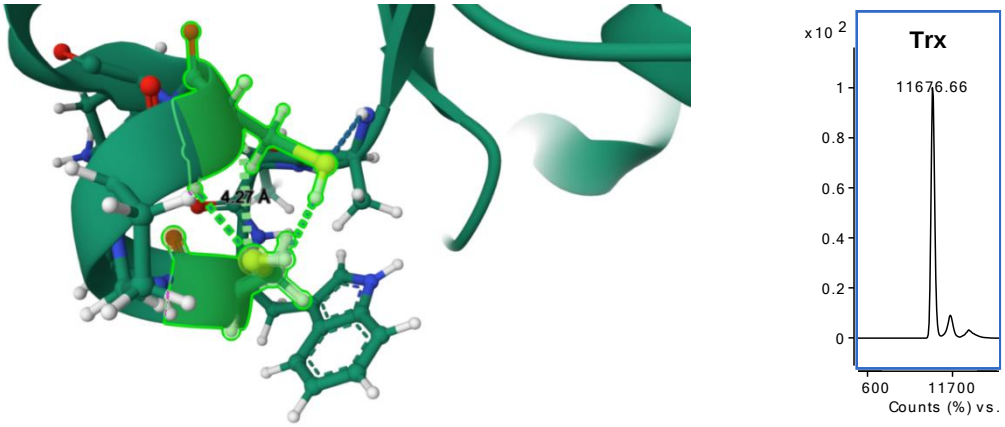
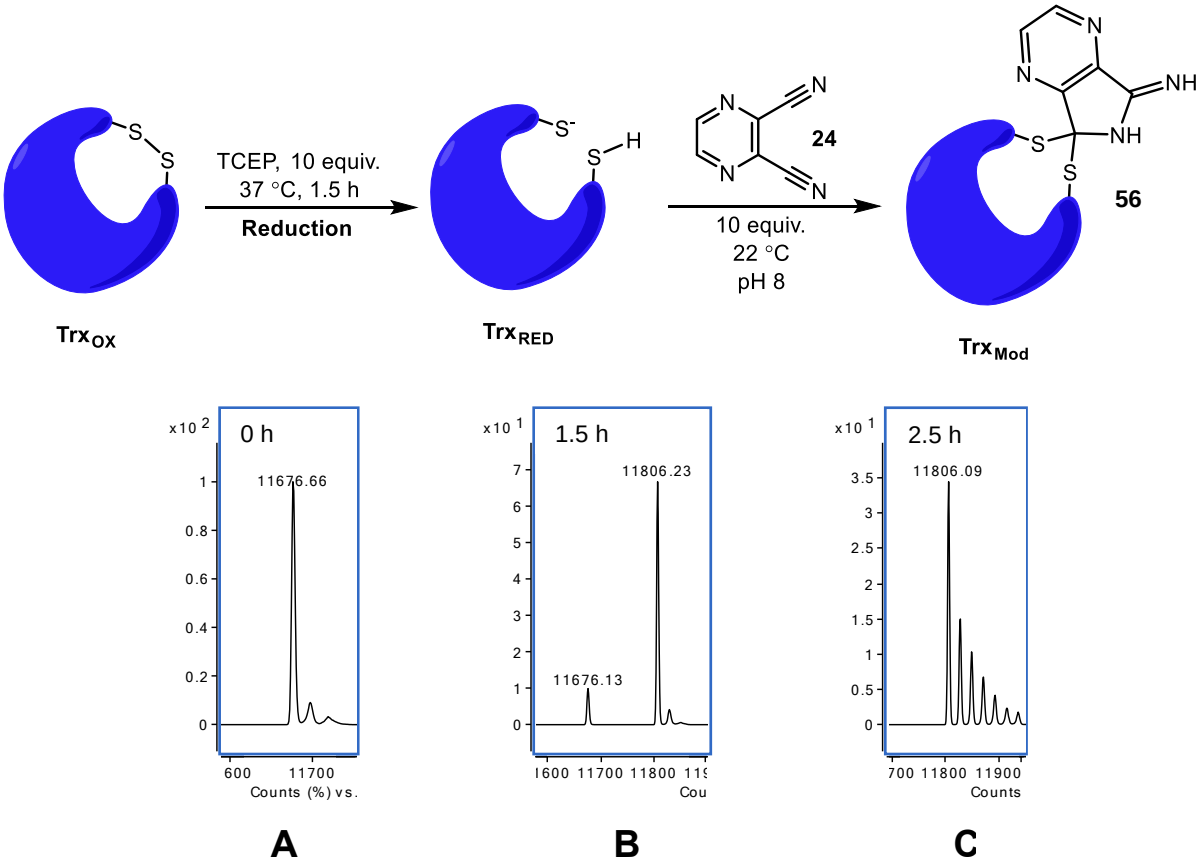


Figure 41. Thioredoxin (Trx), highlighting the distance between the disulfide bond. LCMS trace for native thioredoxin, expected mass = 11676 Da.



Scheme 87. Bioconjugation conditions employed for pyrazine-2,3-dicarbonitrile **24** on reduced Trx. Pyrazine-2,3-dicarbonitrile, 10 equiv., 22 ºC, pH 7.4. Mass of native Trx = 11676 Da. Expected mass of addition = 130 Da. Smaller unlabelled peaks correspond to sodium adducts (M+Na, +23 Da).

Trx_{OX} was reduced with TCEP over the course of 1.5 h at 37 °C, this reduced Trx (Trx_{RED}) was then treated with 10 equivalents of pyrazine-2,3-dicarbonitrile. After 1.5 h, a conversion of 87% can be seen which upon extension to 2.5 h results in the complete conversion to the bridged conjugate (Scheme 87).

Trx_{RED} is able to reduce the disulfide bond between the α and β chains that make up insulin, and this effect results in the precipitation of the free chains and consequent increase in turbidity. This increase can monitored by absorption at 650 nm.[247] A preliminary insulin assay done with sample **A** and **B**, from which clear signs of inhibition could be detected. Minor activity can be seen in the modified sample (Figure 42). However, this can easily be justified with the presence of slight portion of unmodified Trx (13%) and the reducing agent DTT, which can also reduce the insulin chains.

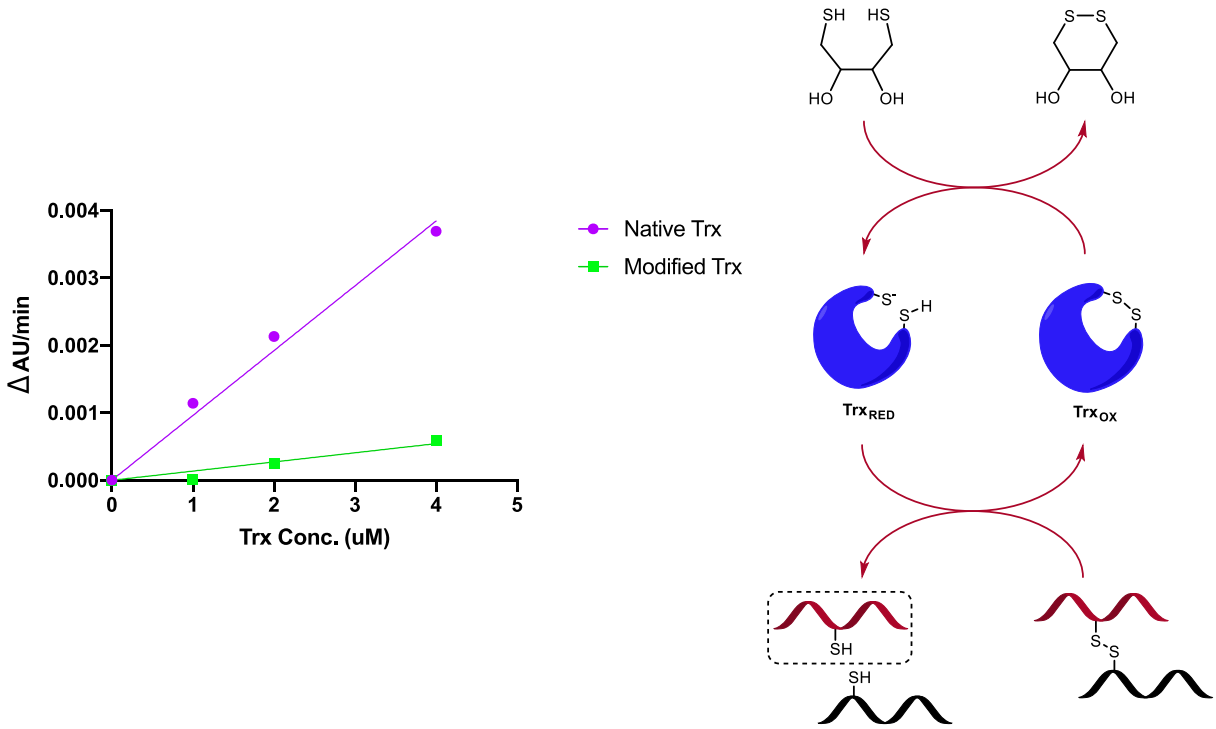
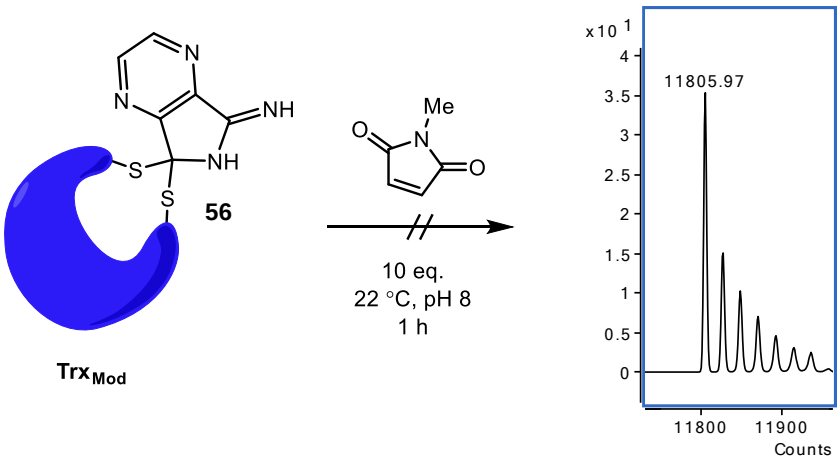


Figure 42. Bioactivity measurement of Trx and modified Trx **56**. The pink line represents native thioredoxin, the green line represents ADTA modified thioredoxin.

Addition of NMM to the Trx conjugate **56**, resulted in no change in the LCMS trace further indicating the modification of the disulfide bond (Scheme 88). Further experiments revealed that 2,3-dicyanoquinoxaline-6-carboxylic acid, provided analogous results to the pyrazine counterpart, yielding complete conversion to the desired product in 2.5 h.



Scheme 88. Incubation of Trx conjugate **56** with NMM (10 equiv., 1 h, 22 °C). No change is seen indicating modification of the disulfide bond.

3.6. Summary and conclusion of nitrile bridging

1,2 aryl bis-nitrile systems have been shown to be efficient disulfide bridging reagents on multiple systems such as Fab, trastuzumab and thioredoxin. Pyrazine-2,3-dicarbonitrile **24** provided the first example of the bis-nitrile trapping system, **24** was shown to react with both ethanedithiol and ethanethiol to produce the corresponding trapped ADTA. **24** was then shown to successfully bridge the Fab disulfide, demonstrating excellent selectivity and stability in physiological and mimic conditions. A functional derivative of the pyrazine core was synthesised in the form of **30**, embellished with two CuAAC compatible alkyne handles. This reagent uncovered a sensitivity towards copper upon attempts to functionalise the alkyne handles, which could potentially provide be developed into a triggered release mechanism.

The pyrazine system was further developed with electron deficient designs, these reagents provided a new layer of reactivity to the nitrile bridge. Diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate **34** and pyrazine-2,3,5,6-tetracarbonitrile **35** bridged Fab fragments both demonstrated cleavage upon exposure to cysteine. This highlights the impact of the electronic effects provided by the electron withdrawing groups. Alternate bridge pairs were also investigated, **36** and **38** possessed an ester and thioester moiety respectively. These reagents also provided clean bridged species with varying rates of reactivity, demonstrating a possible area of tunability.

Quinoxaline derivatives of the bis-nitrile system generated rapid, stable bridged Fab fragments. Click compatible derivatives synthesised afforded access to functional handles on Fab fragments. Carboxyfluorescein functionalised bridged Fab fragments were found to be stable when incubated under physiological conditions over five days, no transfer of fluorescence was observed. The quinoxaline derivative was also found to efficiently bridge the four disulfides present on the full antibody; trastuzumab, confirmed with LCMS and SDS-PAGE analysis.

The dynamic bridging cascade of the bis-nitrile systems was also shown to possess inherent selectivity when applied in a competition experiment between mono (GSH) and di-thiol (Fab) systems. The quinoxaline derivative **47** was able to tolerate up to 4 eq. of GSH, which proved to be enough to prevent the sulfide capping reagent, *N*-methyl maleimide, from modifying the free thiols. Finally, the bis-nitrile systems have been shown to bridge another protein class; namely the redox protein, thioredoxin (Trx). Pyrazine-2,3-dicarbonitrile **24** was shown to inhibit the activity of Trx when compared with native Trx in an insulin reduction assay. This result provides preliminary evidence as potential inhibitor moieties.

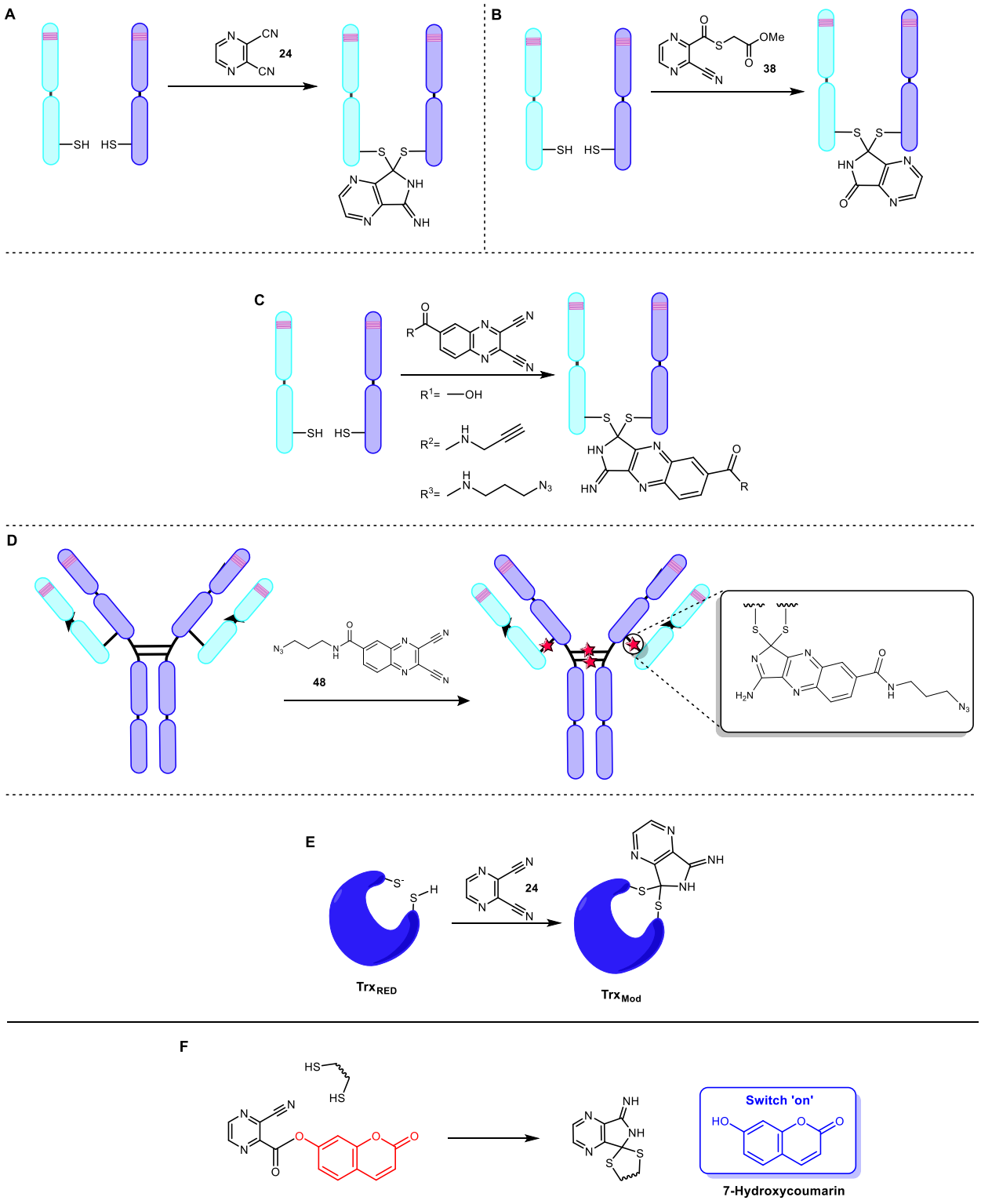


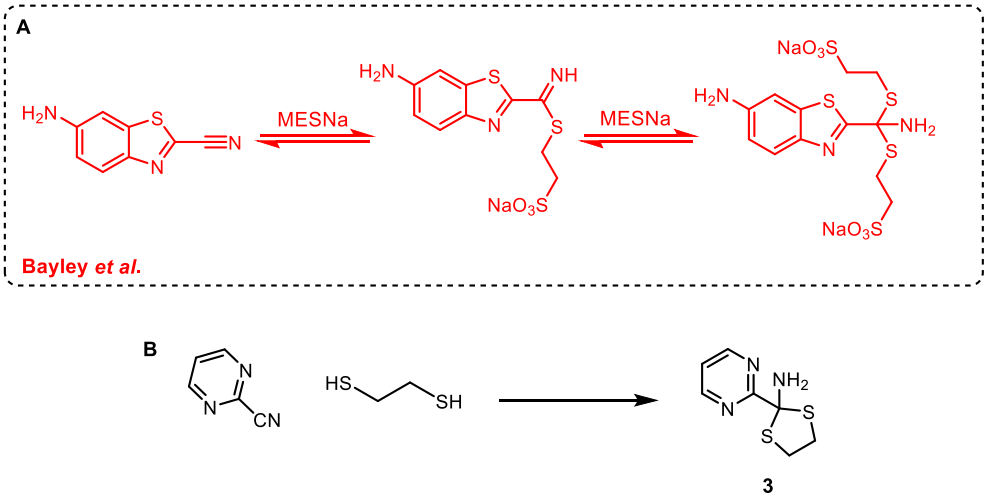
Figure 43. (**A**) Fab bis-nitrile bridging overview, (**B**) Fab thioester-nitrile bridging overview, (**C**) Fab quinoxaline bis-nitrile bridging overview, (**D**) Trastuzumab quinoxaline bis-nitrile bridging overview, (**E**) Trx quinoxaline bis-nitrile inhibition overview, (**F**) Future 7-hydroxycoumarin design.

Further work includes the investigation behind the pyrazine-based bridging reagent copper mediated cleavage mechanism, to generate potential click and release reagents. A detailed study into the tolerance towards copper would provide greater insight into the conditions required for successful click and release. Functional variants of the pyrazine nitrile thioester and ester reagents could also be explored in an attempt to generate useful tools reagents for labelling experiments. These pyrazine nitrile thioester and ester reagents may find use as disulfide detection reagents, use of leaving groups such as 7-hydroxycoumarin may provide a visual indicator. 7-Hydroxycoumarin is a is fluorescently active plant metabolite and food component, however, this fluorescence is quenched upon loss of the free hydroxyl group.[248,249] Release of 7-hydroxycoumarin acts as a switch, which turns “on” in the presence of reduced disulfide bonds or high concentrations of free thiol (Figure 43 **F**). These bis-nitrile reagents could be expanded further as a new class of PDI and redox protein inhibitors (e.g. Trx) for use in assays. The core scaffold could also be developed to further increase affinity for target active sites.

4. Chapter 4: Amino dithioacetals (ADTAs)

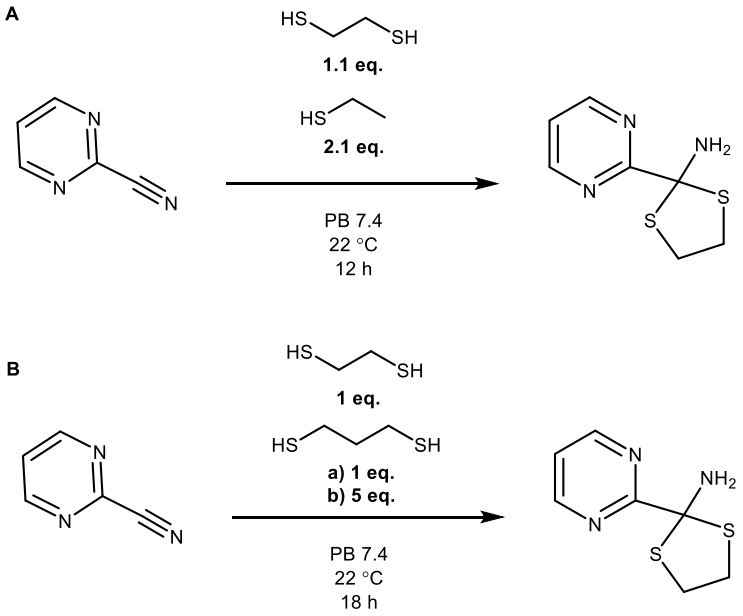
4.1. ADTA formation

As seen in chapters 2 and 3, dual addition of thiol moieties to nitrile reagents are reactions with exciting potential. Bayley *et al.* described the formation of a tetrahedral product which was unable to be isolated due to its highly reversible properties. Through the use of the commercially available ethanedithiol (EDT), a stable analogue of this tetrahedral product was formed. This was termed the nitrile bis-thiol (NBT) reaction, and the resulting product was named an amino dithioacetal (ADTA) (Scheme 89). This new reaction class presents a potential future use in protein modification as a bioorthogonal reaction. Chapter 4 looks to explore the selectivity of these reactions as well as develop an understanding behind the stability and reversibility. The chapter will ideally provide insight into the viability of the reaction with regards to both the stability and selectivity relative to bioconjugation conditions for future development.



Scheme 89. (**A**) Previous work by Bayley *et al*. presenting the dual thiol addition to nitrile groups. (**B**) General scheme for the NBT reaction.

The NBT reaction leads to the formation of ADTAs, of these species **3** presents a raised level of interest. The formation of a tetrahedral species has been shown by Bayley *et al*. to be highly reversible, however the use of the EDT provides an added degree of stability, allowing for the isolation of product in excellent yield (Scheme 89). In order to expand upon the specialty behind this reaction, the EDT-NBT reaction was subjected to a series of competition experiments (Scheme 90).



Scheme 90. (**A**) Reaction scheme for the competition NBT formation of ADTA (**3**) with ethanedithiol and ethanethiol. (**B**) Reaction scheme for the competition NBT formation of ADTA (**3**) with ethanedithiol and propanedithiol.

The first of these experiments was the competition between EDT and the monothiol counterpart, ethanethiol (ET). The outcome for this reaction was the sole isolation of the EDT – ADTA, **3**. This shows there is an innate selectivity present for EDT over mono-thiols, attributed to the reversible nature of mono-thiol addition to nitriles illustrated by Bayley *et al*. The next set of experiments looked at the competition between EDT and longer chained di-thiols, to investigate the selectivity of the 5-member ring formation. 1,3-Propanedithiol (PDT) was selected as the competing dithiol, as this would lead to the formation of a 6-membered ring, which would be the hypothesised thermodynamic product due to lower ring strain. The outcome impressively followed that of the prior ET competition reaction, leading to the sole isolation of the EDT – ADTA, **3**. This was found to remain the case upon addition of 5 competing equivalents of PDT. The 5-membered ring formation therefore shows remarkable preference over both monothiol and alternate dithiol systems.

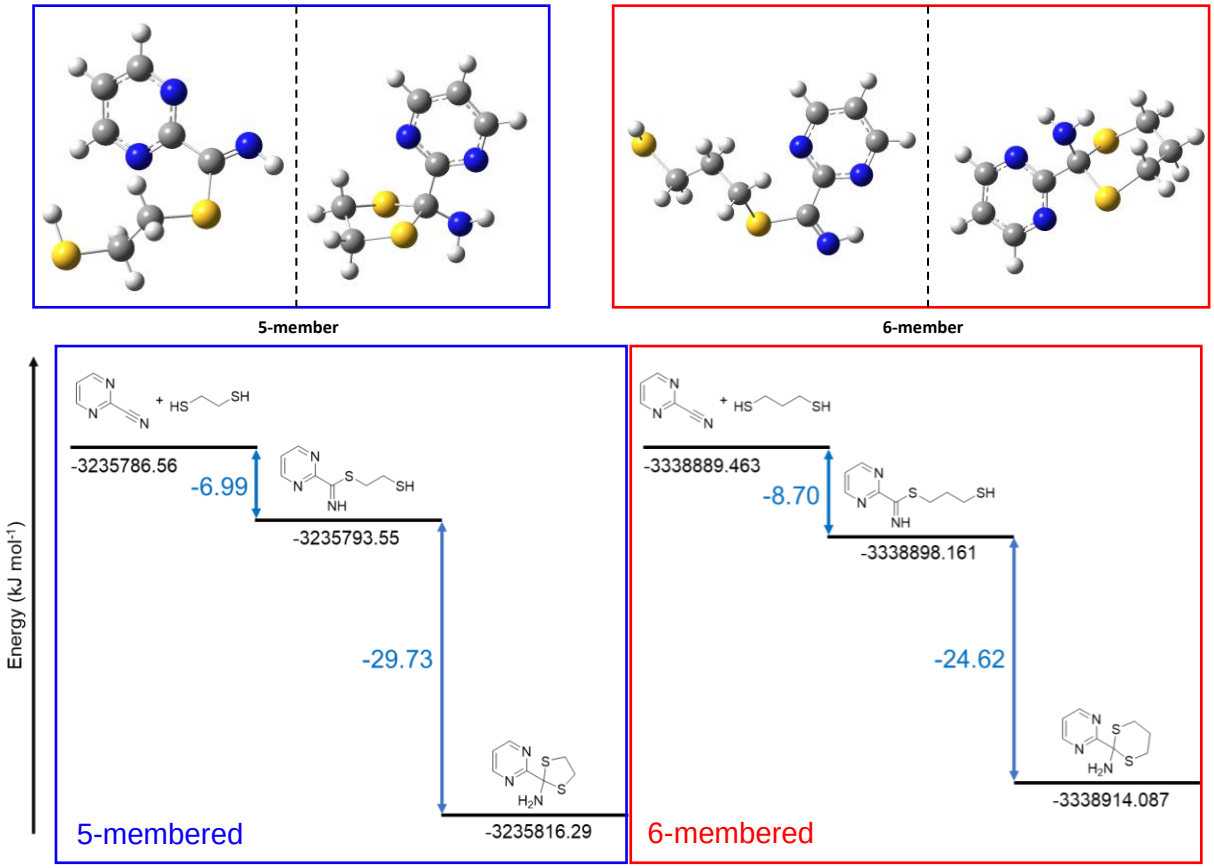


Figure 44. Energy diagram illustrating the energy difference between the 5 and 6-membered ring formation. The blue box highlights the 5-membered ring formation, and the red box highlights the 6-membered ring formation.

To further investigate the preference behind the 5-membered ring formation, DFT calculations were carried out with the help of Prof. Mike Porter. The software CREST (Conformer–Rotamer Ensemble Sampling Tool) was employed in order to identify the lowest energy conformer for the open and closed system for both ring sizes. These were then applied to the DFT study from which the outcome indicates a similar change in energy from the first thiol addition (7 - 8 kJ mol $^{-1}$) in both 5 and 6 membered systems. The second step indicates a slight favour (~ 5 kJ mol $^{-1}$) in the formation of the 5 membered ring over the 6 membered ring (Figure 44). The difference is not as large as initially thought due to the excellent selectivity however upon reconsideration the formation of the 6 membered ring is also likely. This however may not be preferred over the 5 membered ring as seen both computationally, and in particular, experimentally. Further investigation into the 6 membered ring may provide greater insight and possible uses for these systems. Rate of cylisation varies with ring size, 5-membered rings are kinetically favoured over the larger 6-membered rings. The larger 6 membered ring is typically thermodynamically favoured due to the lesser ring strain. The preference for the 5 membered ADTA may stem from the large, diffuse nature of sulfur atoms, which may help mitigate transannular strain.

4.2. ADTA stability

To explore the prospect of ADTAs as bioconjugation tools, the stability to multiple conditions was explored. The first of these conditions was the incubation of ADTA **3** in buffered conditions in order to establish the aqueous stability. No changes were seen in the NMR spectra over 12 h, implying the construct is stable in aqueous conditions (Figure 45).

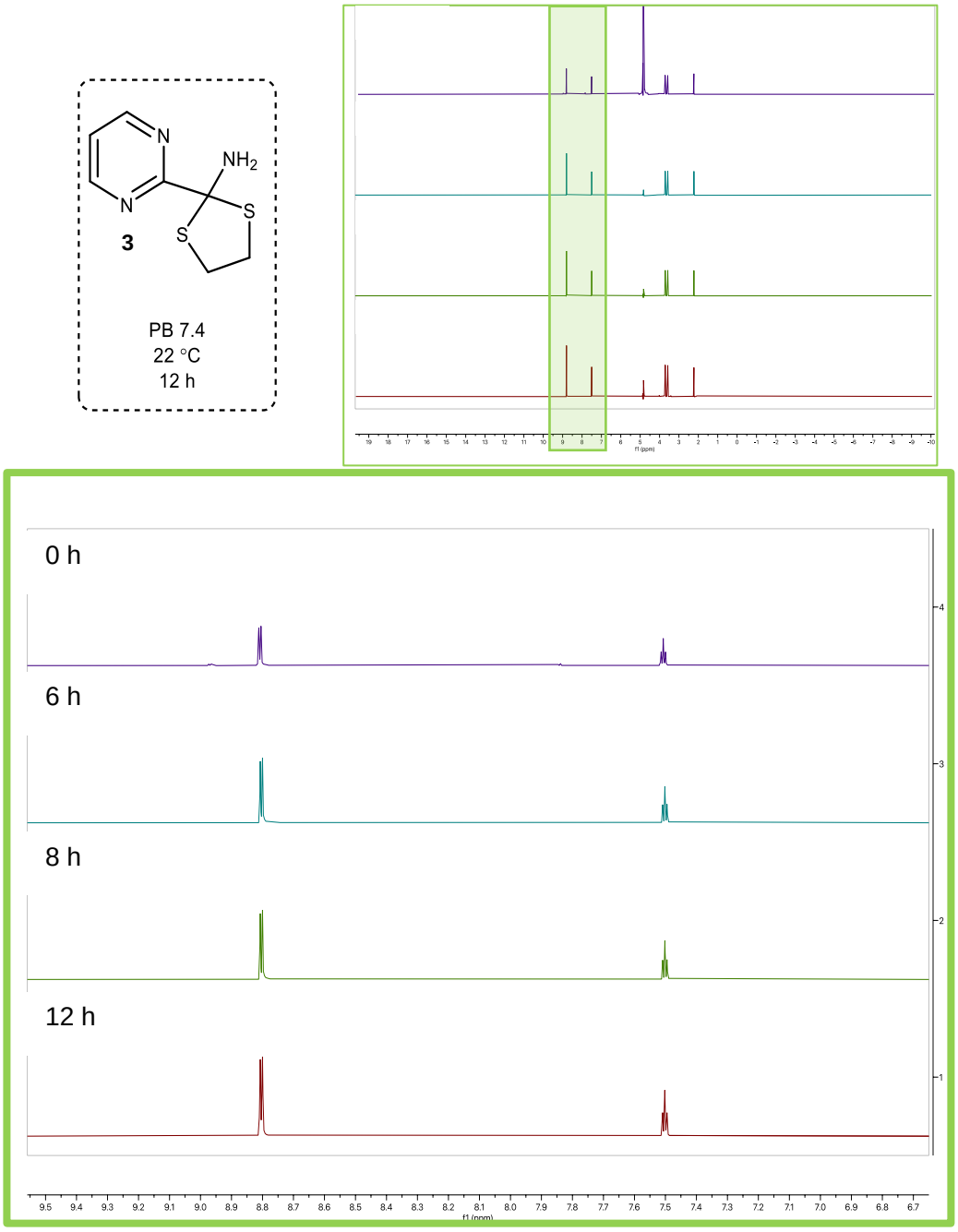


Figure 45. NMR aqueous stability. Aromatic region highlighted and expanded. Over the course of 12 h, no change in the product (**3**) can be seen, indicating aqueous stability.

The next stability experiment investigated the stability of **3** to a stoichiometric quantity of ethanethiol (2 equiv.). The outcome followed that of the aqueous stability where no changes were seen to the NMR spectra over the span of 12 h, indicating stability towards external thiols (Figure 46).

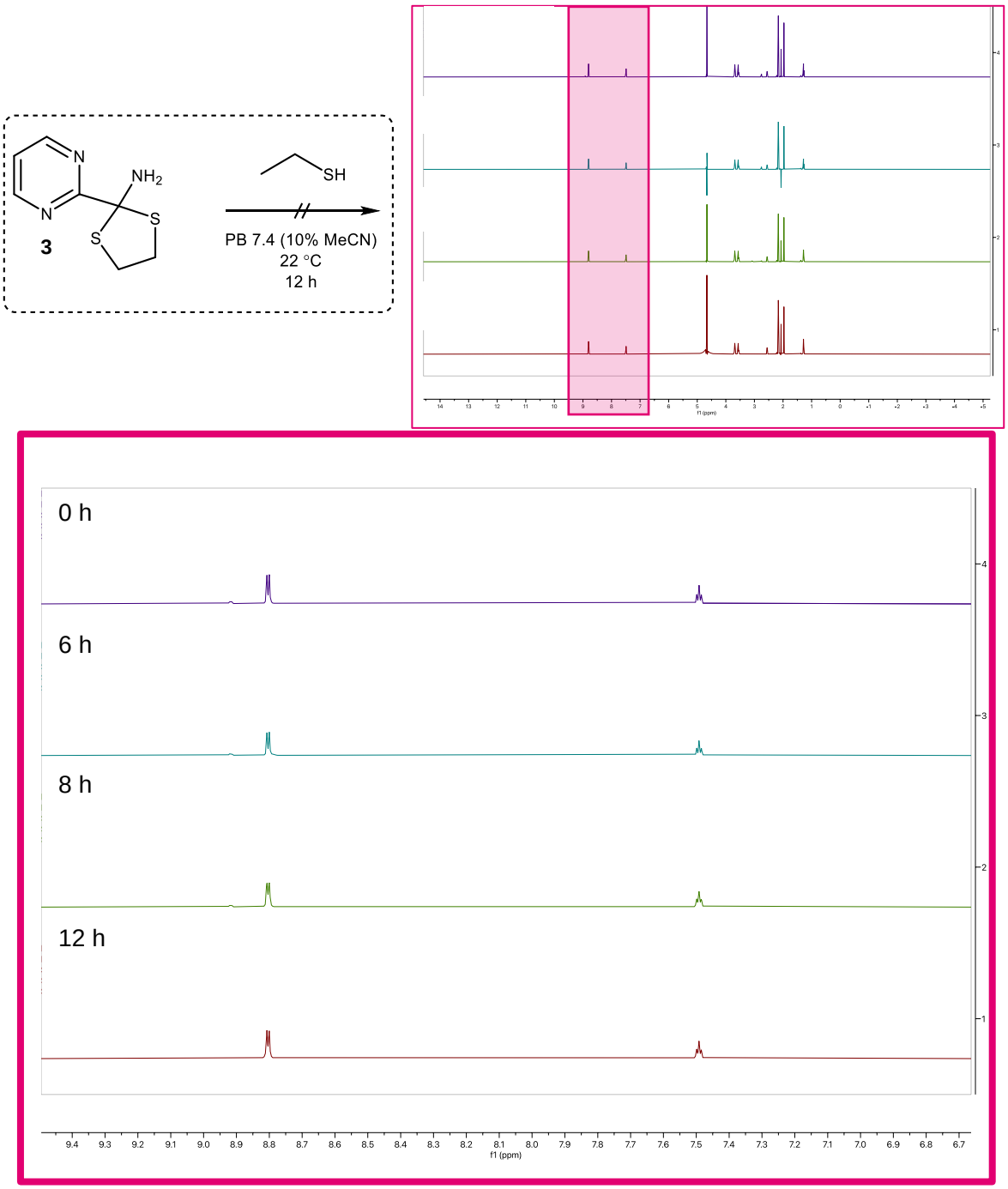


Figure 46. NMR ethanethiol stability. Aromatic region highlighted and expanded. Over the course of 12 h, no change in the product (**3**) can be seen, indicating stability towards stoichiometric amounts of monothiol.

The ADTA, **3** was found to stain with the oxidising reagent DTNB, and this created a basis for the hypothesis of selective reversibility. The hypothesis involved the use of NMM as a cleaving agent due to the rapid thiol reactivity (Figure 47 **A** - **D**). ADTA **3** was reacted with NMM (3 equiv.) in pH 7.4 phosphate buffer and rapid cleavage was observed upon which regeneration of the nitrile bond was observed alongside modification of the EDT moiety with NMM. The ^{13}C NMR shows the loss of the quaternary carbon (87.8 ppm) and reformation of the characteristic nitrile carbon peak (116.3 ppm) whilst the ^{1}H NMR, zoomed into the aromatic region, shows the reformation of the 2-cyano pyrimidine nitrile aromatic proton peaks (9.09 ppm, d and 7.96 ppm).

A

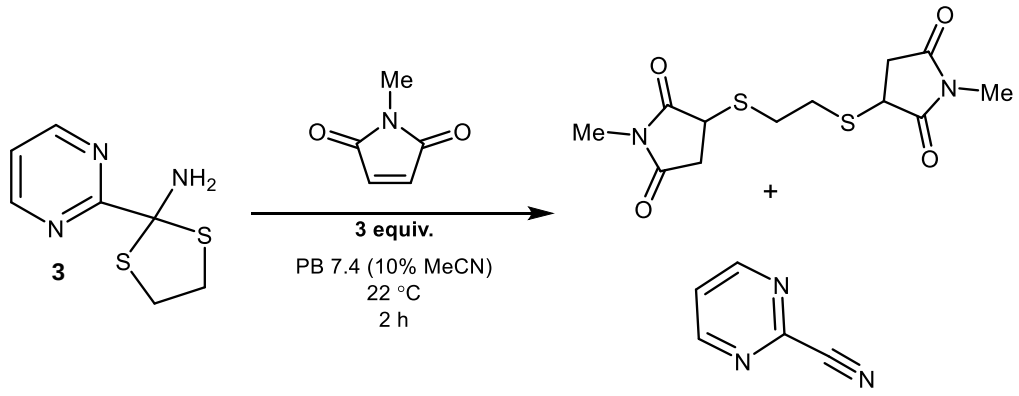


Figure 47. (**A**) ADTA **3** removal with NMM; reaction scheme for maleimide mediated removal of ADTA.

B

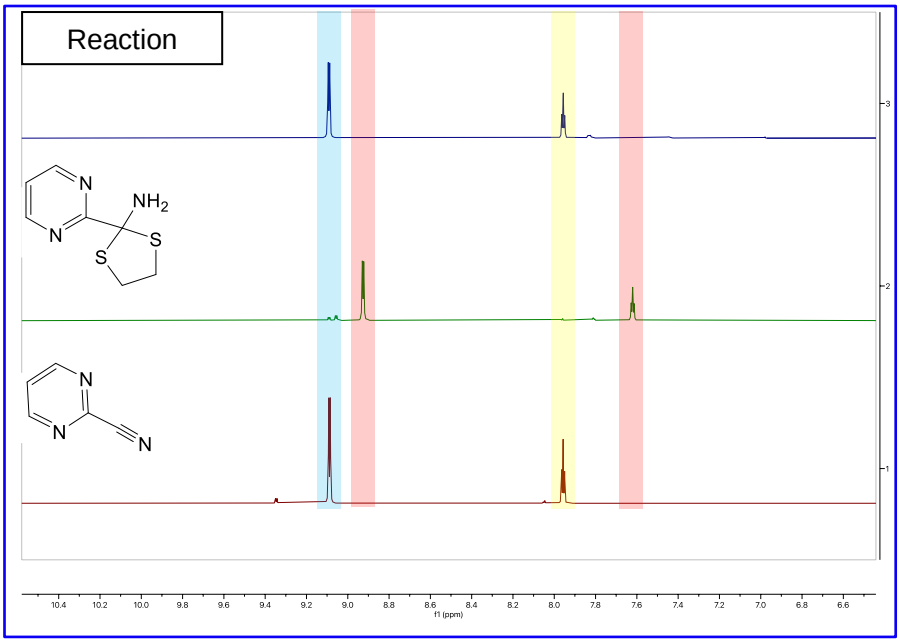


Figure 47. (**B**) ADTA **3** removal with NMM; Water suppression ^{1}H NMR zoomed into the aromatic region showing reformation of the 2-cyano pyrimidine nitrile aromatic proton peaks (9.09 ppm, d - blue and 7.96 ppm, t - yellow).

C

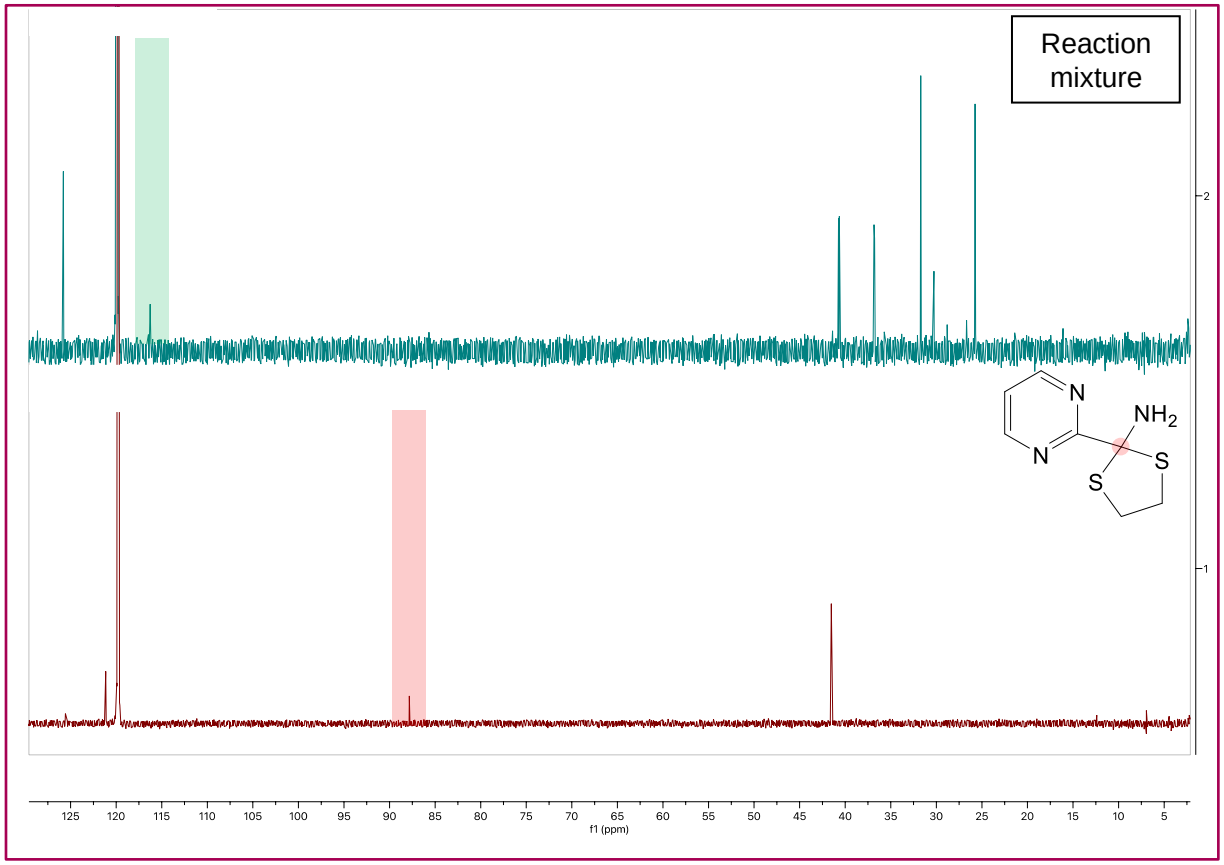


Figure 47. (**C**) ADTA **3** removal with NMM; ^{13}C NMR showing the loss of the quaternary carbon (87.83 ppm), highlighted red (bottom) and reformation of the characteristic nitrile carbon peak (116.27 ppm), highlighted green (top).

D

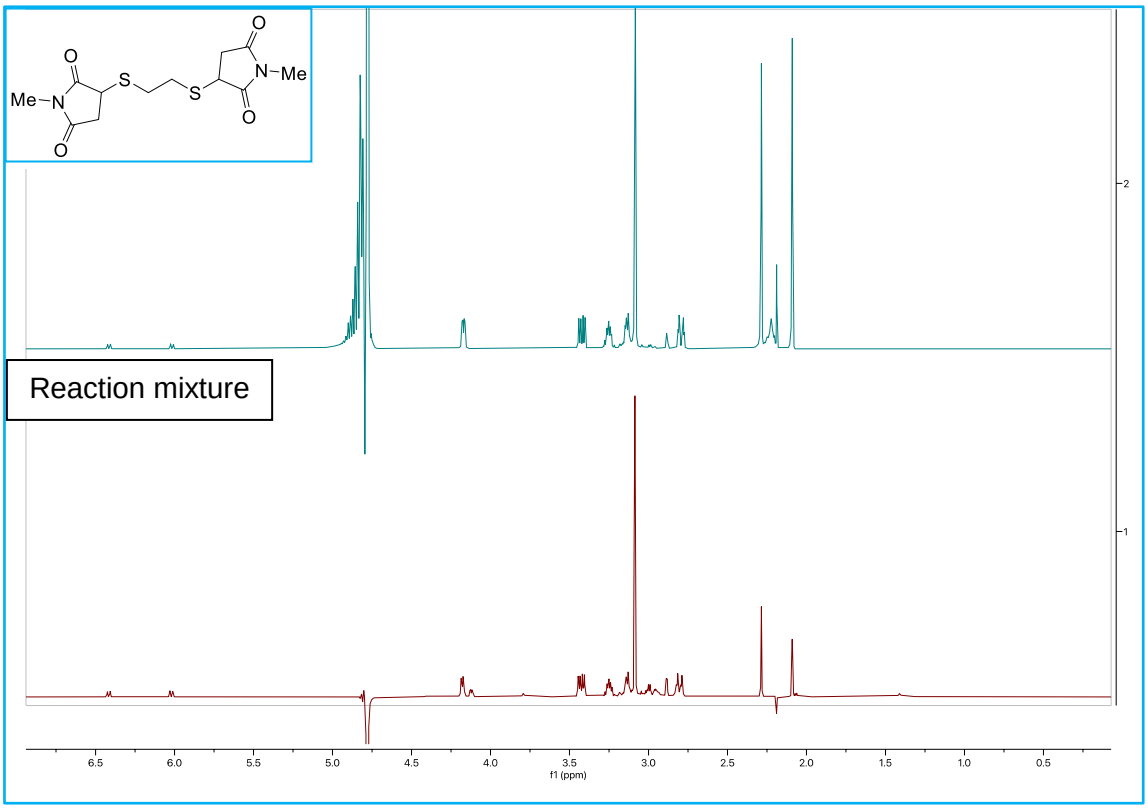


Figure 47. (**D**) ADTA **3** removal with NMM; Stacked comparison between the EDT maleimide product (green), formed *via* the reaction of EDT (1 equiv.) and NMM (3 equiv.) under identical conditions, and the reaction mixture after 2 h. Full spectra available in ESI (**S4**).

4.3. Summary and Further work

The NBT reaction provides a novel platform, upon which a new series of reversible bioconjugation tools can be developed. The reaction is compatible with aqueous conditions and provides a notable degree of selectivity when employed in the presence of alternate mono- and di-thiols namely, ET and PDT. This reaction is also selectively reversible with the addition of highly thiol reactive reagents, such as NMM (2 h). This system can potentially be used in multiple lines of bioconjugation such as protein tags and pull-down assays for example.

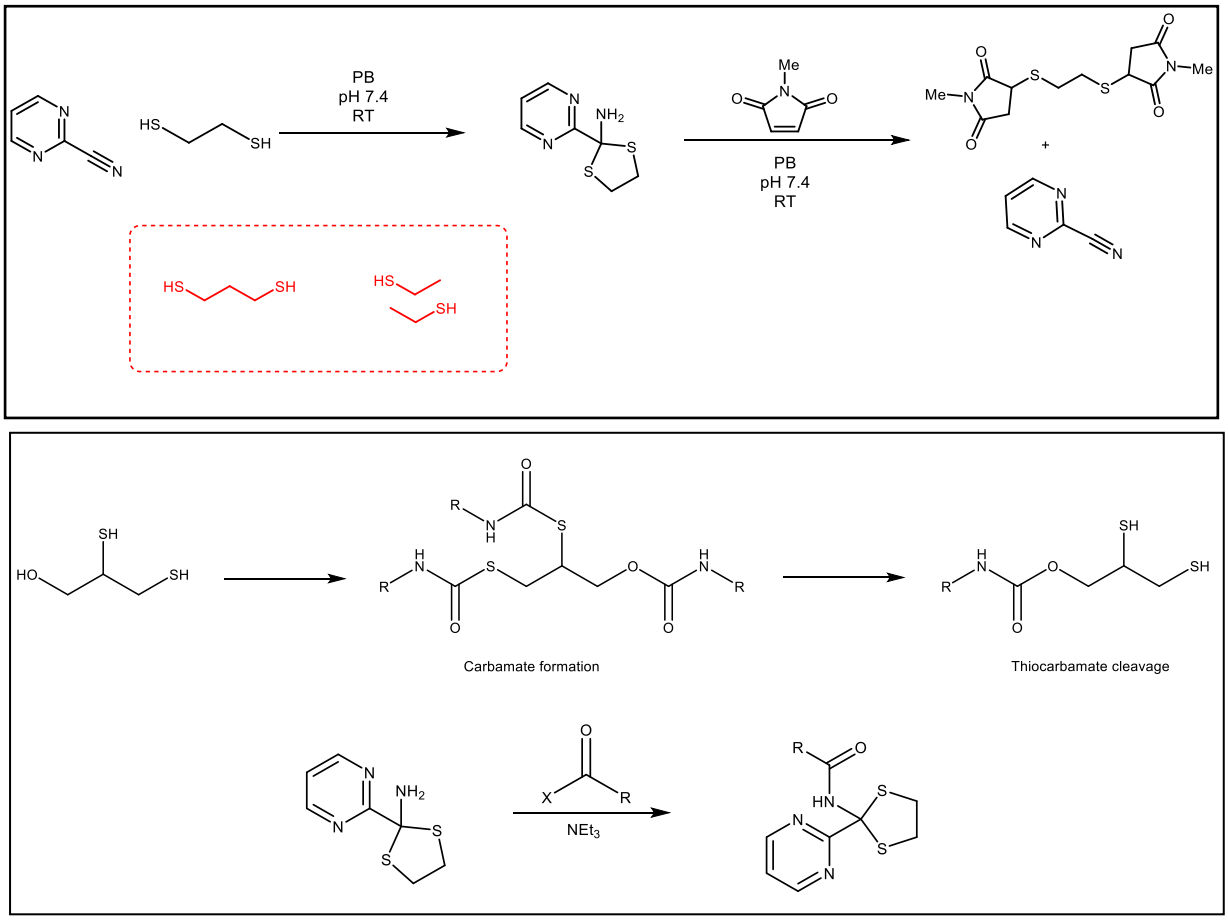


Figure 48. Summary and future plans for the NBT reaction.

The next steps in the development of this reaction as a bioconjugation tool would involve the synthesis of a functional derivative of ethanedithiol. A target compound could introduce a carbamate species as the functional handle starting from the commercially available; 2,3-dimercaptopropan-1-ol. This is one of many potential routes for functionality. The stability of ADTA **3**, should be further investigated with other highly thiol or nitrile reactive reagents in order to develop a broader picture regarding the reactivity of such constructs.

A fundamental aspect of the EDT-NBT reaction is the reversibility upon exposure to particular conditions e.g. exposure to NMM, this can be potentially counteracted *via* the acylation of the ADTA amino group. The resulting amide 'lock' could theoretically provide an irreversible system, adding further layers of complexity to the overall system.

5. Conclusion

This work investigated the use of nitrile reagents in multiple capacities in order to develop a new reactive group for protein modification. Aryl mono-nitriles were introduced as a new series of CLT reagents with varying degrees of reactivity. The conjugation mechanism follows a novel NBT style reaction followed by transfer to proximal lysines to generate amidine functionalised conjugates. A particularly promising reagent was found to be the *p*-amide pyrimidine nitrile, which achieved quantitative conversion to the desired product in just 30 min, with a sequential SPAAC functionalisation. The total reaction time of 6 h provided a final fluorescent CLT conjugate with a DAR of 1.5. These conjugates were found to be stable under multiple conditions including aqueous and serum incubation. Further work is still required as the conjugates are yet to undergo full LCMS/MS analysis to confidently confirm the location of the modification. Nitrile CLT is also yet to be applied on other protein systems other than trastuzumab Fab. The disulfide bond and the proximal lysines are located within the constant domains and therefore have a high chance of being present in another IgG1 antibodies, supporting the application on alternate mAb Fab fragments.

The second nitrile system further exploits the NBT reaction *via* the incorporation of a trapping group. The first of these reagents was the 2,3 bis-nitrile, and this reagent class was shown to generate clean disulfide bridges. Functional variants were developed in the form of quinoxaline 2,3 bis-nitrile which could subsequently undergo CuAAC and SPAAC mediated functionalisation. These reagents were also successfully proven on the more complex system of a full native antibody with the four interchain disulfides. Due to the nature of the trapping cascade, disulfide selectivity was thought to be inherent. The reagents were found to preferentially bridge the Fab disulfide when in the presence of up to 4 equiv. of free GSH. The bis-nitrile design was also shown to successfully modify and inhibit the physiologically prevalent enzyme thioredoxin. This provides early insights into possible uses for these reagents as enzyme inhibitors for therapeutic or mechanistic applications. Further work includes the development of this chemistry and reagents on other systems, to fully investigate the limits of the trapping system.

The final system developed the EDT-NBT reaction with 2-cyano pyrimidine. This reaction presents a remarkable level of selectivity when placed under competition conditions with both other bis and mono-thiols. The reaction was found to be highly rewarding with regards to yield,

providing ~90% in the span of an hour. The generated amino-dithioacetal was stable to aqueous incubation as well as exposure to stoichiometric equivalents of the mono-thiol ethanethiol. The amino-dithioacetal was reversible upon addition of the rapid thiol trapping reagent; NMM providing the reformation of 2-cyano pyrimidine. This reaction presents ample scope for use as a protein tag or use in pulldown assays and warrants further development.

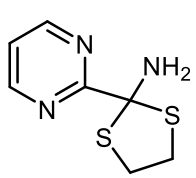
6. Experimental section

6.1. Chemical synthesis general remarks

All chemical reagents and solvents were purchased from commercial sources; Sigma UK, Fisher UK, Fluorochem or VWR UK and used as per manufacturer instructions. All chemical reactions were carried out at atmospheric pressure, under argon, unless stated otherwise. Room temperature (RT) is defined as between 15-25 °C. The term *in vacuo* refers to organic solvent removal using Buchi rotary evaporator between 15-60 °C under reduced pressure. Chemical reactions were monitored using thin layer chromatography (TLC) on pre-coated silica gel plates (254 µm) purchased from VWR, UK. Detection of synthesised compounds was done by UV (254 nm and 365 nm) or chemical stain ($KMnO_4$, ninhydrin). Flash column chromatography was carried out using pre-loaded FlashPure Eco Flex column on a Biotage Isolera Spektra One flash chromatography system. ^{1}H NMR and ^{13}C NMR were obtained at ambient temperature on a Bruker Advance AMX600 instrument operating at 500 MHz, 600 MHz, or 700 MHz ^{1}H and 126 MHz, 151 MHz, or 176 MHz for ^{13}C in the stated solvent. Chemical shifts (δ) are reported in parts per million (ppm) and coupling constants (*J*) in Hertz (Hz). The multiplicity of each signal is indicated as s-singlet, d-doublet, t-triplet, q-quartet, quin-quintet, m-multiplet (i.e., complex peak obtained due to overlap) or a combination of these. All assignments were made with the aid of DEPT, COSY, HSQC, HMBC, or NOESY correlation experiments. Infra-red spectra were recorded on a Bruker ALPHA FT-IR spectrometer operating in ATR mode, with frequencies given in reciprocal centimetres (cm^{-1}). Melting points were taken on a Gellenkamp apparatus and are uncorrected. High- and low-resolution mass spectra were recorded on a VG70 SE mass spectrometer, operating in modes ESI, EI, or CI (+ or -) depending on the sample, at the Department of Chemistry, University College London. Buffers used in organic reactions were prepared with double-deionised water and filter sterilised (0.20 µm).

6.2. Synthesis and characterisation of compounds

2-(Pyrimidin-2-yl)-1,3-dithiolan-2-amine (**3**)

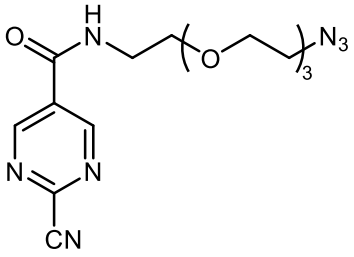


2-Cyano pyrimidine (105 mg, 1 mmol) was dissolved in MeCN (4 mL) followed by addition of 50 mM phosphate buffer pH 7.4 (16 mL). To this mixture was added ethanedithiol (109 µL, 1.2 mmol), the reaction was then stirred at room temperature for 1 h. The reaction was then

extracted with EtOAc (3 x 10 mL). The organic layer was dried over $MgSO_4$ and purified by flash column chromatography (Flashpure, 4 g, eluent 0-10% MeOH: CH_2Cl_2) to afford the title product as a yellow-brown oil (181 mg, 0.91 mmol, 91%). ^{1}H NMR (700 MHz, $CDCl_3$) δ 8.71 (d, 2H, *J* = 4.8 Hz, Ar***H***), 7.15 (t, 1H, *J* = 4.8 Hz, Ar***H***), 3.61 (s, 4H, C***H₂***C***H₂***). ^{13}C NMR (176 MHz, $CDCl_3$) δ 171.5, 157.3, 119.4, 88.5 (***C***), 41.6. HRMS (ESI) theoretical $[C_7H_9N_3S_2\text{-}NH_2]^+$: 183.0051, measured: 183.0046. HRMS (ESI) theoretical full $[C_7H_9N_3S_2\text{-}H]^-$: 198.0154, measured: 198.0154.

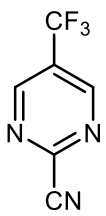
Title product presents a highly unpleasant smell, an IR spectrum was therefore not obtained.

N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide (**5**)



2-Cyanopyrimidine-5-carboxylic acid (80 mg, 0.54 mmol) and EEDQ (160 mg, 0.65 mmol) were dissolved in MeCN (10 mL). The resultant solution was then stirred at RT for 0.5 h. 2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethan-1-amine (95.8 µL, 0.48 mmol) was added and the reaction was stirred at RT for 14 h after which the solvent was removed *in vacuo*. The crude mixture was purified by flash column chromatography (Flashpure, 4 g, eluent 0-100% EtOAc:cyclohexane, followed by 0-10% MeOH:EtOAc) to yield the title product as a colourless oil (85 mg, 0.24 mmol, 51%).^{1}H NMR (500 MHz, $CDCl_3$) δ 9.28 (2H, s, Ar***H***), 3.70-3.66 (12H, m, C***H₂***), 3.62 (2H, t, *J* = 5 Hz, C***H₂***), 3.39 (2H, t, *J* = 5 Hz, C***H₂***). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.1, 157.3, 146.2, 129.6, 115.5 (***CN***), 70.7, 70.6, 70.5, 70.3, 70.0, 69.5, 50.6, 40.4. IR ν_{max} / cm^{-1} 3345, 3067, 2954, 2159, 2117. HRMS (ESI) theoretical $[C_{14}H_{19}N_7O_4+H]^+$: 350.1566, measured: 350.1571.

5-(Trifluoromethyl)pyrimidine-2-carbonitrile (**6**)[250]

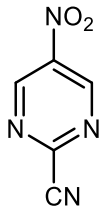


To a solution of DABCO (4.6 mg, 0.041 mmol) in DMSO:H_2O (0.6:0.25 mL) was added finely ground sodium cyanide (15 mg, 0.31 mmol). The flask was then sealed under an atmosphere

of nitrogen and cooled to 0 °C. 2-chloro-5-(trifluoromethyl)pyrimidine (50.3 mg, 0.28 mmol) dissolved in DMSO (0.65 mL) was then added dropwise over 15 min, the reaction was then stirred at RT for 2 h. The reaction was then quenched with water (5 mL) and extracted with diethyl ether (3 x 5 mL). The combined organic layers were washed with brine (1 x 3 mL). The organic phase was dried over $MgSO_4$ and concentrated *in vacuo*. Crude 5-(trifluoromethyl)pyrimidine-2-carbonitrile was purified by flash column chromatography (Flashpure, 4 g, eluent 0-75% CH_2Cl_2:hexane) to obtain the title product as a pale-yellow oil (33.4 mg, 0.19 mmol, 69%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 9.13 (2H, s, Ar***H***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 155.6 (q, J_{CF} = 3.6 Hz), 147.6, 127.0 (q, J_{CF} = 35.2 Hz), 121.8 (q, J_{CF} = 273.2 Hz), 114.8 (***CN***). IR v_{max} / cm^{-1} 2983 (C-H sp^2), 2166 (C≡N), 1328 (C-F). HRMS (ESI) theoretical [$C_6H_2N_2F_3$]: 173.0201, Measured : 173.0200.

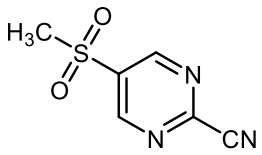
Caution: Title compound is highly volatile.

5-Nitropyrimidine-2-carbonitrile (**7**)



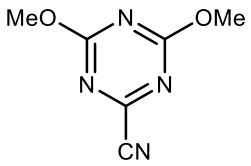
To a solution of DABCO (11.2 mg, 0.1 mmol) in DMSO:H_2O (5:0.5 mL) was added finely ground sodium cyanide (27.0 mg, 0.55 mmol). The flask was then sealed under an atmosphere of nitrogen and cooled to 0 °C. 2-chloro-5-nitropyrimidine (79.8 mg, 0.5 mmol) dissolved in DMSO (0.65 mL) was then added dropwise over 15 min. The resulting solution was stirred at RT for 2 h. After this time water (5 mL) and EtOAc (3 mL) were added to the mixture. The aqueous layer was extracted with EtOAc (3 x 3 mL). The combined organic phases were then washed with water (3 x 3 mL) and brine (5 mL). The organic phase was dried over $MgSO_4$ and concentrated *in vacuo*. Crude 5-nitropyrimidine-2-carbonitrile was purified by flash column chromatography (Flashpure, 4 g, eluent 10-60% EtOAc:cyclohexane) to obtain the title product as a white solid (58.5 mg, 0.39 mmol, 78%). m.p 53-55°C; ^{1}H NMR (600 MHz, $CDCl_3$) δ 9.61 (2H, s, Ar***H***); ^{13}C (151 MHz, $CDCl_3$) δ 153.8, 147.8, 142.4, 114.5 (***CN***). IR v_{max} / cm^{-1} 3059 (C-H sp^2), 2917 (C-H sp^3), 1590 (NO_2). HRMS (ESI) theoretical $[C_5H_2N_4O_2+H]^+$: 150.0183, Measured : 150.0183.

5-(Methylsulfonyl)pyrimidine-2-carbonitrile (**8**)



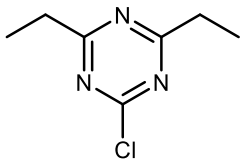
A mixture of 5-bromopyrimidine-2-carbonitrile (500 mg, 2.72 mmol), sodium methane sulfinate (422 mg, 4.14 mmol) and DMSO (5 mL) was stirred at 100 °C under a N_2 atmosphere for 20 h. The reaction was then cooled to RT and water was added (5 mL). The reaction was extracted with 3 x 10 mL EtOAc. The organic layer was washed with water (15 mL) and brine (15 mL) and dried over Na_2SO_4. The organic layer was concentrated *in vacuo* and purified with flash column chromatography (Flashpure , 12 g, eluent 20-50% EtOAc:cyclohexane) to obtain the title product as an off-white solid (55 mg, 0.30 mmol, 11 %). m.p 190-193 °C; ^{1}H NMR (600 MHz, CD_3CN) δ 9.33 (2H, s, Ar***H***), 3.25 (3H, s, C***H$_3$***); ^{13}C (151 MHz, CD_3CN) δ 158.8, 148.1, 138.1, 116.2 (***CN***), 44.8. IR ν_{max} / cm^{-1} 3063 (C-H sp^2), 2925 (C-H sp^3), 1550 (S=O). HRMS (ESI) theoretical $[C_6H_5N_3O_2S+H]^+$: 184.0175, Measured : 184.0175.

4,6-Dimethoxy-1,3,5-triazine-2-carbonitrile (**9**)



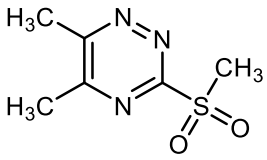
To a solution of 2-chloro-4,6-dimethoxy-1,3,5-triazine (87.8 mg, 0.5 mmol) in DMSO (5 mL) was added DABCO (11.2 mg, 0.1 mmol) and NaCN (27.0 mg, 0.55 mmol) followed by dropwise addition of water (0.5 mL). The resulting solution was stirred at RT for 1.5 h. After this time water (5 mL) and EtOAC (3 mL) were added to the mixture. The aqueous layer was extracted with EtOAc (3 x 3 mL). The combined organic phases were then washed with water (3 x 3 mL) and brine (5 mL). Organic phase was dried over $MgSO_4$ and concentrated *in vacuo*. Crude 4,6-dimethoxy-1,3,5-triazine-2-carbonitrile was purified by flash column chromatography (Flashpure, 4 g, eluent 10-60% EtOAc:cyclohexane) to obtain the title product as a clear oil (68.1 mg, 0.41 mmol, 82%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 4.10 (6H, s, C***H$_3$***); ^{13}C NMR (151 MHz, $CDCl_3$) δ 172.7, 154.9, 114.1 (***CN***), 56.5. IR ν_{max} / cm^{-1} 3039 (C-H sp^2), 2921 (C-H sp^3), 2244 (C≡N); HRMS (ESI) theoretical $[C_6H_6N_4O_2+H]^+$: 167.0564, Measured : 167.0560.

2-Chloro-4,6-diethyl-1,3,5-triazine (**10**)



To a solution of cyanuric chloride (300 mg, 1.63 mmol) and CuI (10.7 mg, 0.056 mmol) in THF (3 mL) was added ethylmagnesium bromide (1 M solution in THF, 4 mL, 4 mmol) dropwise at -15 °C over 20 min under N_2. The reaction was then stirred for a further 3 h at -15 °C, after which the reaction mixture was quenched by the addition of sat. NH_4Cl (10 mL) and extracted with Et_2O (3 x 10 mL). The organic phase was then washed with H_2O, dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude product was purified by flash column chromatography (Flashpure, 12 g, eluent 0-20% EtOAc:cyclohexane) to afford the title product as a yellow oil (89%, 249 mg, 1.45 mmol). ^{1}H NMR (700 MHz, $CDCl_3$) δ 2.93 (4H, q, J = 7.6 Hz, C***H₂***), 1.36 (6H, t, J = 7.6 Hz, C***H₃***). ^{13}C NMR (176 MHz, $CDCl_3$) δ 185.0, 171.9, 32.1, 11.3. IR v_{max} / cm^{-1} 2926 (C-H sp^3). HRMS (ESI) theoretical $[C_7H_{10}N_3{}^{35}Cl+H]^+$: 172.0636, Measured : 172.0634.

5,6-Dimethyl-3-(methylsulfonyl)-1,2,4-triazine[215]

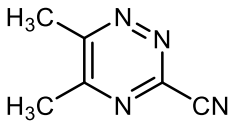


To a solution of freshly distilled 2,3-butanedione (0.226 mL, 2.57 mmol) and sodium bicarbonate (198 mg, 2.36 mmol) in water (4 mL) at RT was added a solution of *S*-methyl isothiosemicarbazide hydroiodide (500 mg, 2.15 mmol) in water (2.5 mL). The mixture was stirred for 16 h at RT. The aqueous layer was then extracted with dichloromethane (DCM) (3 x 10 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Crude 5,6-dimethyl-3-(methylthio)-1,2,4-triazine was carried forward in the next step.

Under a nitrogen atmosphere, mCPBA (753 mg, 4.36 mmol) was added in small portions at 0 °C to a solution of 5,6-dimethyl-3-(methylthio)-1,2,4-triazine (278 mg, 1.79 mol) in dry DCM (11 mL). After complete addition, the mixture was stirred for 2 h at RT and the suspension was filtered through a pad of Celite®. The filtrate was then washed with an aq. saturated solution of sodium bicarbonate (2 x 50 mL) and an aqueous saturated solution of sodium thiosulfate (2 x 50 mL). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (Flashpure, 12 g, eluent 10- 80%

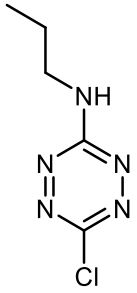
EtOAc:cyclohexane) to obtain the title product as a white solid (175 mg, 0.93 mmol, 52%). m.p 130–131°C; ^{1}H NMR (600 MHz, $CDCl_3$) δ 3.46 (3H, s, Ar-C***H₃***), 2.81 (3H, s, S-C***H₃***), 2.71 (3H, s, Ar-C***H₃***); ^{13}C NMR (151 MHz, $CDCl_3$) δ 165.3, 162.3, 161.2, 39.8, 22.2, 20.1. IR ν_{max} / cm^{-1} 3009 (C-H sp^2), 2927 (C-H sp^3), 1313 (S=O); HRMS (ESI) theoretical $[C_6H_9N_3O_2S+H]^+$: 188.0488, Measured : 188.0488.

5,6-Dimethyl-1,2,4-triazine-3-carbonitrile (**11**)



To a solution of 5,6-dimethyl-3-(methylsulfonyl)-1,2,4-triazine (93.6 mg, 0.5 mmol) in DMF (3 mL) was NaCN (27.5 mg, 0.56 mmol). The resulting solution was stirred at RT for 1.5 h. After this time water (5 mL) and EtOAc (3 mL) were added to the mixture. The aqueous layer was extracted with EtOAc (3 x 3 mL). The combined organic phases were then washed with water (3 x 3 mL) and brine (5 mL). The organic phase was dried over $MgSO_4$ and concentrated *in vacuo*. Crude 4,6-dimethoxy-1,3,5-triazine-2-carbonitrile was purified by flash column chromatography (Flashpure, 4 g, eluent 10-60% EtOAc:cyclohexane) to obtain the title product as a brown oil (40.9 mg, 0.305 mmol, 61%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 2.79 (3H, s, Ar-C***H₃***), 2.63 (3H, s, Ar-C***H₃***); ^{13}C NMR (151 MHz, $CDCl_3$) δ 160.3, 160.1, 145.6, 114.7 (***CN***), 21.9, 20.3. IR ν_{max} / cm^{-1} 2926 (C-H sp^3), 2333 (C≡N); HRMS (ESI) theoretical $[C_6H_6N_4+H]^+$: 135.0665 Measured : 135.0662.

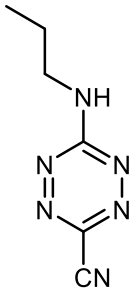
6-Chloro-*N*-propyl-1,2,4,5-tetrazin-3-amine



To a stirred solution of 3,6-dichlorotetrazine (100 mg, 0.66 mmol) in CH_2Cl_2 (4.5 mL) was added propylamine (60 µL, 0.7 mmol), then DIPEA (0.45 mL, 0.26 mmol) after 30 min at RT, silica gel was added. The mixture was then concentrated *in vacuo* and the resultant free-flowing powder was purified by silica gel column chromatography (Biotage 4 g column, 5-40% EtOAc:cyclohexane) to give the title product as a bright orange gum (94.7 mg, 0.54 mmol, 82%). ^{1}H NMR (700 MHz, $CDCl_3$) δ 6.19 (1H, s, N***H***), 3.55 (2H, q, J = 7.3 Hz, NC***H₂***), 1.73 (2H, sx, 7.3 Hz, C***H₂***), 1.01 (3H, t, J = 7.3 Hz, C***H₃***). ^{13}C NMR (125 MHz, $CDCl_3$) δ 161.8 (Tz), 160.7

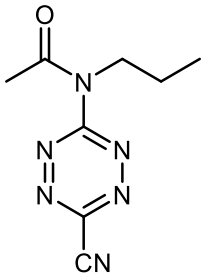
(Tz), 43.6, 22.4, 11.3. IR ν_{max} / cm^{-1} 3272 (N-H), 2962 (C-H sp^3), 1725, 1582. HRMS (ESI): theoretical $[C_5H_8N_5Cl+H]^+$ 174.0541, Measured : 175.0540.

6-(Propylamino)-1,2,4,5-tetrazine-3-carbonitrile (**12**)



To a stirred solution of 6-chloro-N-propyl-1,2,4,5-tetrazin-3-amine (100 mg, 0.57 mmol) in DMF (1 mL) at 0 ºC was added sodium cyanide (28 mg, 0.57 mmol) then collidine (74 uL, 0.11 mmol). The resulting solution turned brown. After 2 h, the solution was diluted with toluene (5 mL) and silica gel was added. The volatiles were removed *in vacuo* and the crude mixture was purified by silica gel column chromatography (Biotage 4 g column, 5-65% EtOAc:cyclohexane) to give the title product as an orange-red oil (56.2 mg, 0.34 mmol, 60%).^{1}H NMR (600 MHz, $CDCl_3$) δ 3.68 (2H, q, J = 7.4 Hz, NC***H₂***), 1.78 (2H, sx, J = 7.4 Hz, C***H₂***CH_3), 1.05 (3H, t, J = 7.4 Hz, C***H₃***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 159.1, 145.5, 113.9 (**CN**), 43.6, 22.3, 11.4. IR ν_{max} / cm^{-1} 2966 (C-H sp^3), 2245 (C≡N), 1582. HRMS (ESI): theoretical $[C_6H_8N_6+H]^+$: 165.0883, Measured: 165.0885.

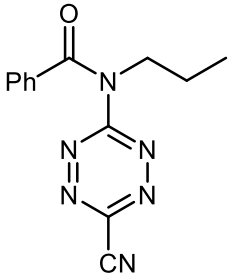
N-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylacetamide (**13**)



To a stirred solution of 6-(propylamino)-1,2,4,5-tetrazine-3-carbonitrile (25.8 mg, 0.15 mmol) in THF (0.30 mL) at RT was added DMAP (2.8 mg, 0.02 mmol), then Ac_2O (32 µL, 0.34 mmol). The reaction was stirred at RT for 15 minutes. Silica gel was added, and the mixture was concentrated *in vacuo*. The free-flowing powder residue was submitted to silica gel column chromatography (Biotage 4 g column, 4-30% EtOAc:cyclohexane), giving the title product as a pink oil (16 mg, 0.077 mmol, 51%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 4.26 (2H, t, *J* = 7.4 Hz, NC***H₃***), 2.66 (3H, s, COC***H₃***), 1.72 (2H, sx, *J* = 7.4 Hz, C***H₂***), 0.98 (3H, t, *J* = 7.4 Hz, C***H₃***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 171.2, 160.3, 147.6, 112.9 (***CN***), 47.6, 27.4, 21.2, 11.3. IR ν_{max} / cm^{-}

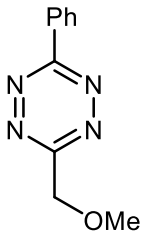
[1] 2925 (C-H sp^3), 1707 (C=O). HRMS (ESI): theoretical [$C_8H_{10}ON_6$]: 206.09106, Measured : 206.0916.

N-(6-Cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylbenzamide (**14**)



To a stirred solution of 6-(propylamino)-1,2,4,5-tetrazine-3-carbonitrile (24 mg, 0.15 mmol) in CH_2Cl_2 (1 mL) at RT was added DMAP (4.46 mg, 0.032 mmol), NEt_3 (40.7 µL, 0.29 mmol) and then benzoyl chloride (21 µL, 0.175 mmol). The reaction was stirred at RT for 1 h. Silica gel was added, and the mixture was concentrated *in vacuo*. The free-flowing powder residue was submitted to silica gel column chromatography (Biotage 4 g column, 0-30% EtOAc:cyclohexane), giving the title product as a pink oil (13 mg, 0.049 mmol, 32%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.54 (1H, t, J = 7.4 Hz, Ar***H***), 7.47 (2H, d, J = 7.4 Hz, Ar***H***), 7.39 (2H, t, J = 7.4 Hz, Ar***H***), 4.37 (2H, t, J = 7.3 Hz, NC***H₂***), 1.91 (2H, sx, J = 7.3 Hz, C***H₂***), 1.06 (3H, t, J = 7.3 Hz, C***H₃***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 171.4, 161.6, 147.0, 134.7, 134.2, 133.1, 130.7, 129.3, 128.7, 112.9 (***CN***), 49.7, 21.8, 11.5. IR ν_{max} / cm^{-1} 3064 (C-H sp^2), 2934 (C-H sp^3), 1711 (C=O). HRMS (ESI) theoretical $[C_{13}H_{12}N_6O+H]^+$: 269.1145, Measured : 269.1141.

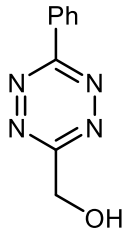
3-(Methoxymethyl)-6-phenyl-1,2,4,5-tetrazine (**15**)



Methoxyacetonitrile (1.56 mL, 21.0 mmol), phenyl nitrile (0.72 mL, 7.0 mmol) and $Ni(OTf)_2$ (1.25 g, 3.5 mmol) added to 64% hydrazine hydrate (4.50 mL, 35 mmol). The purple-red solution was stirred at 60 °C in a sealed tube under N_2 for 16 h before being cooled to RT and diluted with H_2O (10 mL). The reaction mixture was poured into a solution of sodium nitrate (2.42 g, 35.0 mmol) in H_2O (10 mL) at 0 °C. The mixture was slowly acidified with 4M HCl until the effervescence stopped (Caution: toxic nitrogen oxide gas is formed). The reaction was then extracted with EtOAc (5 x 30 mL). The organic layers were combined and dried over $MgSO_4$. The solution was concentrated *in vacuo* to give the crude product as a purple oil, which was purified by flash column chromatography (Flashpure, 12 g, eluent 0-10% EtOAc:cyclohexane), affording the title product as a pink solid (570 mg, 2.81 mmol, 40%). mp

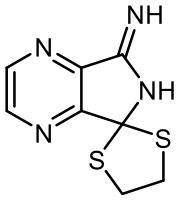
57-60 °C. ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.63 (2H, d, *J* = 7.1 Hz, Ar***H***), 7.67 – 7.63 (1H, m, Ar***H***), 7.63 – 7.59 (2H, m, Ar***H***), 5.11 (2H, s, C***H₂***), 3.64 (3H, s, C***H₃***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 166.3, 165.2, 133.2, 131.6, 129.5, 128.4, 72.4, 59.8. IR ν_{max} / cm^{-1} 3051 (C-H sp^2), 2930 (C-H sp^3). HRMS (ESI) theoretical $[C_{10}H_{11}ON_4]^+$ 203.0929, Measured 203.0927.

(6-Phenyl-1,2,4,5-tetrazin-3-yl)methanol (**16**)



3-(Methoxymethyl)-6-phenyl-1,2,4,5-tetrazine (100 mg, 0.5 mmol) was dissolved in anhydrous DCM (1 mL). The flask was sealed, cooled to -70 °C and put under an N_2 atmosphere. To this solution was added BBr_3 (1.0 M solution in DCM, 1 mL, 1 mmol) dropwise over 10 min resulting in an orange solution. The reaction was stirred at -70 °C for 1 h before being gently warmed to RT with further stirring for another 2 h. The solution was then diluted with ice cold dH_2O upon which the vibrant purple colour returned. The aqueous layer was extracted with CH_2Cl_2 (3 x 5 mL), and the combined organic layers were washed with brine and dried over $MgSO_4$. The solvent was removed *in vacuo* and the crude product was purified by flash column chromatography (Flashpure, 4 g, eluent 0-35% EtOAc:cyclohexane), affording the title product as a purple solid. (88.5 mg, 0.47 mmol, 94%). mp 90-95 °C. ^{1}H NMR (600 MHz, CDCl3) δ 8.63 (2H, d, J = 7.1 Hz, Ar***H***), 7.68 – 7.64 (1H, m, Ar***H***), 7.64 – 7.60 (2H, m, Ar***H***), 5.34 (2H, d, J = 6.3 Hz, C***H₂***), 3.05 (1H, t, J = 6.3 Hz, O***H***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 167.4, 165.6, 133.2, 131.6, 129.5, 128.4, 63.0. IR ν_{max} / cm^{-1} 3212 (O-H), 3047 (C-H sp^2), 2924 (C-H sp^3). HRMS (ESI) theoretical $[C_9H_8ON_4+H]^+$ 189.0771, Measured 189.0771.

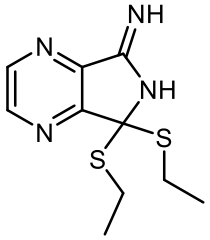
Spiro[pyrrolo[3,4-b]pyrazine-5,2'-[1,3]dithiolan]-7(6H)-imine (**25**)



Pyrazine-2,3-dicarbonitrile (130 mg, 1 mmol) was dissolved in MeCN (4 mL) followed by addition of 50 mM phosphate buffer pH 7.4 (16 mL). To this mixture was added ethanedithiol (109 µL, 1.2 mmol), the reaction was then stirred at room temperature for 1 h. The reaction was then extracted with EtOAc (3 x 10 mL). The organic layer was dried over and $MgSO_4$ purified by flash column chromatography (Flashpure, 4 g, eluent 0-10% MeOH:CH_2Cl_2) to afford the title product as a red oil (206 mg, 0.92 mmol, 92%). ^{1}H NMR (500 MHz, $CDCl_3$) δ

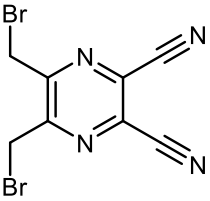
8.61 (1H, d, *J* = 2.7 Hz, Ar***H***), 8.51 (1H, d, *J* = 2.7 Hz, Ar***H***), 3.79 (4H, s, C***H₂***C***H₂***). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.8, 161.1, 144.8, 144.0, 143.8, 90.5 (**C**), 41.7. IR ν_{max} / cm^{-1} 3345 (N-H), 3082 (C-H sp^2), 2930 (C-H sp^3), 1657 (C=NH). HRMS (ESI) theoretical $[C_8H_8N_4S_2+H]^+$: 225.0258, measured: 225.0263.

7,7-Bis(ethylthio)-6,7-dihydro-5H-pyrrolo[3,4-b]pyrazin-5-imine (**26**)



2-Cyano pyrimidine (105 mg, 1 mmol) was dissolved in MeCN (4 mL) followed by addition of 50 mM phosphate buffer pH 7.4 (16 mL). To this mixture was added ethanedithiol (109 µL, 1.2 mmol), the reaction was then stirred at room temperature for 1 h. The reaction was then extracted with EtOAc (3 x 10 mL). The organic layer was dried over $MgSO_4$ and purified by flash column chromatography (Flashpure, 4 g, eluent 0-10% MeOH:CH_2Cl_2) to afford the title product as a red oil (229 mg, 0.90 mmol, 90%).^{1}H NMR (600 MHz, $CDCl_3$) δ 8.60 (1H, d, *J* = 2.3 Hz, Ar***H***), 8.56 (1H, d, *J* = 2.3 Hz, Ar***H***), 2.88-2.73 (4H, m, 2 x C***H₂***), 1.19 (6H, t, *J* = 7.6 Hz, 2 x C***H₃***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 165.9, 161.2, 144.4, 144.1, 143.4, 81.8 (**C**), 25.4, 14.3. IR ν_{max} / cm^{-1} 3344 (N-H), 3080 (C-H sp^2), 2931 (C-H sp^3), 1657 (C=NH). HRMS (ESI) theoretical $[C_{10}H_{14}N_4S_2+H]^+$: 255.0729, measured: 255.0733.

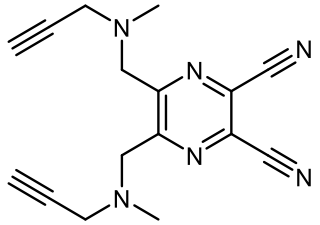
5,6-Bis(bromomethyl)pyrazine-2,3-dicarbonitrile (**29**)[251]



A solution of 1,4-dibromo-butane-2,3-dione (2.0 g, 8.2 mmol) in MeCN (8 mL) was added to a solution of diaminomaleonitrile (887 mg, 8.2 mmol) in MeCN (8 mL). The solution was heated at 70 °C for 1 h, the solvent was removed *in vacuo* and the resulting solid residue was purified with flash column chromatography (Flashpure, 12 g, eluent 100% CH_2Cl_2) to obtain the title product as a pale yellow solid (1.6 g, 5.0 mmol, 61 %); m.p. 136-139 °C. ^{1}H NMR (600 MHz, CD_3CN) δ 4.78 (s, 4H, BrC***H₂***); ^{13}C NMR (151 MHz, CD_3CN) δ 156.1, 133.2, 114.2 (***CN***), 29.0. IR ν_{max} / cm^{-1} 2978 (C-H sp^3), 2244 (C≡N). HRMS (ESI) theoretical $[C_8H_4Br_2{}^{79}N_4+H]^+$:

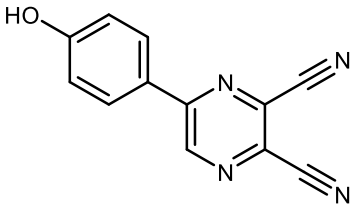
314.8875, measured: 314.8872. HRMS (ESI) theoretical $[C_8H_4Br_2{}^{81}N_4+H]+$: 316.8875, measured: 316.8851.

5,6-Bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile (**30**)



A solution of 5,6-bis(bromomethyl)pyrazine-2,3-dicarbonitrile (150 mg, 0.47 mmol) in acetone (5 mL) was added to a solution of *N*-methyl propargylamine (87.0 µL, 1.04 mmol) and pyridine (83.8 µL, 1.04 mmol) in acetone (5 ml) at RT. The reaction was stirred for 45 min after which the reaction was concentrated *in vacuo* and diluted with CH_2Cl_2 (8 mL). The organic layer was then washed with aq. $CuSO_4$ (3 x 5 mL), water (5 mL) and dried over $MgSO_4$. The solution was then concentrated *in vacuo* and purified using flash column chromatography (Flashpure, 4 g, eluent 0-35% EtOAc:cyclohexane) to obtain the title product as a yellow oil (21 mg, 0.072 mmol, 16%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 3.40 (d, 4H, *J* = 2.2 Hz, 2 x NC***H₂***), 2.34 (s, 6H, 2 x C***H₃***), 2.30 (t, 2H, *J* = 2.4 Hz, 2 x C≡C***H***); ^{13}C NMR (151 MHz, $CDCl_3$) δ 158.4, 131.1, 113.1 (***CN***), 77.8, 74.4, 58.4, 46.4, 42.4. IR ν_{max} / cm^{-1} 3288 (C-H sp^2), 2950 (C-H sp^3), 2231 (C≡N), 2107 (C≡C); HRMS (ESI) theoretical $[C_{16}H_{16}N_6+H]^+$: 293.1509, measured: 293.1509.

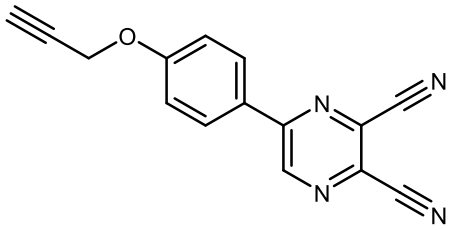
5-(4-hydroxyphenyl)pyrazine-2,3-dicarbonitrile (**31**)[252]



4'-Hydroxyacetophenone (1 g, 7.35 mmol) was dissolved in a dioxane/water mixture (9:1, v/v, 20 mL). To this solution was added selenium dioxide (2.45 g, 22.1 mmol), the resulting mixture was heated to reflux overnight. After this time the reaction was allowed to cool to RT. DAMN (1.59 g, 14.7 mmol) and conc. HCL (1.5 mL) were then added before the reaction was refluxed for a further 1.5 h. The reaction was cooled to RT and filtered; the resulting residue was washed with THF. The filtrate was then concentrated under reduced pressure to yield the crude product which was purified by flash column chromatography (Flashpure, 12 g, eluent 0-10% CH_2Cl_2:MeOH) yielding the title product as a yellow solid (522 mg, 2.35 mmol, 32%); m.p. 158-160 °C. ^{1}H NMR (600 MHz, $CDCl_3$) δ 9.21 (1H, s. Ar***H***), 8.08 (2H, d, J = 8.5 Hz, Ar***H***),

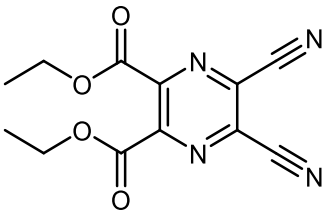
7.03 (2H, d, J = 8.5 Hz, Ar***H***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 160.0, 154.3, 143.4, 139.6, 133.3, 130.3, 129.7, 125.3, 116.9, 113.4, 113.2, 111.5. IR ν_{max} / cm^{-1} 3460 (C-H sp^2), 2236 (C≡N). HRMS (ESI) theoretical $[C_{12}H_6N_4O\text{-}H]^-$: 221.0458 measured: 221.0465.

5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile (**32**)



5-(4-hydroxyphenyl)pyrazine-2,3-dicarbonitrile (145 mg, 0.65 mmol) was dissolved in DMF (2 mL), to this solution was added potassium carbonate (135 mg, 0.975 mmol). Propargyl bromide (~80% in toluene, 111.5 uL, 0.65 mmol) was added to the mixture and the reaction was left to stir at RT overnight. The reaction was diluted with toluene (5 mL) and the solvent was removed under reduced pressure. The resulting crude residue was purified by flash column chromatography with CH/EtOAc (0-60%) yielding the title product as a clear oil (98.4 mg, 0.38 mmol, 58%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.22 (1H, s, Ar***H***), 8.13 (2H, d, J = 8.9 Hz, Ar***H***), 7.17 (2H, d, J = 8.9 Hz, Ar***H***), 4.81 (2H, d, J = 2.4 Hz, C***H₂***), 2.59 (1H, t, J = 2.4 Hz, C≡C***H***). ^{13}C NMR (126 MHz, $CDCl_3$) δ 161.6, 154.4, 143.5, 133.4, 130.0, 129.9, 125.9, 116.3, 113.5, 113.2, 77.6, 76.7, 56.2, 27.1. IR ν_{max} / cm^{-1} 3450 (C-H sp^2), 2220 (C≡N/C≡C). HRMS (ESI) theoretical $[C_{15}H_8N_4O+H]^+$: 261.0771 measured: 261.0771.

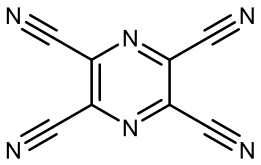
Diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate (**34**)[233]



A mixture of diethyl tartrate (1.03 g, 5.0 mmol) and NBS (3.56 g, 20 mmol) in dry $C_2H_4Cl_2$ (7 mL) was heated under reflux for 12 h under N_2 atmosphere. After completion, Na_2SO_3 was added to quench the residual bromine. The reaction mixture was evaporated under reduced pressure to remove $C_2H_4Cl_2$. Then the residue was diluted with diethyl ether (20 mL), filtered, and the solid washed with ether (10 mL). The combined filtrate was evaporated under reduced pressure to yield a yellow oil - crude diethyl 2,3-dioxosuccinate (677 mg, approximately 3.35 mmol). To a solution of crude diethyl 2,3-dioxosuccinate in AcOH (3.35 mL) and EtOH (80 mL) was added DAMN (362 mg, 3.35 mmol). The reaction was stirred at 80 °C for 2 h after which

the reaction was concentrated under reduced pressure. The crude product was purified with flash column chromatography (Flashpure, 12 g, eluent 100% CH_2Cl_2) to yield the title product as a white solid (504.9 mg, 1.84 mmol, 55%); m.p. 81-85 °C. ^{1}H NMR (600 MHz, $CDCl_3$) δ 4.48 (4H, q, *J* = 7.2 Hz, 2 x C***H₂***), 1.41 (6H, t, *J* = 7.2 Hz, 2 x C***H₃***). ^{13}C NMR (151 MHz, $CDCl_3$) δ 161.6, 146.5, 133.5, 112.0 (***CN***), 64.3, 27.0. IR ν_{max} / cm^{-1} 2940 (C-H sp^3), 2247 (C≡N), 1724 (C=O). HRMS (ESI) theoretical $[C_{12}H_{10}N_4O_4+H]^+$: 275.0775 Measured : 275.0735.

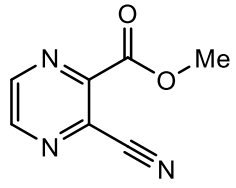
Pyrazine-2,3,5,6-tetracarbonitrile (**35**)[236]



To a solution of DAMN (500 mg, 4.6 mmol) in TFA (6 mL) was added DDQ (525 mg, 2.3 mmol) in a portion wise manner over 10 min. The beige mixture was stirred at RT for 4 h after which the reaction was warmed to 70 °C and filtered, the residue was washed with warm TFA and water. The beige residue boiled in benzene for 10 min. The solution was filtered while hot and concentrated under reduced pressure. The resulting solid was washed with water and chloroform before being dried under vacuum to yield the title compound as an off-white solid (66%, 540.4 mg, 3.04 mmol) m.p. 265-270 °C. ^{13}C NMR (176 MHz, Acetone) δ 135.79, 112.80 (**CN**). IR ν_{max} / cm^{-1} 2250 (C≡N).

Compound failed to ionise therefore HRMS was not recorded.

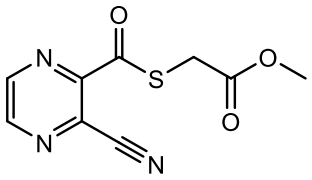
Methyl 3-cyanopyrazine-2-carboxylate (**36**)[237]



To a stirred suspension of pyrazine-2,3-dicarboxylic acid monoamide (200 mg, 1.20 mmol) in CH_2Cl_2 (2 mL) at 0 °C were added TEA (0.35 mL, 2.51 mmol) and methyl chloroformate (2.03 mL, 2.63 mmol). The mixture was allowed to reach room temperature and stirred for 2 hrs. Water and $NaHCO_3$ (sat.) 1:1 were added (2 mL), and the mixture was extracted with CH_2Cl_2 (3 x 2 mL). The organic layer was washed with H_2O, dried over $MgSO_4$, and concentrated to give the methyl 3-cyanopyrazine-2-carboxylate. Crude product was purified using flash column chromatography (Flashpure, 12 g, eluent 0-50% EtOAc:cyclohexane) to obtain the

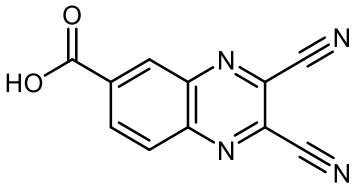
title product (157.4 mg, 0.96 mmol, 80%). ^{1}H NMR (600 MHz, CD_3CN) 8.92 (1H, d, J=2.3 Hz, Ar***H***), 8.90 (1H, d, J=2.3 Hz, Ar***H***), 4.02 (3H, s, OC***H$_3$***); ^{13}C NMR (151 MHz, CD_3CN) 163.7, 148.5, 147.8, 147.0, 131.4, 116.2 (***CN***), 54.1. IR v_{max} / cm^{-1} 3033 (C-H sp^2), 2956 (C-H sp^3), 2060 (C≡N), 1731 (C=O). HRMS (ESI) theoretical $[C_7H_5N_3O_2+H]^+$: 164.0455 Measured : 164.0455.

Methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate (**38**)



To a solution of pyrazine-2,3-dicarboxylic acid monoamide (96.1 mg, 0.58 mmol) in THF (10 mL) was added DMAP (7.1 mg, 0.058 mmol) followed by EDC.HCl (220 mg, 1.15 mmol) and methyl thioglycolate (0.155 mL, 1.73 mmol). The reaction was stirred for 16 h at RT after which the mixture was concentrated *in vacuo*. To a solution of crude methyl 2-((3-carbamoylpyrazine-2-carbonyl)thio)acetate (31 mg, 0.12 mmol) in CH_2Cl_2 (5 mL) at 0 °C were added NEt_3 (35 µL, 0.25 mmol) and methyl chloroformate (20 µL, 0.263 mmol), and the mixture was allowed to reach room temperature and stirred for 2 hrs. Water and $NaHCO_3$ (sat.) (volume ratio = 1:1) were added (3 mL), and the mixture was extracted with CH_2Cl_2 (3 x 3 mL). The organic layer was washed with H_2O and dried over $MgSO_4$. The crude product was purified using flash column chromatography (Flashpure, 4 g, eluent 0-65% EtOAc:cyclohexane) to obtain the title product as a yellow oil (13.6 mg, 0.057 mmol, 48%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.95 (1H, d, *J* = 2.3 Hz, Ar***H***), 8.88 (1H, d, *J* = 2.3 Hz, Ar***H***), 3.92 (2H, s, C***H$_2$***), 3.78 (3H, s, C***H$_3$***); ^{13}C NMR (151 MHz, $CDCl_3$) δ 189.0, 168.6, 148.5, 147.4, 145.8, 126.7, 114.4 (***CN***), 53.2, 31.6. IR v_{max} / cm^{-1} 3015 (C-H sp^2), 2955 (C-H sp^3), 2233 (C≡N), 1737 (C=O),1675 (COS); HRMS (ESI) theoretical $[C_9H_7N_3O_3S+H]^+$: 238.0281, measured : 238.0279.

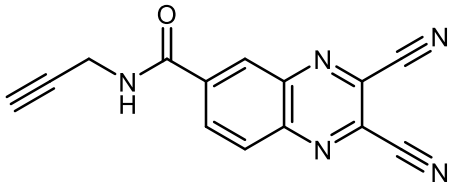
2,3-Dicyanoquinoxaline-6-carboxylic acid (**47**)[253]



DAMN (540 mg, 5 mmol) was dissolved in MeCN (25 mL), to this solution was added DDQ (1.14 g, 5 mmol). The reaction was stirred for 20 min at RT after which the mixture was filtered and concentrated *in vacuo* to yield DISN (the crude intermediate was carried forward). DISN

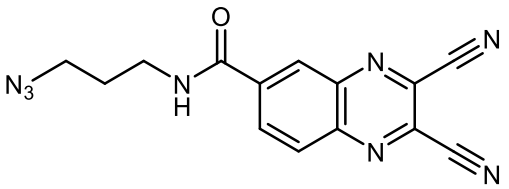
(531 mg, 5 mmol) was dissolved in TFA (8 mL), to this mixture was added 1,2-diaminobenzoic acid (767mg, approximately 5 mmol) portion wise. The reaction was then stirred overnight at RT under N_2 atmosphere. The solvent was evaporated *in vacuo*, the residue was then purified *via* flash column chromatography (Flashpure, 12 g, eluent 0-30% CH_2Cl_2:EtOAc and 1% AcOH) to yield the title product as a yellow solid (538 mg, 2.4 mmol, 48%); m.p. 217-220 °C; ^{1}H NMR (700 MHz, MeOD) δ 8.85 (1H, s, Ar***H***), 8.64 (1H, d, *J* = 8.8 Hz, Ar***H***), 8.35 (1H, d, *J* = 8.8 Hz, Ar***H***). ^{13}C NMR (176 MHz, MeOD) δ 167.0, 144.3, 142.3, 137.6, 135.1, 133.8, 133.2, 132.7, 131.21, 114.78 (***CN***), 114.75 (***CN***). IR ν_{max} / cm^{-1} 3067 (C-H sp^2), 2940 (C-H sp^3), 2253 (C≡N),1692 (C=O). HRMS (ESI) theoretical $[C_{11}H_4N_4O_2\text{-}H]^-$: 223.0262, measured: 223.0258.

2,3-Dicyano-*N*-(prop-2-yn-1-yl)quinoxaline-6-carboxamide (**49**)



2,3-Dicyanoquinoxaline-6-carboxylic acid (110 mg, 0.49 mmol) was dissolved in a mixture of anhydrous THF (2.5 mL) and toluene (12.5 mL). To this solution, thionyl chloride (370 µL, 0.45 mmol) was added and mixture was refluxed for 3 h at 100 °C. After this time the excess thionyl chloride and solvents were removed under reduced pressure and the resulting residue was dissolved in anhydrous THF (2 mL), propargylamine (311 µL, 4.9 mmol) in anhydrous THF (0.5 mL) was then slowly dropped into the mixture which was then refluxed for a further 3 h at 60 °C. The crude mixture was concentrated *in vacuo* and extracted from water (5 mL) with EtOAc (3 x 5 mL). The organic layer was then washed with brine, dried over $MgSO_4$ and purified by flash column chromatography (Flashpure, 4 g, eluent 0-50% EtOAc:cyclohexane) to yield **S4** as an off-white gum (57.5 mg, 0.22 mmol, 45%). ^{1}H NMR (700 MHz, MeOD) δ 8.68 (1H, s, Ar***H***), 8.51 (1H, d, *J* = 8.8 Hz, Ar***H***), 8.35 (1H, d, *J* = 8.8 Hz, Ar***H***), 4.24 (2H, d, *J* = 2.5 Hz, C***H$_2$***), 2.66 (1H, t, *J* = 2.5 Hz, C≡C***H***). ^{13}C NMR (176 MHz, MeOD) δ 166.9, 143.7, 142.4, 140.7, 134.0, 133.6, 133.3, 131.3, 129.9, 114.81 (**CN**), 114.79 (**CN**), 80.1, 72.4, 30.2. IR ν_{max} / cm^{-1} 3286 (N-H), 3015 (C-H sp^3), 2955 (C-H sp^2), 2242 (C≡N), 2121 (C≡C), 1704 (C=O). HRMS (ESI) theoretical $[C_{14}H_7N_5O\text{-}H]^-$: 260.0572, measured : 260.0578.

N-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide (**48**)[254]

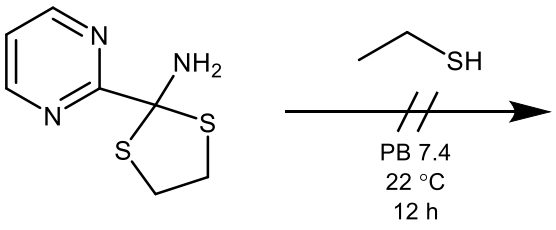


NaN_3 (450 mg, 6.9 mmol) in water (2 mL) was added to water solution (3 mL) of 3-bromopropylamine hydrobromide (500 mg, 2.3 mmol) and heated at 90 °C for 24 h. The solution was then alkalized with potassium hydroxide (approximately pH 10) and extracted with diethyl ether (3 x 5 mL), dried over anhydrous $MgSO_4$ and concentrated *in vacuo* to yield a yellow oil (174 mg, 55%). The product was used directly in the next step due to its instability.

2,3-Dicyanoquinoxaline-6-carboxylic acid (100 mg, 0.44 mmol) was dissolved in anhydrous THF (10 mL), thionyl chloride (325 µL, 0.45 mmol) in anhydrous toluene (7 mL) was added to this solution and mixture was refluxed for 3 h at 100 °C. After this time, the excess thionyl chloride and solvents were removed *in vacuo* and the resulting residue was dissolved in anhydrous THF (2 mL). 3-Azidopropan-1-amine (120 mg, 1.2 mmol) in anhydrous THF (0.5 mL) was then slowly dropped into the mixture which was then refluxed for a further 3 h at 60 °C. The crude mixture was concentrated *in vacuo* and extracted from water (5 mL) with EtOAc (3 x 5 mL). The organic layer was then washed with brine, dried over $MgSO_4$ and purified by flash column chromatography (Flashpure, 4 g, eluent 0-20% CH_2Cl_2:EtOAc) to yield the title product as a white solid (78 mg, 0.25 mmol, 58%). m.p.139-142 °C . ^{1}H NMR (700 MHz, $CDCl_3$) δ 8.55 (1H, s, Ar***H***), 8.52 (1H, d, *J* = 8.8 Hz, Ar***H***), 8.35 (1H, d, *J* = 8.8 Hz, Ar***H***), 3.67 (2H, q, *J* = 6.4 Hz, C***H₂***), 3.54 (2H, t, *J* = 6.4 Hz, C***H₂***), 1.99 (2H, p, *J* = 6.4 Hz, C***H₂***). ^{13}C NMR (176 MHz, $CDCl_3$) δ 164.86, 142.70, 141.40, 140.54, 133.83, 131.82, 131.57, 130.80, 128.01, 113.23 (***CN***), 50.03, 38.87, 28.62. IR ν_{max} / cm^{-1} 3083 (C-H sp^2), 2955 (C-H sp^3), 2110 (C≡N), 1737 (C=O). HRMS (ESI) theoretical $[C_{14}H_{10}N_8O\text{-}H]^-$: 305.0905 Measured : 305.0905.

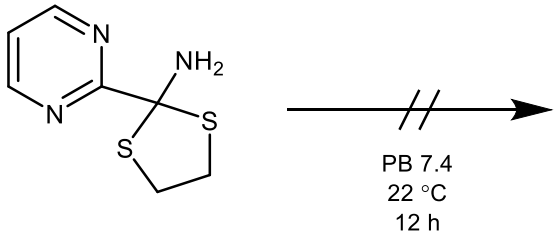
ADTA (**3**) thiol stability studies

ADTA (**3**) incubation with ethanethiol



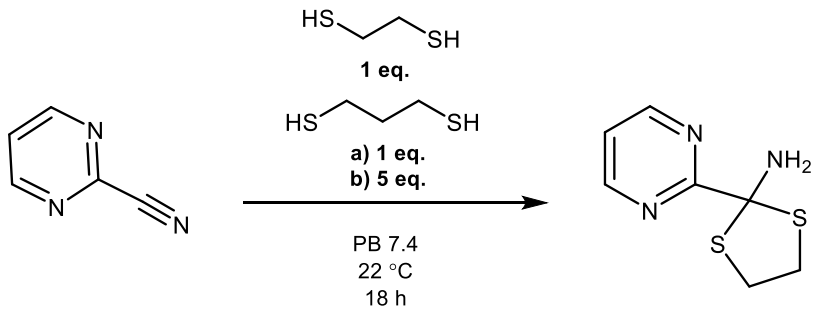
2-(Pyrimidin-2-yl)-1,3-dithiolan-2-amine was dissolved in MeCN (10%), to this solution was added 250 mM phosphate buffer pH 7.4 (10 mM final concentration) followed by ethanethiol (9 µL, 0.12 mmol). The mixture was tracked *via* NMR. Over 12 h, no changes were seen to the starting material.

ADTA (**3**) aqueous stability



2-(Pyrimidin-2-yl)-1,3-dithiolan-2-amine was dissolved in 250 mM phosphate buffer pH 7.4 (10 mM), The mixture was tracked *via* NMR. Over 12 h, no changes were seen to the starting material.

Ethanedithiol Vs. Propanedithiol NBT competition experiment



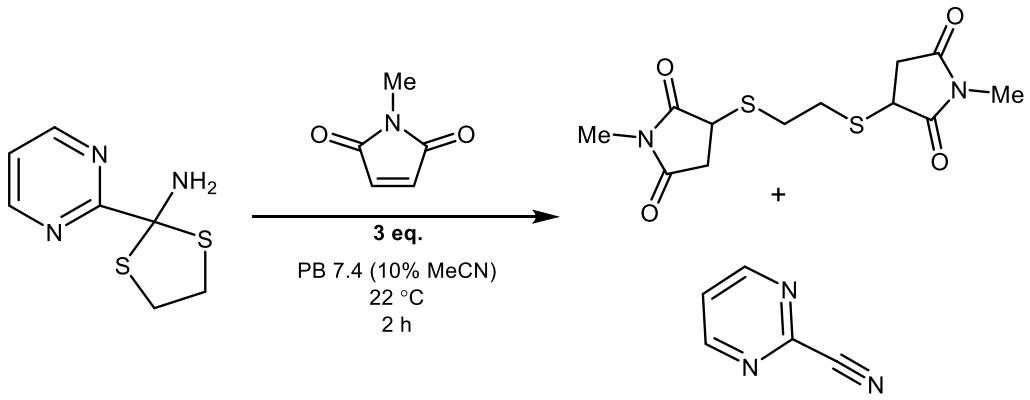
Ethanedithiol (49.6 µL, 0.59 mmol) and propanedithiol ((a) 59.1 µL, 0.59 mmol, 1 equiv. / (b) 295 µL, 2.95 mmol, 5 equiv.) were dissolved in MeCN (5 mL) followed by addition of 50 mM phosphate buffer pH 7.4 (13 mL). To this mixture was added 2-cyano pyrimidine (62 mg, 0.59 mmol) dissolved in MeCN (2 mL). The reaction was stirred for 18 h at RT before being extracted with EtOAc (3 x 10 mL). The organic layer was dried over $MgSO_4$ and purified by flash column chromatography (Flashpure, 4 g, eluent 0-10% MeOH : CH_2Cl_2) to afford 2-(pyrimidin-2-yl)-1,3-dithiolan-2-amine as a yellow-brown oil ((a) 88.0 mg ,0.44 mmol, 75% (b) 84.5 mg, 0.42 mmol, 72%). ^{1}H NMR (700 MHz, $CDCl_3$) δ 8.71 (2H, d , *J* = 4.8 Hz), 7.15 (1H, t, *J* = 4.8 Hz), 3.61 (4H, s).

*NMR as before (**3**)*

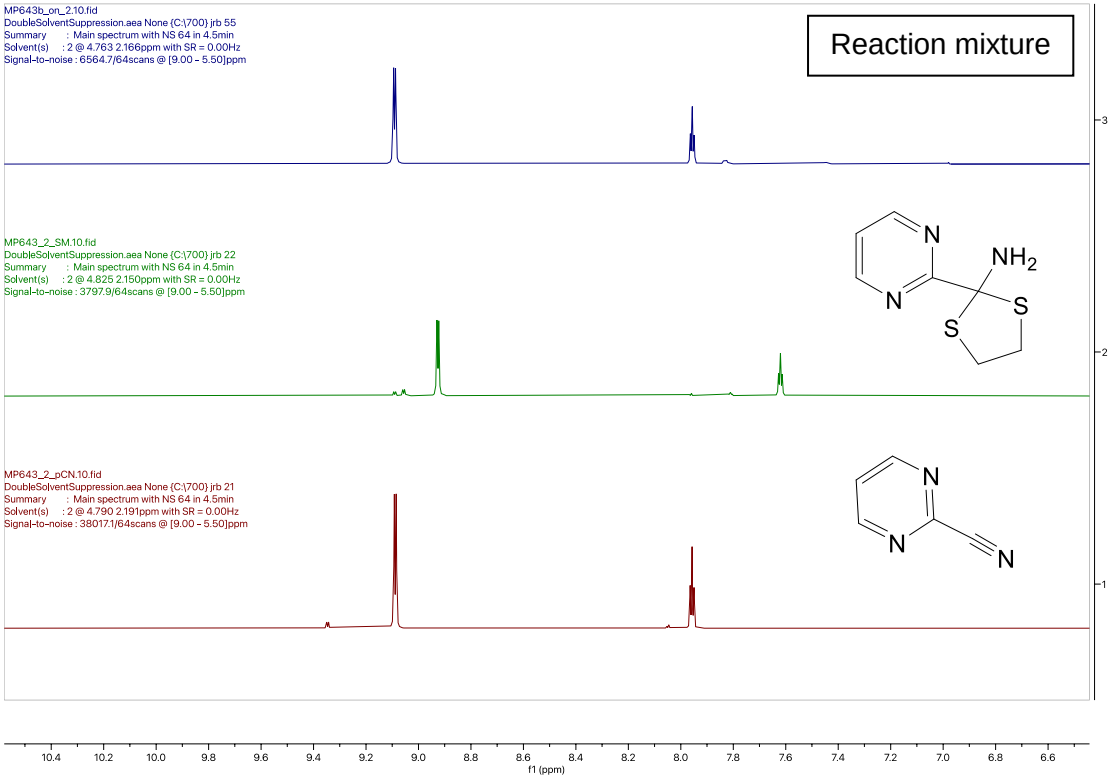
ADTA (**3**) reversibility- *N*-methyl maleimide

A 10 mM solution of 2-(pyrimidin-2-yl)-1,3-dithiolan-2-amine was prepared in PB (pH 7.4)/ 20% MeCN (0.4 mL). To this solution was added *N*-methyl maleimide (3 equiv.) in 0.1 mL MeCN. The reaction was analysed *via* a double solvent suppression ^{1}H NMR method and ^{13}C NMR.

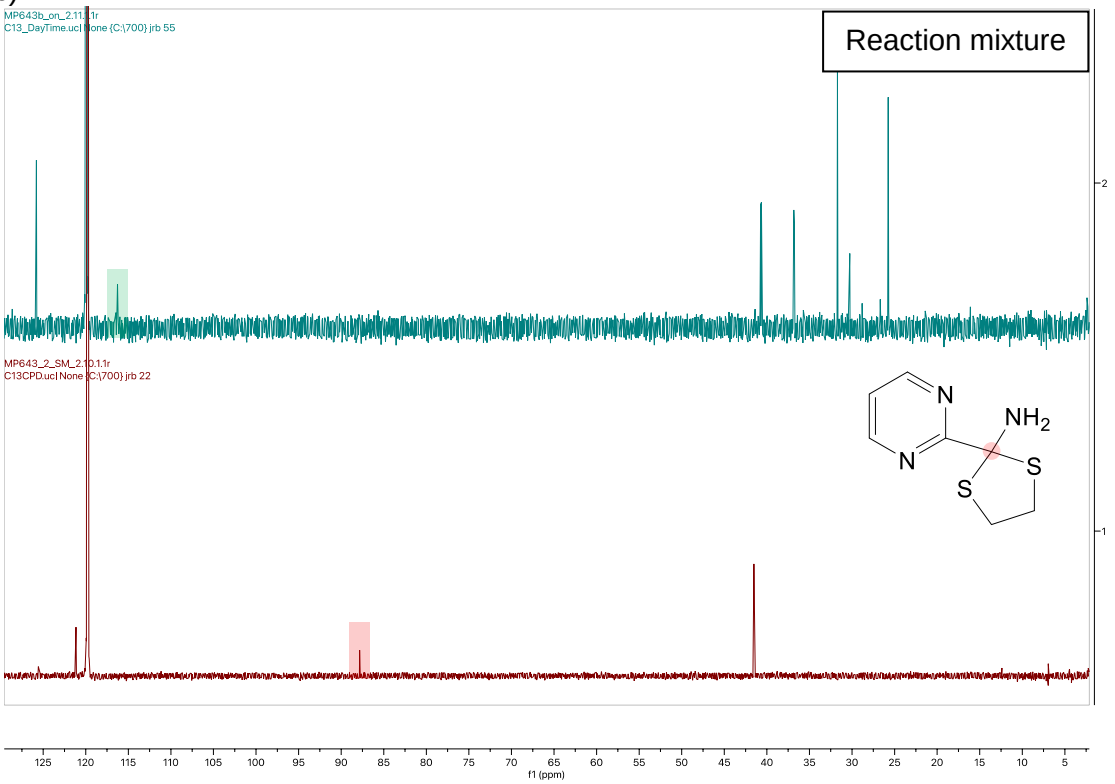
(a)



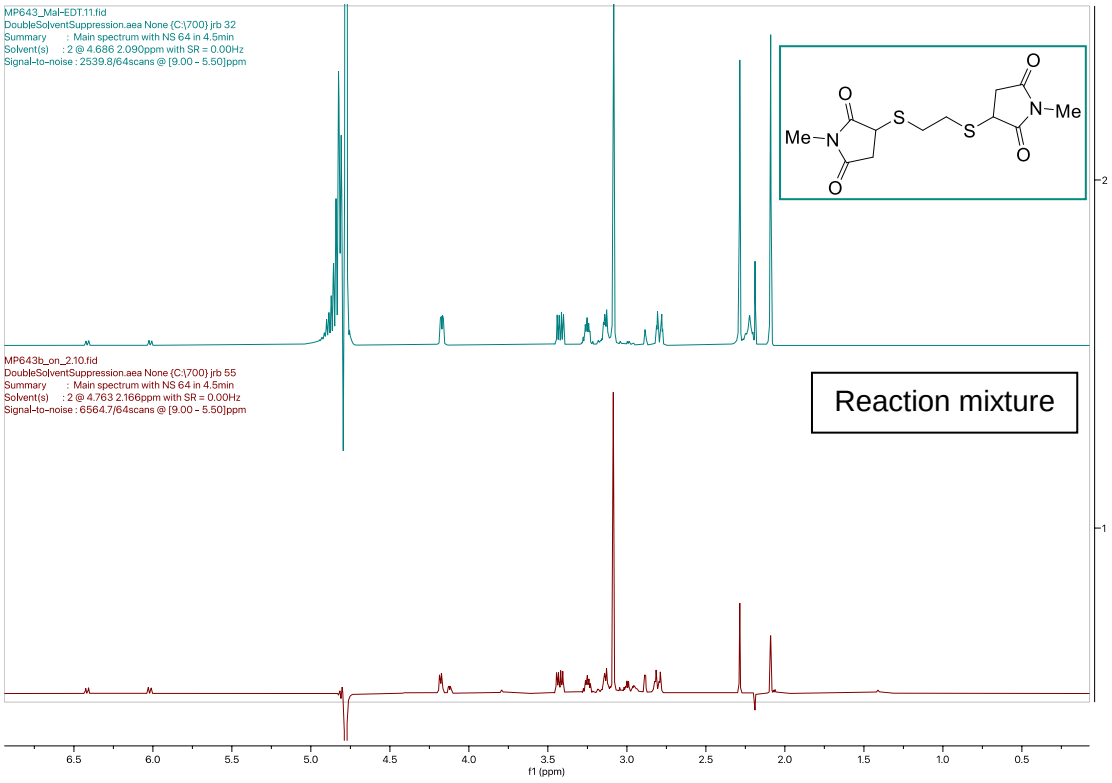
(b)



(c)


Reaction mixture
MP643_2_SM_2.10.1.1r
C13CPD.ucl None {C:\700} jrb 22
N
NH2
N
S
S
f1 (ppm)

(d)


MP643_Mal-EDT.11.fid
DoubleSolventSuppression.aea None {C:\700} jrb 32
Summary : Main spectrum with NS 64 in 4.5min
Solvent(s) : 2 @ 4.686 2.090ppm with SR = 0.00Hz
Signal-to-noise : 2539.8/64scans @ [9.00 – 5.50]ppm
O
Me–N
O
S
S
O
N–Me
O
MP643b_on_2.10.fid
DoubleSolventSuppression.aea None {C:\700} jrb 55
Summary : Main spectrum with NS 64 in 4.5min
Solvent(s) : 2 @ 4.763 2.166ppm with SR = 0.00Hz
Signal-to-noise : 6564.7/64scans @ [9.00 – 5.50]ppm
Reaction mixture
f1 (ppm)

(e)

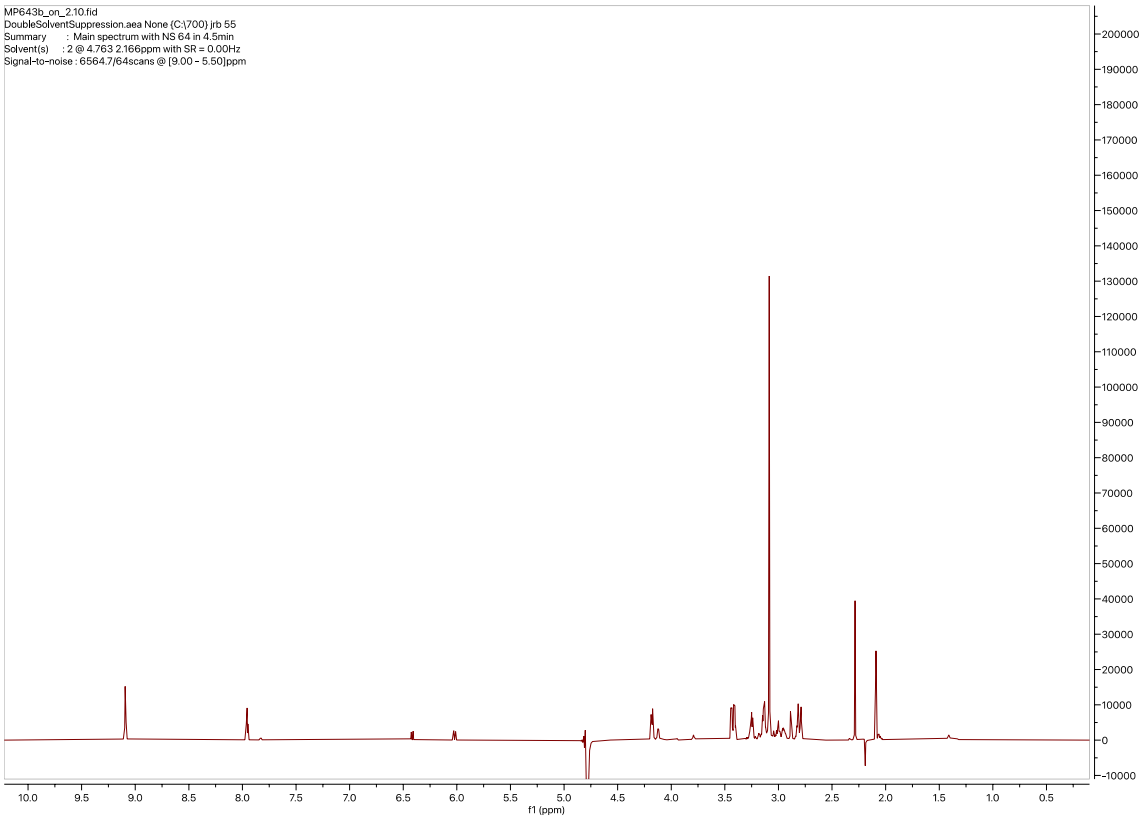
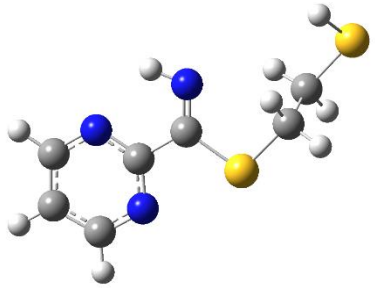


Figure S4; ADTA **3** removal with *N*-methyl maleimide; (a) reaction scheme for maleimide mediated removal of ADTA (b) Water suppression ^{1}H NMR zoomed into the aromatic region showing reformation of the 2-cyano pyrimidine nitrile aromatic proton peaks (9.09 ppm, d and 7.96 ppm, t) (c) ^{13}C NMR showing the loss of the quaternary carbon (87.83 ppm) and reformation of the characteristic nitrile carbon peak (116.27 ppm) (d) Stacked comparison between the EDT maleimide product (green), formed *via* the reaction of EDT (1 equiv.) and *N*-methyl maleimide (3 equiv.) under identical conditions, and the reaction mixture after 2 h. (e) Full reaction NMR spectra.

DFT general remarks and calculated co-ordinates (S5)

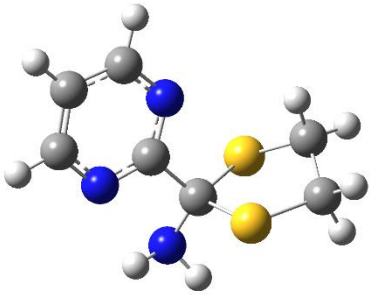
Calculations were performed with Gaussian 16.[6] The geometry optimization of all the minima involved was carried out at the M062X level of theory with the 6-31+G(d,p) basis set. Structures were optimised in the solvent (water) phase and confirmed as minima by the existence of zero imaginary vibrational frequencies.

M062X/6-31+G(d,p) – 5-membered Acyclic optimised geometry
-1232.448505 Hartree / -3235793.55 kJ mol^{-1}



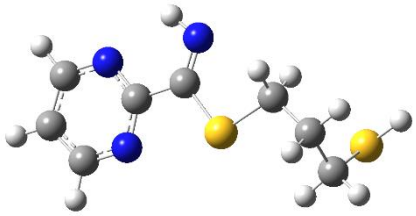
Symbol	X	Y	Z
C	1.771327	0.178937	-0.08098
N	2.28153	-1.03277	0.126555
C	3.601208	-1.09316	0.341227
C	4.398759	0.042692	0.347921
C	3.757785	1.255064	0.120776
N	2.441791	1.331469	-0.09499
H	4.022705	-2.07928	0.511206
H	5.466156	-0.01194	0.522686
H	4.309502	2.190149	0.110567
C	0.288687	0.27719	-0.33145
C	-2.26214	-0.70607	-0.59744
H	-2.24016	0.025523	-1.40839
H	-2.8082	-1.58993	-0.9364
C	-2.89765	-0.12825	0.658073
H	-3.00743	-0.90128	1.422185
H	-2.29312	0.681239	1.069856
H	0.374952	2.136443	-0.53732
N	-0.31271	1.377298	-0.55635
S	-0.56073	-1.27628	-0.3011
H	-4.23344	1.478076	-0.49315
S	-4.59008	0.487565	0.342571

M062X/6-31+G(d,p) – 5-membered cyclic optimised geometry
-1232.457165 Hartree / -3235816.287 kJ mol^{-1}



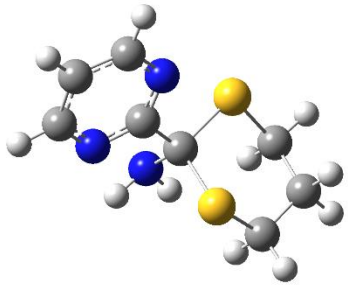
Symbol	X	Y	Z
C	-0.94853	0.154282	0.183062
N	-1.15342	-1.16095	0.093175
C	-2.40909	-1.55447	-0.1306
C	-3.45951	-0.65423	-0.2725
C	-3.13205	0.688836	-0.17538
N	-1.87835	1.101244	0.053315
H	-2.57714	-2.6256	-0.20051
H	-4.47481	-0.98464	-0.45346
H	-3.88601	1.463417	-0.28201
C	0.483017	0.651463	0.380146
C	2.270096	-1.38229	-0.01664
C	2.658122	-0.35407	-1.06601
H	3.151641	-1.90446	0.365271
H	1.560779	-2.11042	-0.41505
H	3.501384	0.251117	-0.72723
H	2.915568	-0.83531	-2.01293
N	0.467352	1.907342	1.075389
H	1.386039	2.339228	1.033206
H	-0.21515	2.530801	0.652163
S	1.532412	-0.47905	1.369836
S	1.20611	0.706877	-1.36516

M062X/6-31+G(d,p) – 6-membered Acyclic optimised geometry
-1271.718972 Hartree / -3338898.161 kJ mol^{-1}



Symbol	X	Y	Z
C	2.121481	0.20796	0.041964
N	2.204696	-1.06882	-0.32541
C	3.414586	-1.63498	-0.23782
C	4.52507	-0.93351	0.210331
C	4.3205	0.39365	0.568632
N	3.1191	0.972574	0.487197
H	3.488615	-2.67633	-0.53583
H	5.501664	-1.39652	0.278351
H	5.137111	1.011021	0.930459
C	0.772891	0.874798	-0.05011
C	-1.9016	1.01983	-0.63867
H	-2.05007	1.345877	0.393343
H	-1.6318	1.889578	-1.24135
C	-3.15946	0.362206	-1.19555
H	-3.94901	1.122289	-1.19999
H	-2.99662	0.069321	-2.23983
C	-3.65266	-0.86776	-0.43807
H	-2.90724	-1.66811	-0.44205
H	-4.55207	-1.26111	-0.91503
H	-4.99293	0.323441	1.140924
S	-0.51213	-0.15723	-0.67811
S	-4.01096	-0.57435	1.332675
N	0.565782	2.08595	0.291314
H	1.449839	2.484642	0.62206

M062X/6-31+G(d,p) – 6-membered Cyclic optimised geometry
-1271.725038 Hartree / -3338914.087 kJ mol^{-1}



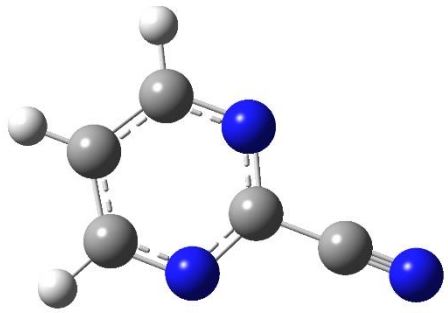
Symbol	X	Y	Z
C	1.260163	0.030843	0.164636
N	1.689469	1.036411	-0.59495
C	3.004879	1.079554	-0.84127
C	3.883825	0.128981	-0.34336
C	3.326383	-0.87889	0.432931
N	2.016661	-0.93502	0.691827
H	3.355917	1.903193	-1.4561
H	4.946422	0.170164	-0.54846
H	3.941047	-1.66517	0.861546
C	-0.23462	-0.08725	0.442091
C	-2.77366	1.081499	0.450458
C	-2.61469	-1.27989	-0.49404
C	-3.27976	0.087293	-0.58781
H	-3.3123	2.028706	0.362139
H	-2.94412	0.710121	1.466725
H	-3.04741	-1.9695	-1.22348
H	-2.76382	-1.72017	0.498075
H	-4.35497	-0.05161	-0.42326
H	-3.14691	0.499288	-1.59382
N	-0.4262	-0.56595	1.78964
H	-1.41152	-0.72291	1.984017
H	0.052938	-1.4586	1.891393
S	-0.83176	-1.27094	-0.87975
S	-1.02511	1.555665	0.231868

Pyrimidine-CN

This was part of the attempt to see the difference between the starting materials with no attachment vs single attachment vs double attachment. As a side note, the pyrimidine was 39 kJ mol^{-1} lower in energy in the solvent vs gas phase optimisation. The energy of the single conformer (**MP-153**) was -935727.6014 kJ mol^{-1} – (shown below).

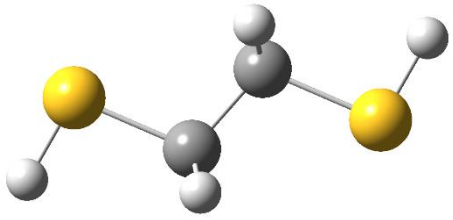
M062X/6-31+G(d,p) – Pyrimidine-CN

-356.399772 Hartree / -935727.6014 kJ mol^{-1}



Symbol	X	Y	Z
C	-0.52736	0	0.000015
N	0.050947	1.198221	0.000005
C	1.389414	1.185036	-0.00001
C	2.114655	0	-1.7E-05
C	1.389414	-1.18504	-8E-06
N	0.050947	-1.19822	0.000002
H	1.884578	2.150356	-1.5E-05
H	3.197716	0	-3.1E-05
H	1.884578	-2.15036	-2.2E-05
C	-1.98278	0	0.000012
N	-3.14002	0	0.00001

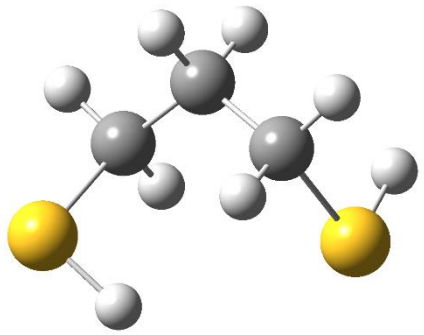
M062X/6-31+G(d,p) – Ethanedithiol



Symbol	X	Y	Z
C	0.504163	0.569163	-7E-06
H	0.383595	1.189401	-0.8906
H	0.383583	1.189415	0.890576
C	-0.50416	-0.56916	-7E-06
H	-0.38358	-1.18942	0.890576

H	-0.3836	-1.1894	-0.8906
S	2.187535	-0.1574	0.000009
H	2.856912	1.00627	-8.7E-05
S	-2.18754	0.157398	0.000009
H	-2.85691	-1.00627	-8.6E-05

M062X/6-31+G(d,p) – Propanedithiol



Symbol	X	Y	Z
C	-1.0108	0.589125	-0.77402
H	-1.72532	1.28108	-1.22575
H	-0.5232	0.042577	-1.58368
C	0.001966	1.36721	0.057809
H	-0.52705	1.995522	0.784094
H	0.540687	2.040895	-0.62067
C	1.010906	0.523795	0.829165
H	1.717478	1.171793	1.350815
H	0.515545	-0.09131	1.585806
S	-2.03812	-0.58867	0.178039
H	-1.07188	-1.49078	0.421439
S	1.9514	-0.66558	-0.19858
H	2.448858	0.237481	-1.06116

DFT Energy Difference Calculations

Addition	SETFE	kJ/mol
2-PyrCN + EDT (MP-153 + MP-154)	-1232.45	-3235787
2-PyrCN + PDT (MP-153 + MP-155)	-1271.72	-3338889

Difference	SETFE	kJ/mol
Vs 5 Closed (Cyclic) (MP-027)	-0.011323	-29.7285365
Vs 5 Open (Acyclic) (MP-030)	-0.002663	-6.9917065

Difference	SETFE	kJ/mol
Vs 6 Closed (Cyclic) (MP-064)	-0.009379	-24.6245645

Vs 6 Open (Acyclic) (MP-103)	-0.003313	-8.6982815

6.3. Bioconjugation general remarks

All conjugation experiments were performed in a standard polypropylene Eppendorf safe-lock tubes (1.5 or 2.0 mL) at atmospheric pressure with mixing at the temperature stated. All reagents and solvents were purchased from commercial sources; Sigma UK, Fisher UK or VWR UK and used as per manufacturer instructions. Buffers were prepared with double-deionised water and filter sterilised (0.20 µm). Borate Buffer Saline (BBS) contains 50 mM sodium borate, 50 mM sodium chloride, and 5 mM ethylenediaminetetraacetic acid (EDTA) at pH 8.5 or 7.4. Phosphate Buffer Saline (PBS) contains 10 mM phosphate, 138 mM NaCl, 2.7 mM KCl, pH 7.4. Conjugation Buffer (CB) contains 40 mM phosphate, 20 mM NaCl, 6 mM EDTA, pH 7.4. Ultrapure DMF was purchased from Sigma and stored under dry conditions. Ultrafiltration was carried out using Amicon Ultra-4 Centrifugal Filter Units with molecular weight cutoff (MWCO) of 10 kDa or in Vivaspin 500 centrifugal concentrators (Sartorius, UK) with molecular weight cut-off (MWCO) of 10 kDa. Centrifugation was performed using Eppendorf 5415R fixed angle bench rotor operating at 14000 rcf at 20 °C or in an Eppendorf 5810 swing-bucket rotor centrifuge operating at 3220 rcf at 20 °C. Trastuzumab (Ontruzant^{TM}) was purchased from UCLH in its clinical formulation (Samsung Bioepis, lyophilised). Ontruzant Fab was prepared by a sequential enzymatic digest of the full antibody with pepsin and papain, following a literature procedure.[255]

UV-Vis spectroscopy was used to determine protein concentration using NanoDrop One/One Microvolume UV-Vis Spectrophotometer (Thermo Fisher) operating at room temperature. Sample buffer was used as blank for baseline correction with extinction coefficients: ε280 = 68590 M^{-1} cm^{-1} for Fab, ε280 = 215380 $M^{-1}cm^{-1}$ for trastuzumab,[256] ε280 = 14180 for Thioredoxin,[257] ε505 = 74000 M^{-1} cm^{-1} for Azide-fluor 488. A correction factor at 280 nm of 0.178 for Fluorescein, 0.11 for Azide-fluor 488 was employed. Antibody conjugate concentration was determined using the same extinction coefficient as for native trastuzumab Fab (Quinoxaline nitriles were found to have negligible absorbance at 280 nm). Protein conjugation reactions were monitored by 12% glycine-SDS-PAGE with 6% stacking gel under nonreducing conditions. Samples were mixed 1:1 with SDS non-reducing loading buffer (composition for 5X SDS: 10 g (8 mL) of glycerol, 4 mL dH2O, 1.6 mL of 10 % (w/v) SDS, 1 mL of 0.5 M Tris-HCl pH 6.8, 0.025 g Coomassie Brilliant Blue R-250) and heated at 75 °C for 5 min before being loaded to the gel. Page Ruler Plus Pre-Stained Protein Ladder (Thermo Scientific) was used. Samples were run at 200 V, 60 min at 1 X running buffer. The 10 X running buffer contained 30 g of Tris base, 144 g of glycine and 10 g of SDS in 1 L of dH2O,

with final pH of 8.3. Gels were stained in Coomassie Blue Stain (10% Ammonium Sulfate (100 g), 0.1% Coomassie Brilliant Blue R-250 (500 mg), 3% Phosphoric Acid (30 mL), ethanol (200 mL), and water (1 L). Gels were de-stained with water. The SDS-PAGE gels (Figure S82 and 83) have had lanes cut out of the gel that included data not related to the specific experiments.

LCMS general remarks

For protein analysis, liquid chromatography mass spectrometry (LCMS) was performed at the Department of Chemistry, UCL, on an Agilent 1100/1200 LC system, with a 6510A QTOF mass spectrometer or a 6530A QTOF mass spectrometer. Samples for LCMS analysis were prepared in LCMS grade water using ZebaSpin (7 kDa) as per manufacturer's instructions. All Fab LCMS samples were prepared at 5 µM. For trastuzumab, LCMS samples were prepared at 6.5 µM, and the antibody was deglycosylated with 1 µL PNGase (NEB) for 16 h at 37 ºC prior to LCMS submission. Sodium adducts are commonly found as [M+23] in some of the obtained LCMS data. 2 µL of sample was injected through a 100 µL loop into the Agilent PLRPS 1000 Å column (150 mm x 2.1 mm, 8 µM particle size) at 60 ºC. Proteins were separated by gradient elution of mobile phase A (water with 0.1% formic acid) and mobile phase B (acetonitrile with 0.1% formic acid) as per the gradient(s) specified in Table S.1 (Agilent 6530) and Table S.2 (Agilent 6510). The flow rate was 0.8 mL/min (Agilent 6530) or 0.3 mL/min (Agilent 6510), and the 6530/6510 QTOF was prepared in positive mode. Other LCMS parameters include: VCap = 3500 V, gas temperature = 350 ºC, dry gas flow rate = 10 L/min, nebuliser = 30 psi, fragmentor = 380 V, skimmer = 65 V, acquisition rate = 0.5 spectra/s. Data was acquired with a LCMS scan within the m/z range 700-7000. Deconvolution of raw LCMS data to zero charge mass spectra was carried out using the maximum entropy deconvolution algorithm within the Agilent MassHunter software (v B.07.00). The number of compound additions is shown on the LCMS (0, 1, 2, 3... etc). For labelling of protein conjugates: LC refers to light chain, HC refers to heavy chain, HHLL refers to full antibody, and HL refers to half antibody. Deconvolution artefacts include species with masses corresponding to LC-LC (2 x LC) or HC-HC (2 x HC). Due to ionisation differences between smaller and larger proteins and subsequent ion detection, the HC cannot be observed in some of the obtained LCMS data.

Table S.1 LCMS gradient for protein LCMS analysis on Agilent 6530.

Time (min)	**Solvent A (water + 0.1% formic acid) (%)**	**Solvent B (acetonitrile+ 0.1% formic acid) (%)**
0.00	80	20
1.00	80	20
6.50	40	60

7.50	40	60
7.60	80	20
8.50	80	20

Table S.2 LCMS gradient for protein LCMS analysis on Agilent 6510.

Time (min)	**Solvent A (water + 0.1% formic acid) (%)**	**Solvent B (acetonitrile+ 0.1% formic acid) (%)**
0.0	85	15
2.0	85	15
3.0	68	32
4.0	68	32
14.0	65	35
18.0	5	95
20.0	5	95
22.0	85	15
25.0	85	15

General procedure for the preparation of trastuzumab (HerceptinTM) Fab fragment.[255]

Trastuzumab (0.50 mL, 44.0 μM, 6.41 mg/mL) was buffer exchanged into NaOAc buffer (20 mM NaOAc, pH 3.1) via ultrafiltration (Amicon Ultra-4 Centrifugal Filter units with 10 kDa MWCO). Immobilized pepsin (0.15 mL) was also washed with NaOAc buffer (20 mM NaOAc, pH 3.1), using a Pierce™ centrifuge column. Trastuzumab (0.50 mL) was then added to the pepsin and the mixture was incubated at 37 ºC for 5 h, under constant agitation (1,100 rpm). The resin was then removed from the digest using a Pierce™ centrifuge column and washed with digest buffer (3 × 0.4 mL, 50 mM phosphate, 150 mM NaCl, 1 mM EDTA, pH 6.8). The digest was combined with the washes and the volume was adjusted to 0.5 mL. Next, immobilized papain (0.65 mL) was activated with 10 mM DTT in digest buffer (50 mM phosphate, 150 mM NaCl, 1 mM EDTA, pH 6.8). This was incubated at 37 ºC for 1.5 h under constant agitation (1,100 rpm). The resin was washed with digest buffer (4 × 0.4 mL, 50 mM phosphate, 150 mM NaCl, 1 mM EDTA, pH 6.8) without DTT, using a PierceTM centrifuge column. $F(ab')_2$ (0.5 mL) was added to the washed papain and the mixture incubated at 37 ºC for 20 h under constant agitation (1,100 rpm). The resin was separated and washed with conjugation buffer (4 × 0.4 mL, 40 mM phosphate, 20 mM NaCl, 6 mM EDTA, pH 7.4). The digest was combined with the washes and the buffer was exchanged completely for

conjugation buffer via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 150 µM. Aliquots of Fab were stored at -20 °C for up to 6 months.

Native Fab

Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

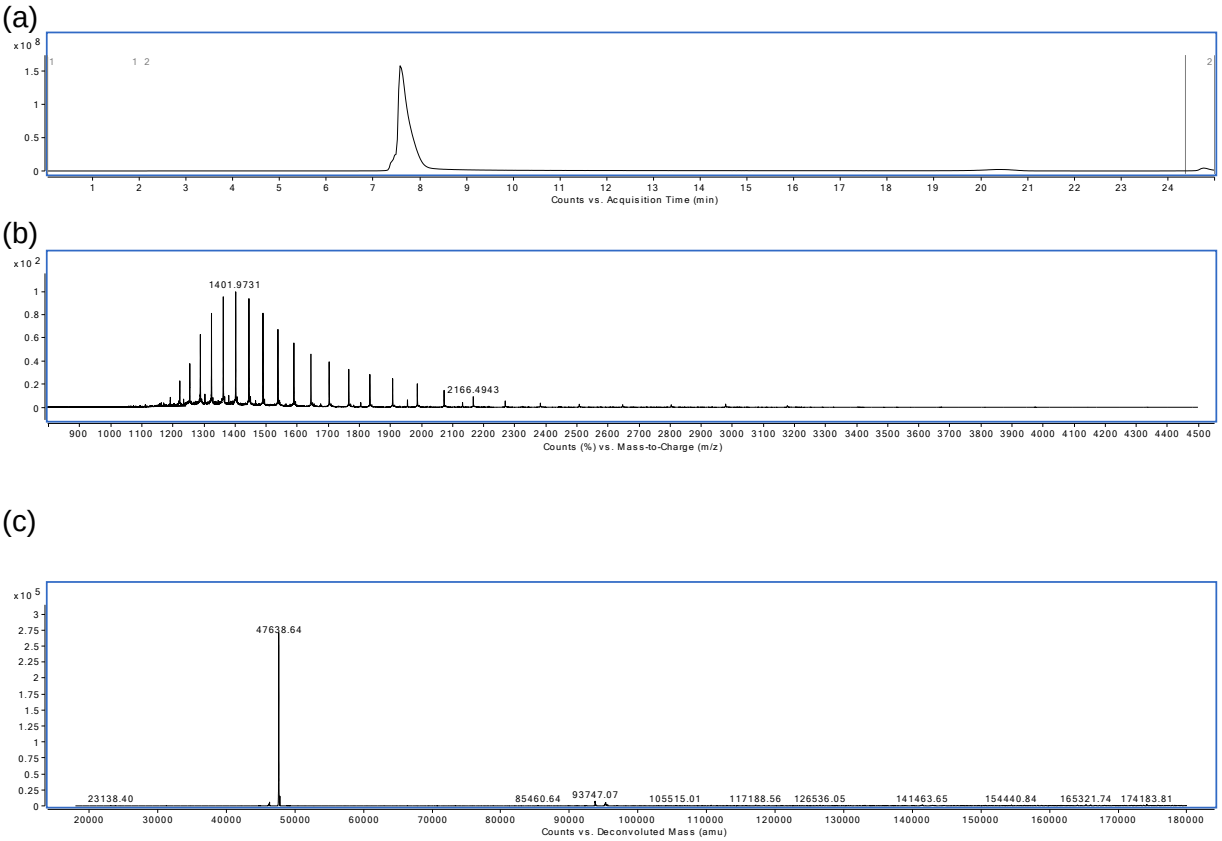
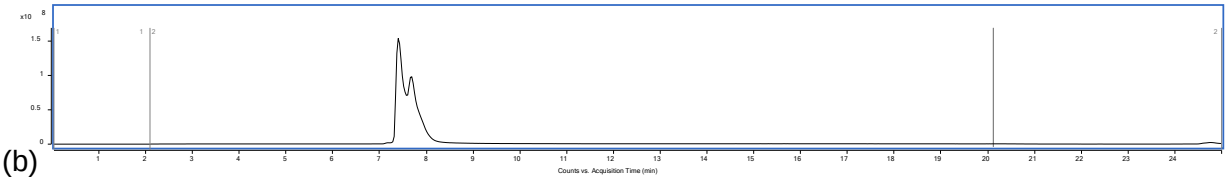


Figure S6: LCMS analysis of Native Fab; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected native Fab mass: 47638 Da. Observed: 47638 Da.

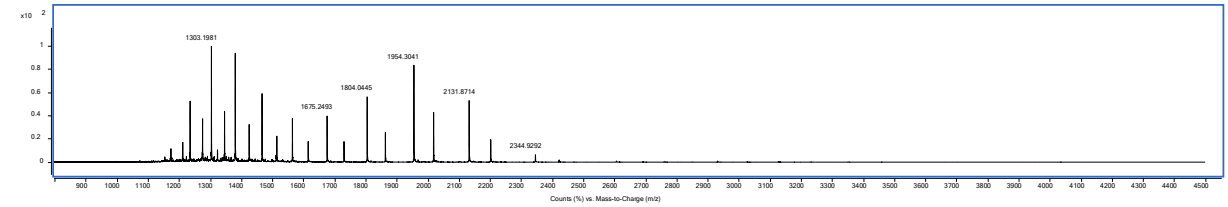
Reduced Fab

Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. The sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1} .

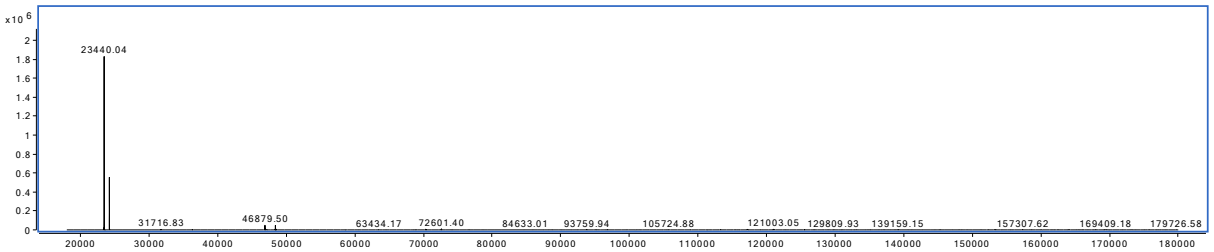
(a)



(b)



(c)



(d)

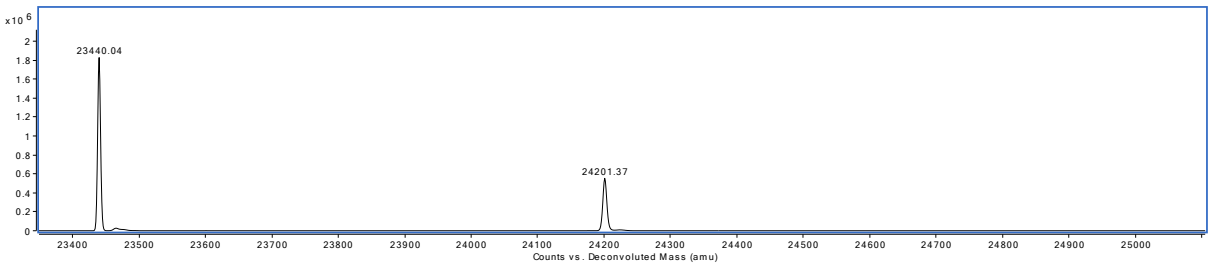
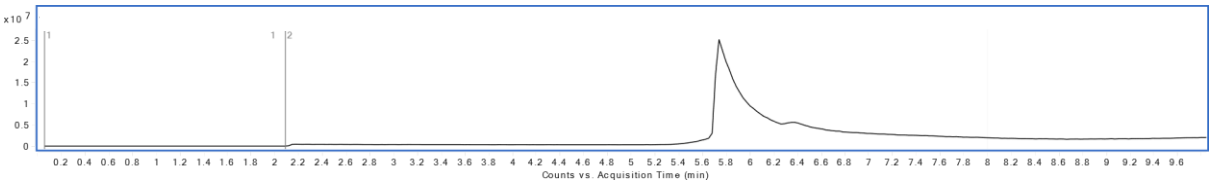


Figure S7: LCMS analysis of Reduced Fab; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected: native light chain (**LC**) 23438 Da, native heavy chain (**HC**) 24200 Da. Observed: 23440 Da, 24201 Da.

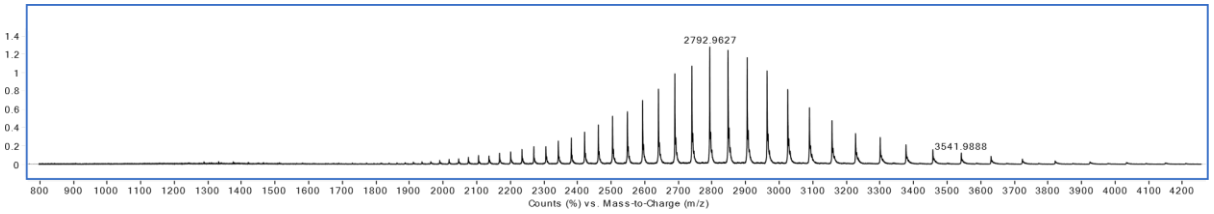
Trastuzumab

Trastuzumab full antibody (80 µL, 25 µM, 3.75 mg/mL) in conjugation buffer was buffer swapped into 50 mM ammonium acetate pH 6.9 (7 kDa MWCO, ZebaSpin) and adjusted to 5 µM with ammonium acetate (pH 6.9). To this (40 µL) was added 0.67 µL of PNGase F and the reaction was left at 37 °C for 16 h. Concentration was determined photometrically using ε_{280}=215380 M^{-1} cm^{-1}.

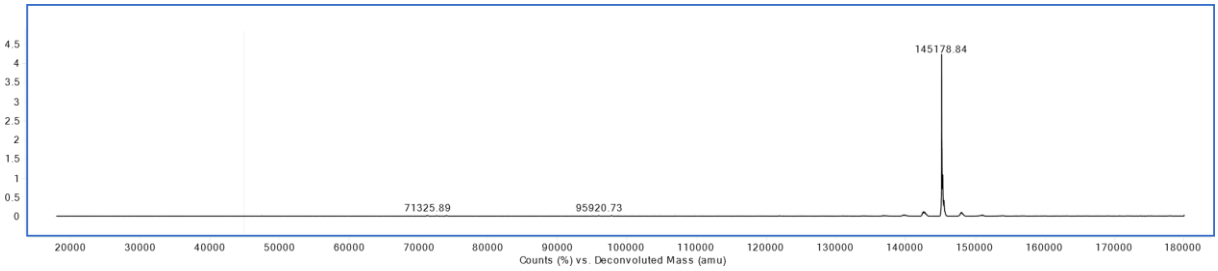
(a)



(b)



(c)



(d)

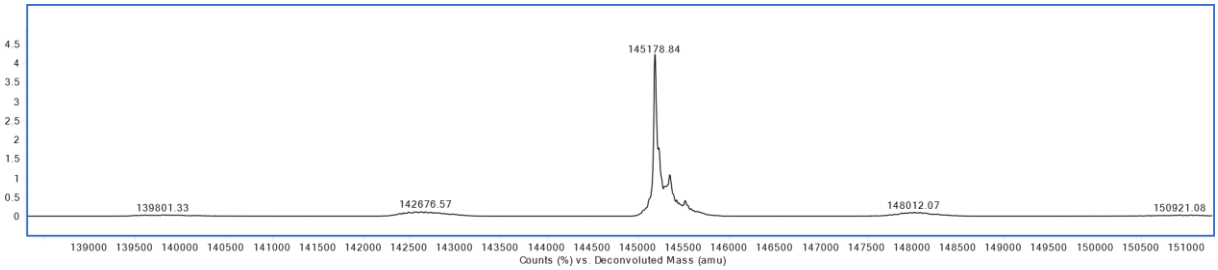


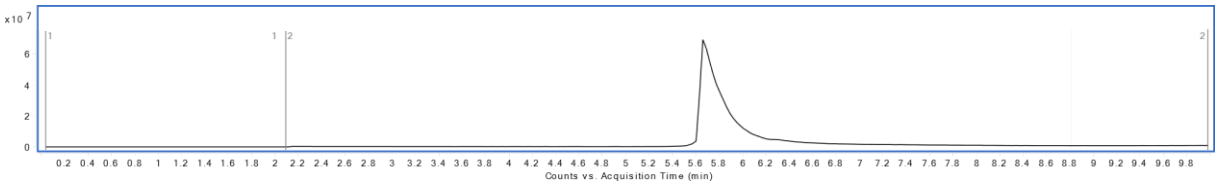
Figure S8: LCMS analysis of Trastuzumab; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; native full antibody expected 145177 Da. Observed 145178 Da.

Reduced Trastuzumab

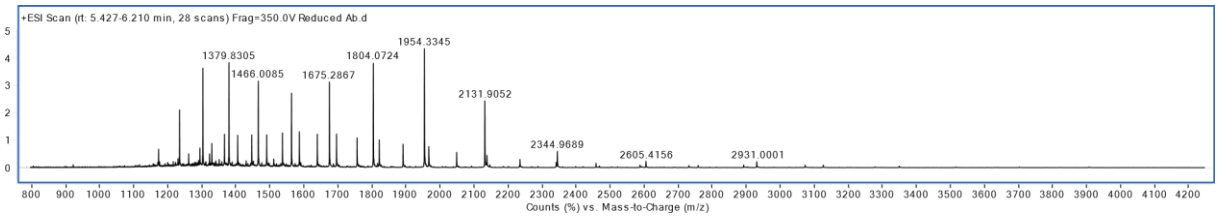
Trastuzumab full antibody (80 µL, 25 µM, 3.75 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (1.3 µL, 15 mM in dH_2O, 10 equiv.). The mixture was incubated at 37 °C for 2 h, 300 rpm. The excess reagent was removed *via* ultrafiltration (10 kDa MWCO) into 50 mM ammonium acetate pH 6.9 (7 kDa MWCO, ZebaSpin) and adjusted to 5 µM with ammonium acetate (pH 6.9). To this (40 µL) was added 0.67 µL of PNGase F

and the reaction was left at 37 °C for 16 h. Concentration was determined photometrically using ε_{280}=215380 M^{-1} cm^{-1}.

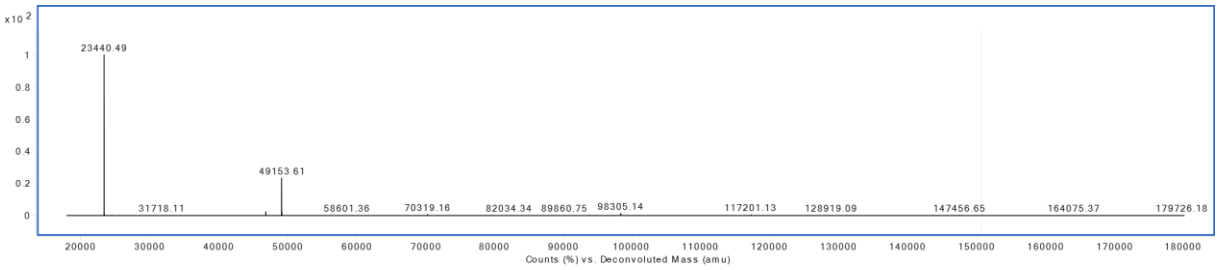
(a)



(b)



(c)



(d)

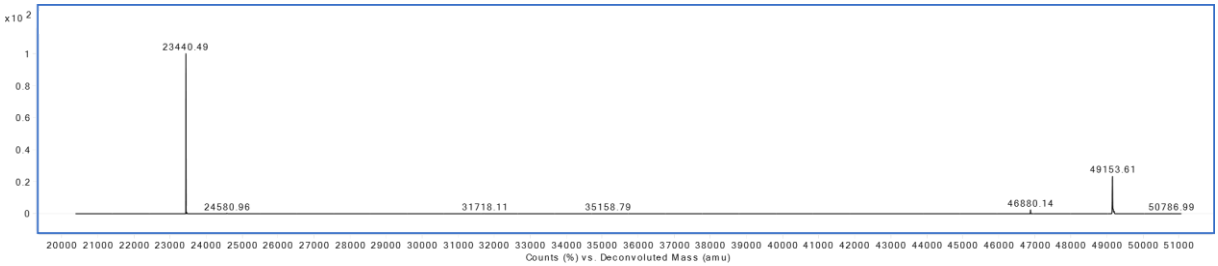
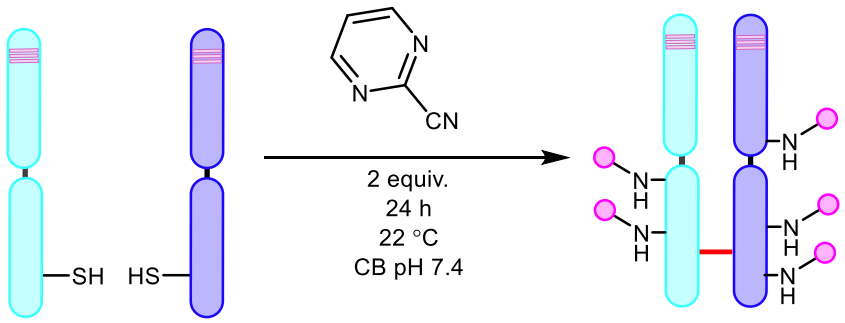


Figure S9: LCMS analysis of Reduced Trastuzumab; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of reduced trastuzumab: **LC** 23439 Da, **HL** 49153 Da. Observed: **LC** 23440 Da, **HL** 49153 Da.

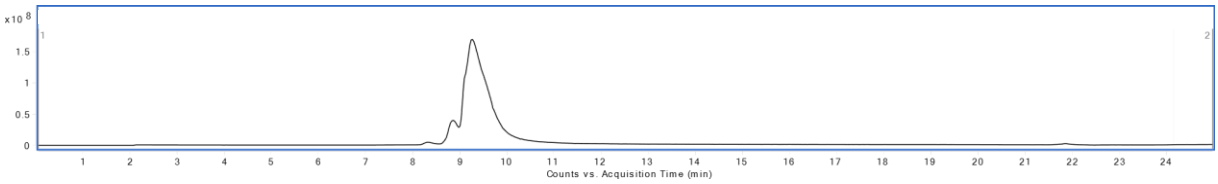
CLT reaction of Fab with 2-cyanopyrimidine.

Low equivalent reactions

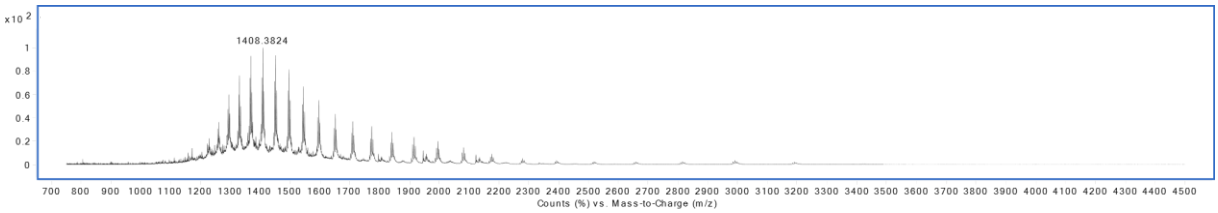


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 2-Cyanopyrimidine 1 was then added (0.4 µL, 15 mM in DMF, 2 equiv.) was added and incubated at 22 °C for 24 h. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

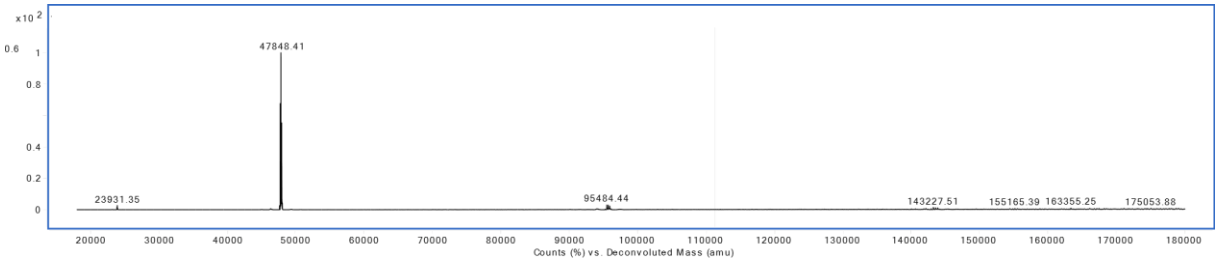
(a)



(b)



(c)



(d)

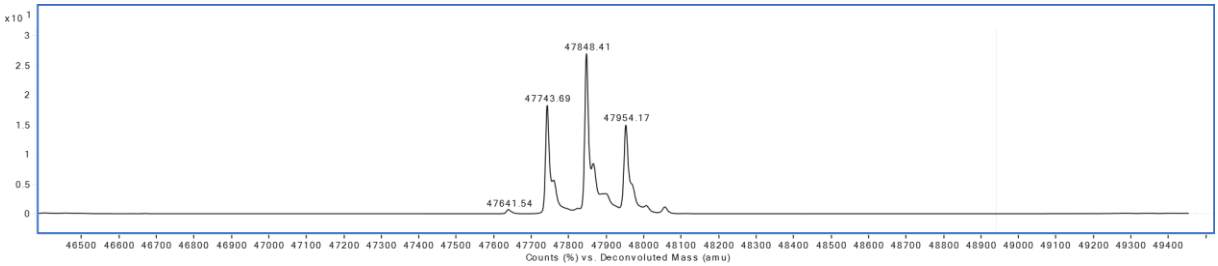
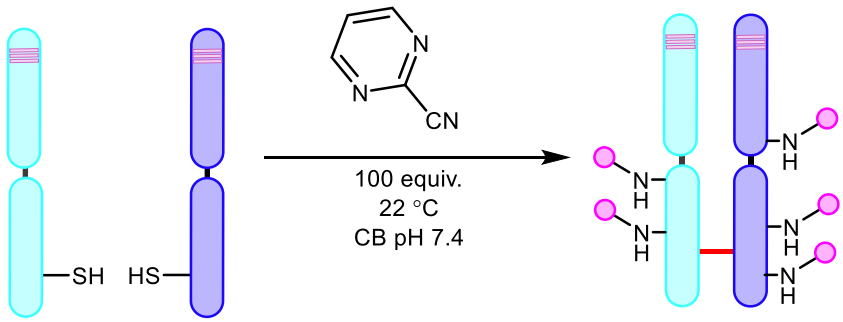


Figure S10: LCMS analysis of conjugate **17**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47638 Da (native, 0 additions), 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions). Observed: 47641 Da, 47743 Da, 47848 Da, 47954 Da.

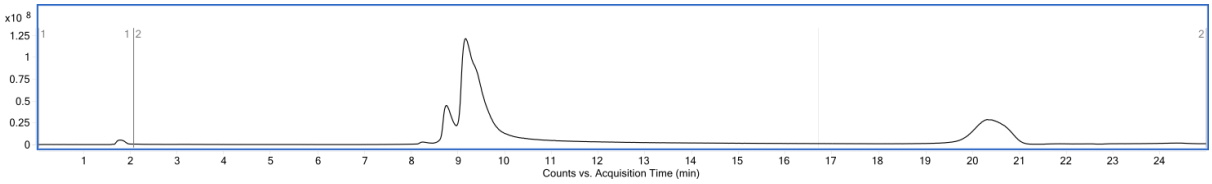
High equivalent reactions



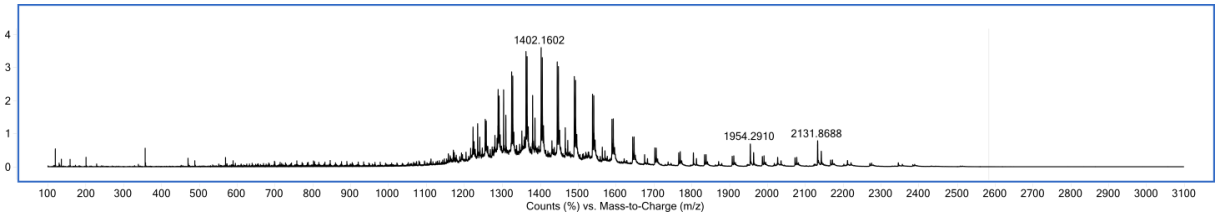
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 2-Cyanopyrimidine was then added (2 µL, 150 mM in DMF, 100 equiv.) and the reaction was incubated at 22 °C. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

15 min

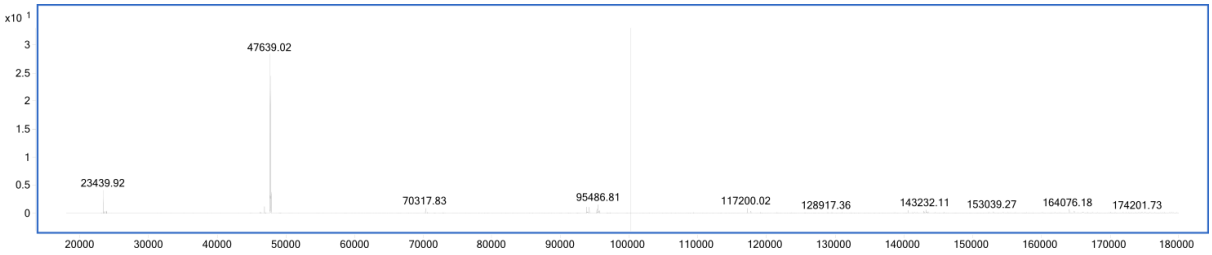
(a)



(b)



(c)



(d)

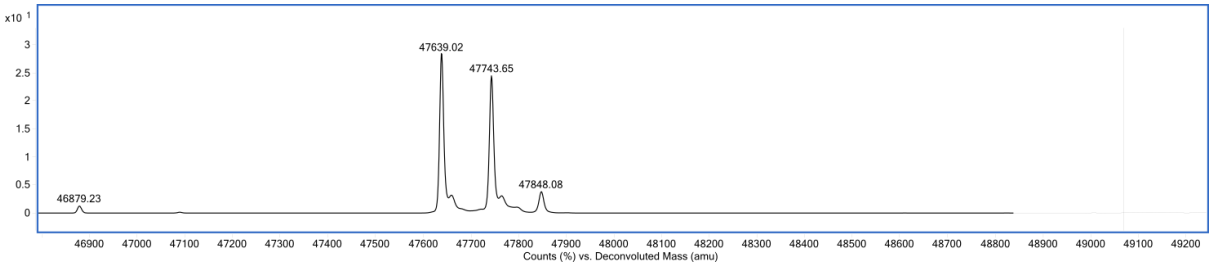
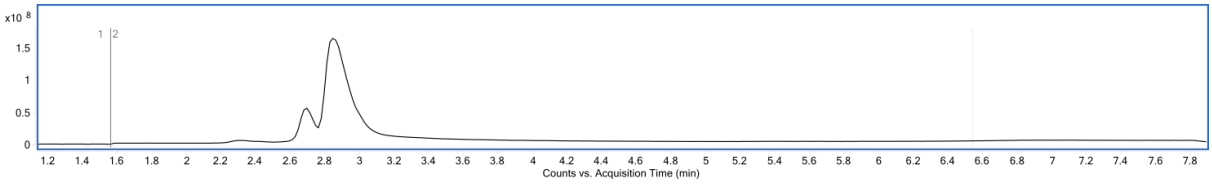


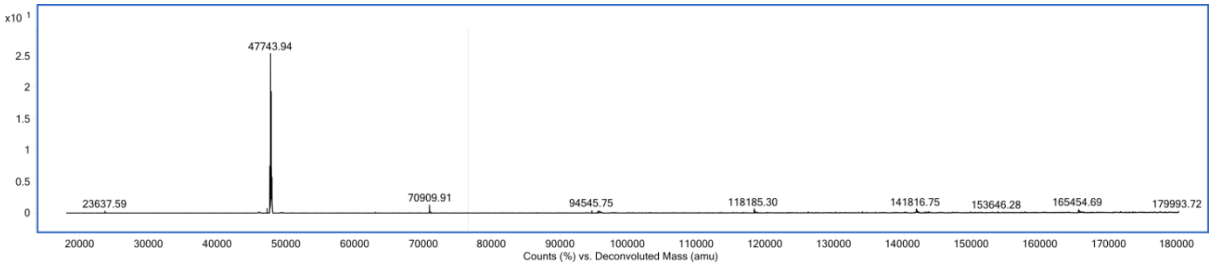
Figure S11: LCMS analysis of conjugate **18**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47639 Da (native, 0 additions), 47743 Da (1 addition), 47848 Da (2 additions). Observed: 47639 Da, 47744 Da, 47848 Da.

30 min

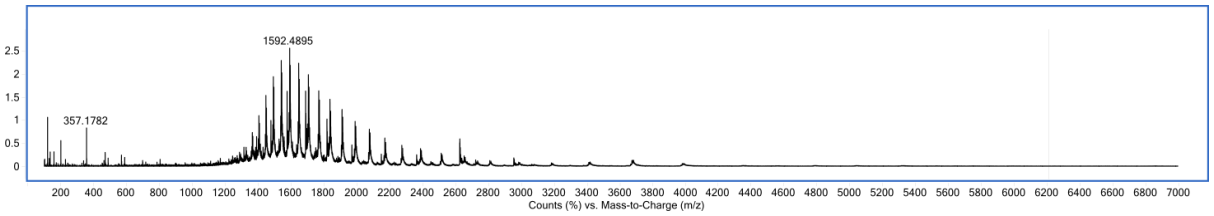
(a)



(b)



(c)



(d)

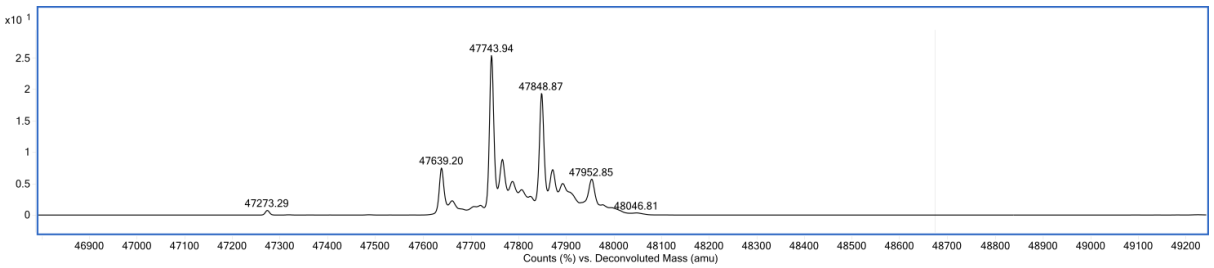
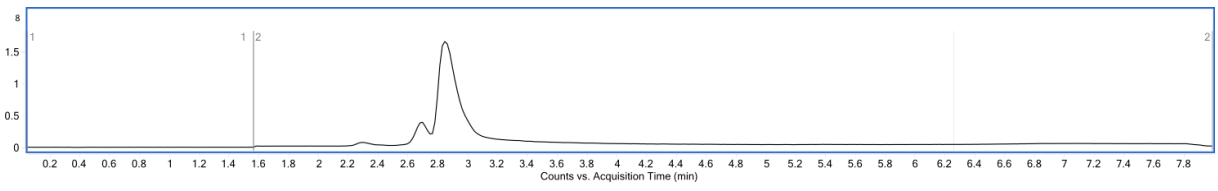


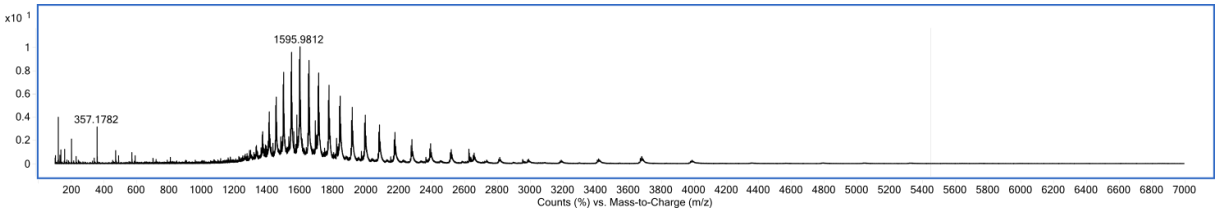
Figure S12: LCMS analysis of conjugate **18**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47639 Da (native, 0 additions), 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions). Observed: 47639 Da, 47744 Da, 47849 Da, 47953 Da.

60 min

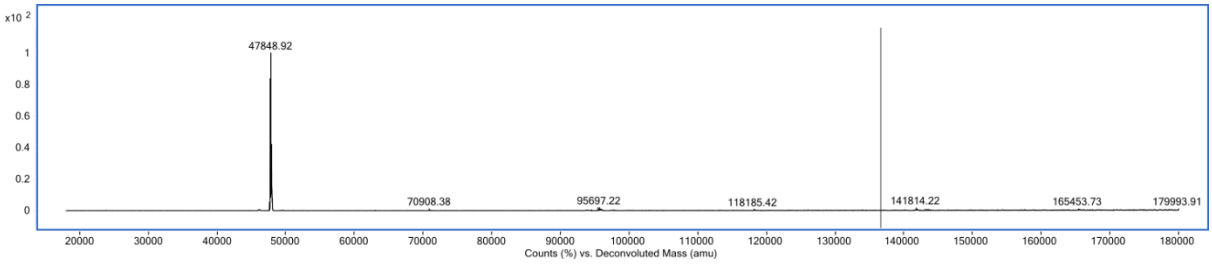
(a)



(b)



(c)



(d)

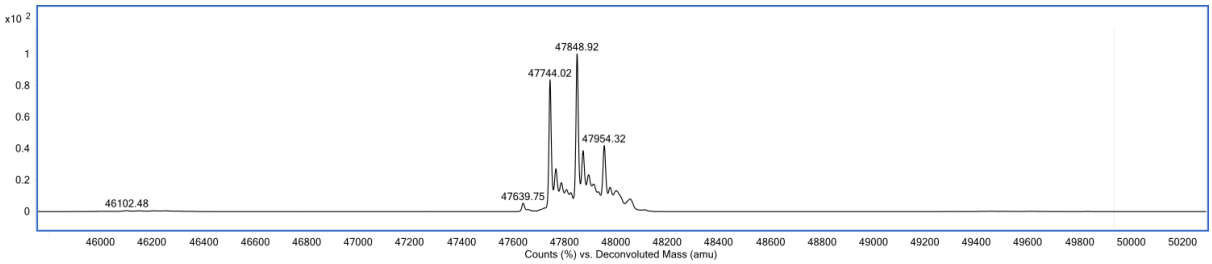
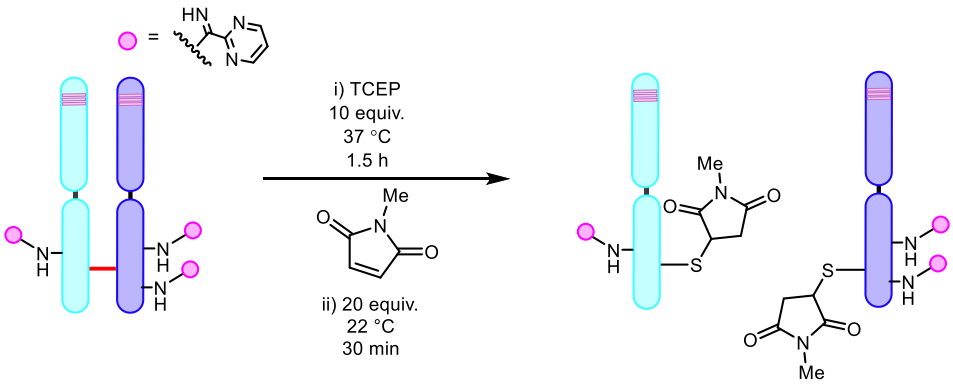


Figure S13: LCMS analysis of conjugate **18** a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47639 Da (native, 0 additions), 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions). Observed: 47639 Da, 47744 Da, 47849 Da, 47954 Da.

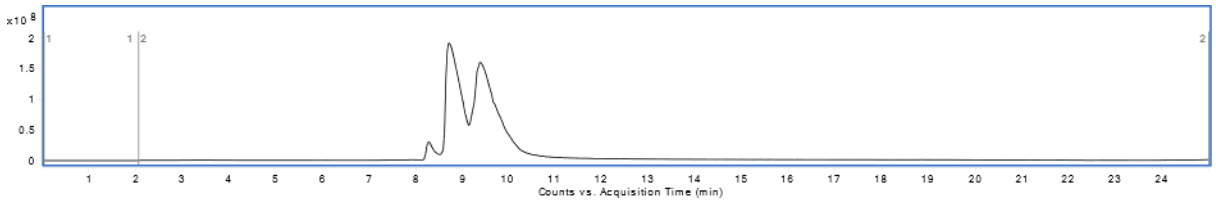
CLT reaction of Fab with 2-cyanopyrimidine (Reduce and cap)



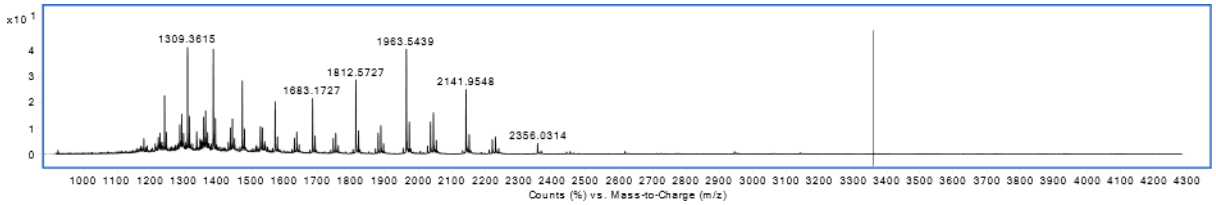
Fab conjugate **17** was buffer exchanged to conjugation buffer via ultrafiltration (10 kDa MWCO). To this solution was added tris(2-carboxyethyl)phosphine (TCEP) (15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Following this NMM (20 mM in DMF, 20 equiv.) was added, and the reaction was left at RT for 30 min. Excess reagent was

then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

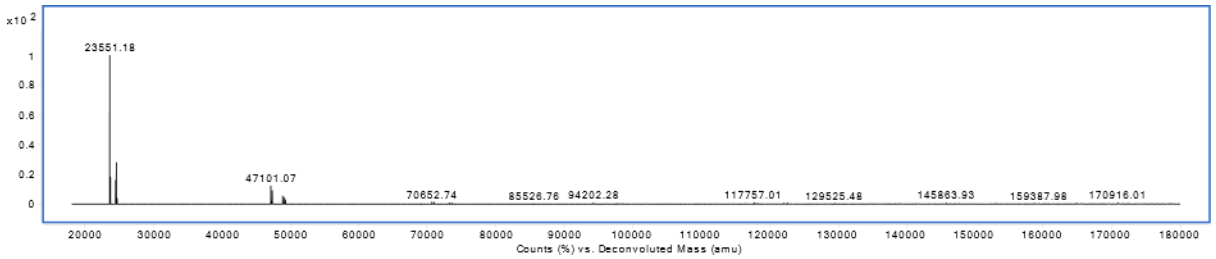
(a)



(b)



(c)



(d)

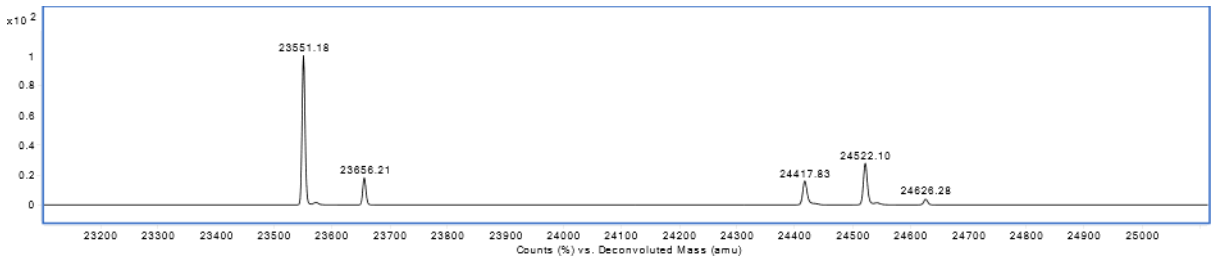
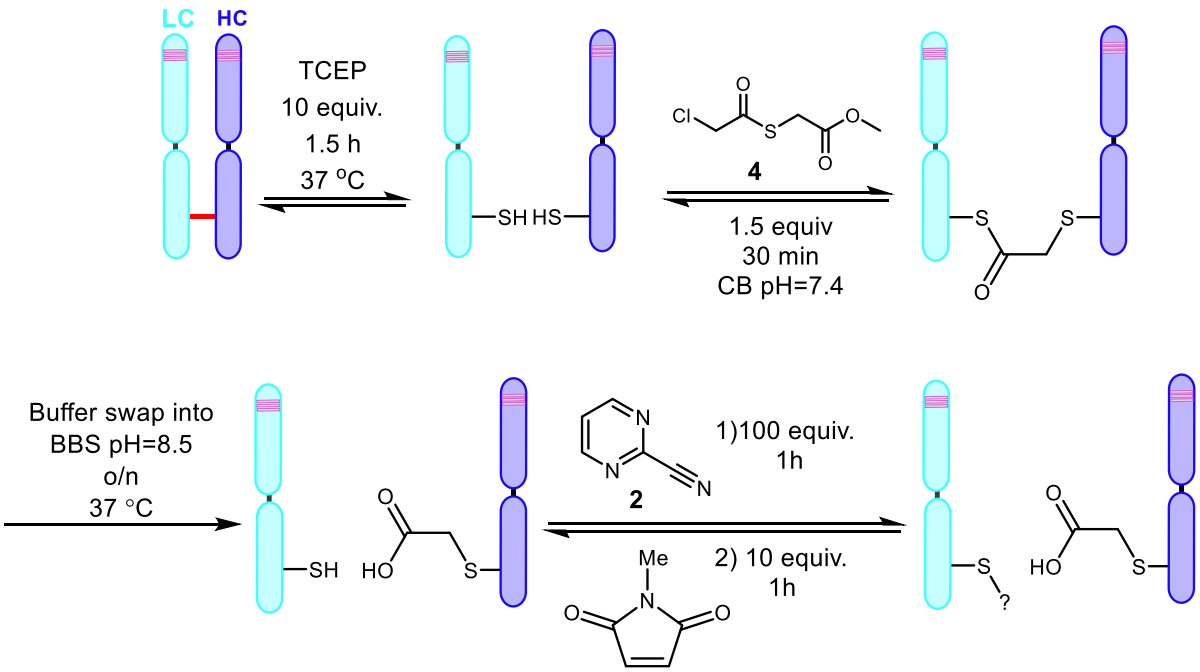


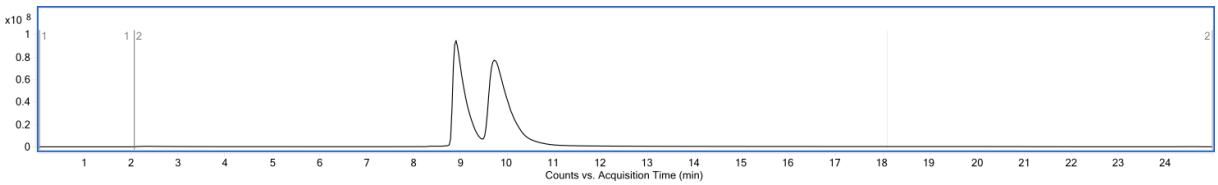
Figure S14: LCMS analysis of reduced and capped conjugate **17**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of capped pyrimidine CLT conjugate: 23550 Da (NMM capped **LC** with 0 additions), 23665 Da (NMM capped **LC** with 1 addition); 24416 Da (NMM capped **HC** with 1 addition), 24521 Da (NMM capped **HC** with 2 additions), 24626 Da (NMM capped **HC** with 3 additions). Observed: **LC** 23551 Da, 23656 Da; **HC** 24418 Da, 24522 Da, 24626 Da. Peaks around 47101 Da (2 x 23551 Da) are proposed to be artifacts formed during ionisation.

Mechanistic study of nitrile CLT

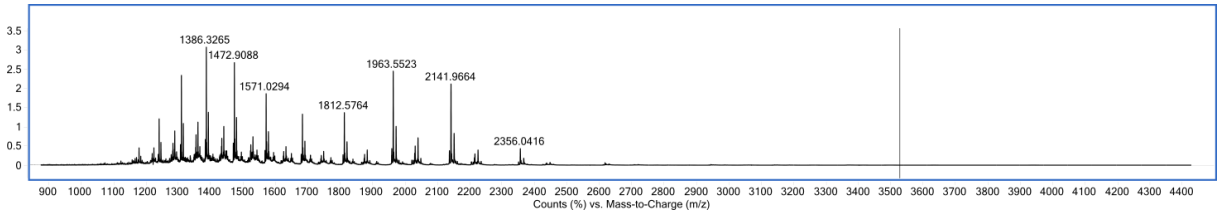


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Following this methyl 2-[(2-chloroacetyl)sulfanyl]acetate (0.4 µL, 12 mM in DMF, 1.5 equiv.) was added and the sample was incubated at 22 °C for 30 min. The sample was then buffer swapped into BBS pH 8.5 and left at 37 °C for 12 h. 2-Cyanopyrimidine (100 equiv.) was then added, and the reaction was left for 1 h at 22 °C. The excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) and NMM (10 equiv.) was added for 1 h at 22 °C. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

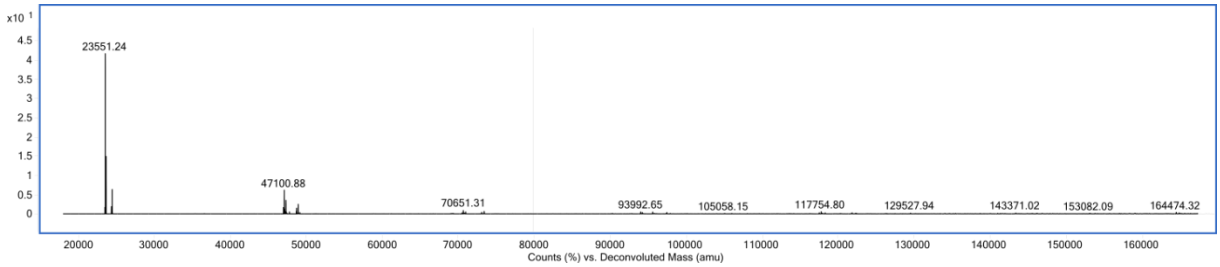
(a)



(b)



(c)



(d)

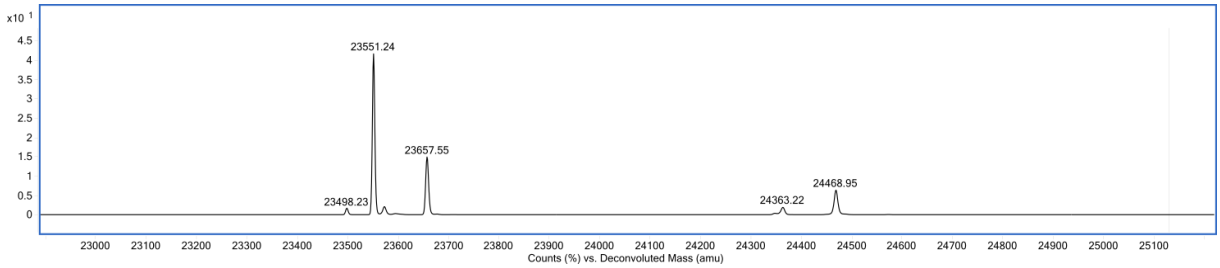
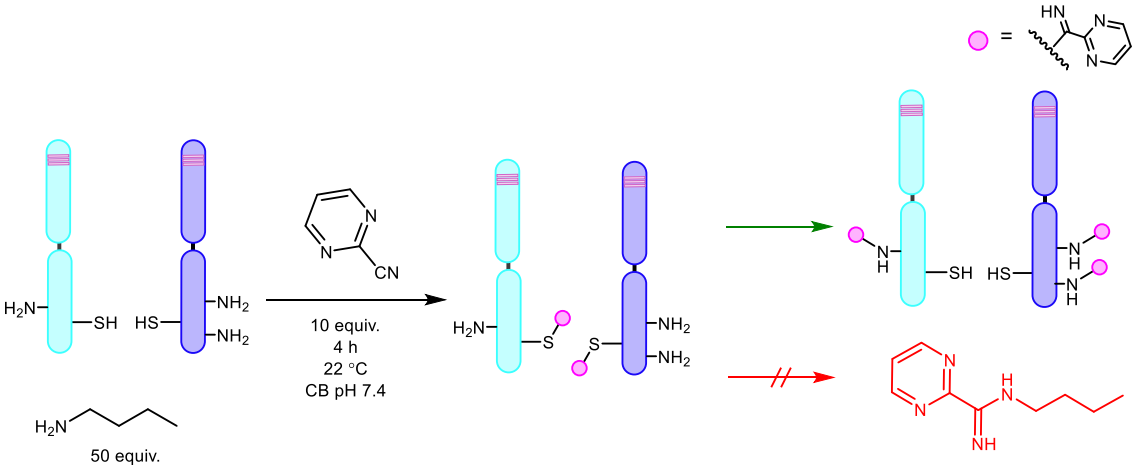


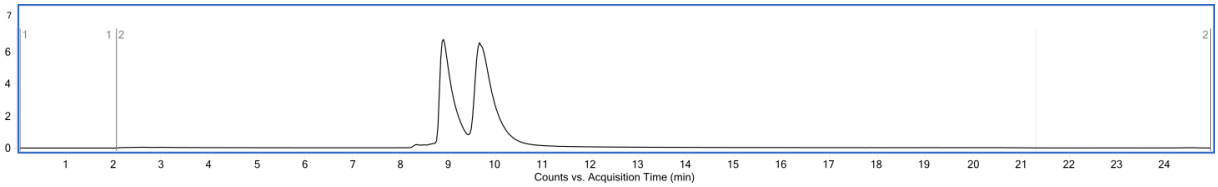
Figure S15: LCMS analysis of mechanistic study; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of capped pyrimidine CLT conjugate: 23498 Da (α-chlorothioester capped **LC** with 0 additions), 23550 Da (NMM capped **LC** with 0 additions), 23655 Da (NMM capped **LC** with 1 addition); 24364 Da (α-chlorothioester capped **HC** with 1 addition), 24469 Da (α-chlorothioester capped **HC** with 2 additions). Observed: **LC** 23498 Da, 23551 Da, 23658 Da; **HC** 24363 Da, 24469 Da.

Intramolecular selectivity competition experiment between Fab and *n*-butylamine

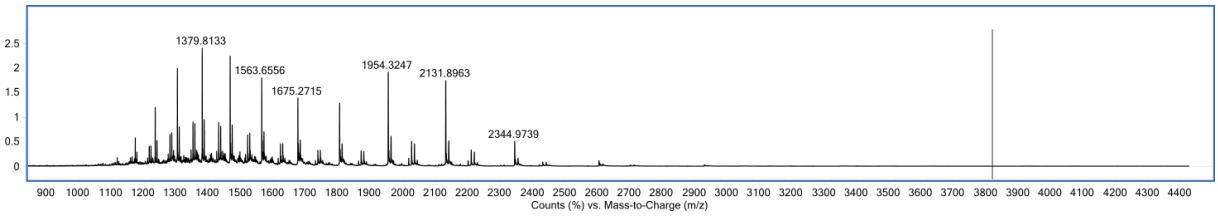


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. To this, *n*-butylamine (0.5 µL, 300 mM in DMF, 50 equiv.) was added followed by 2-cyanopyrimidine (0.2 µL, 150 mM in DMF, 10 equiv.), the reaction was incubated at 22 °C for 4 h. Excess reagent was then removed *via* ultrafiltration (7 kDa MWCO, ZebaSpin) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

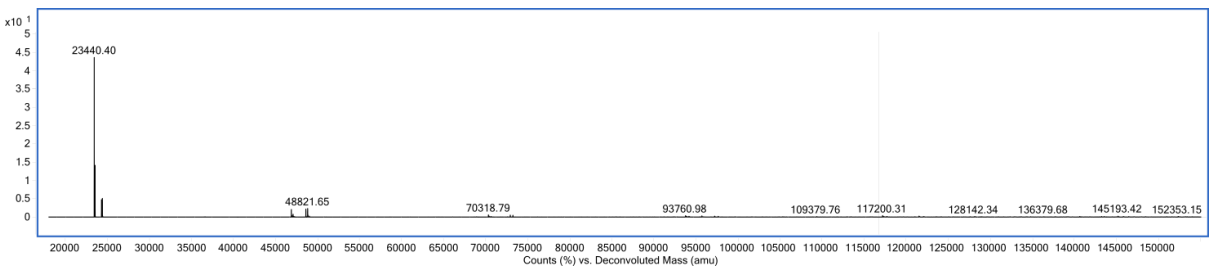
(a)



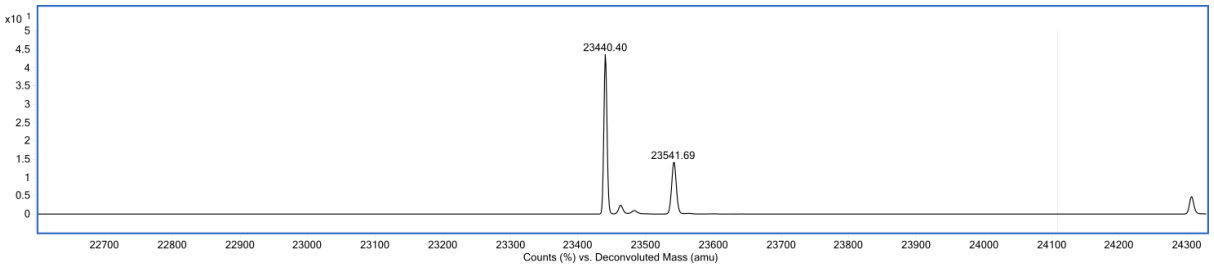
(b)



(c)



(d)



(e)

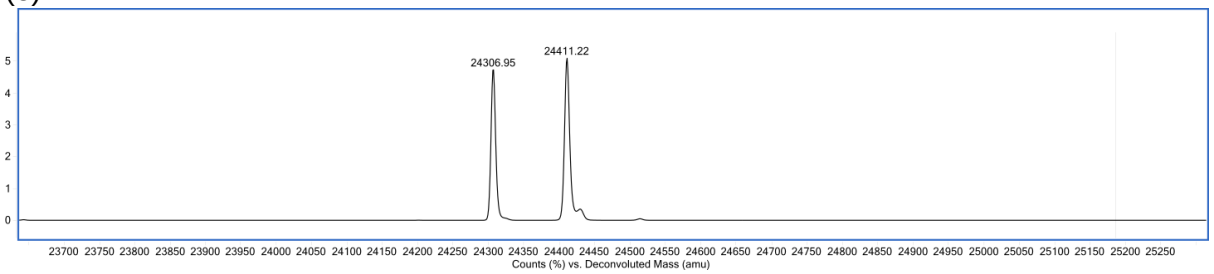
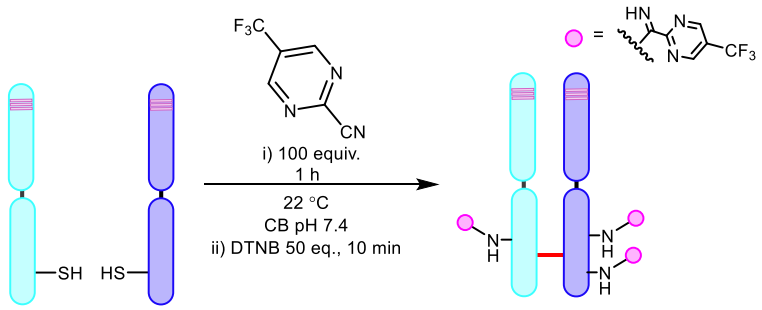


Figure S16: LCMS analysis of reduced and capped conjugate **S7b**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum **LC** region; Expected mass of pyrimidine CLT conjugate: 23440 Da (**LC** with 0 additions), 23545 Da (**LC** with 1 addition), e) zoomed in deconvoluted ion series mass spectrum **HC** region; Expected mass of pyrimidine CLT conjugate: 24305 Da (**HC** with 1 addition), 24410 Da (**HC** with 2 additions). Observed: **LC** 23440 Da, 23542 Da; **HC** 24306 Da, 24411 Da.

5-(trifluoromethyl)pyrimidine-2-carbonitrile **6** reactions on Fab

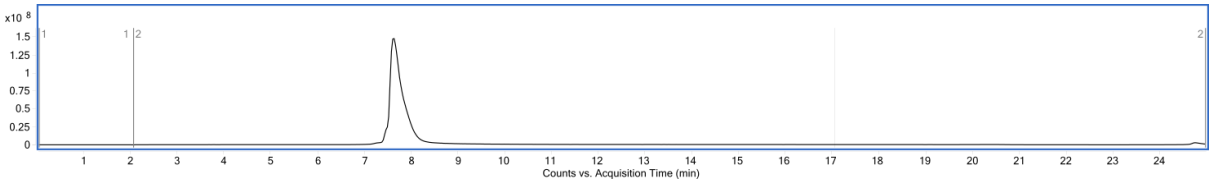
High equivalent reactions



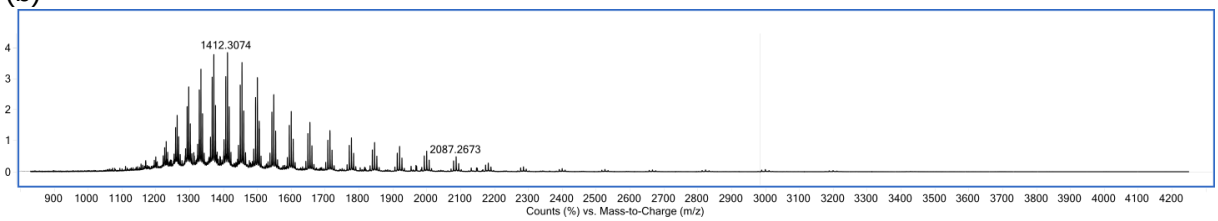
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5-(trifluoromethyl)pyrimidine-2-carbonitrile was then added (2.0 µL, 150 mM in DMF, 100 equiv.) was added and incubated at 22 °C for 1 h. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. The sample was desalted into HPLC grade water (7 kDa MWCO,

ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

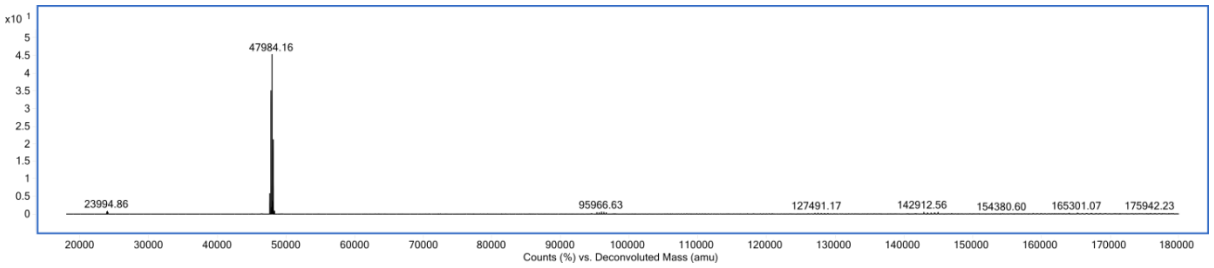
(a)



(b)



(c)



(d)

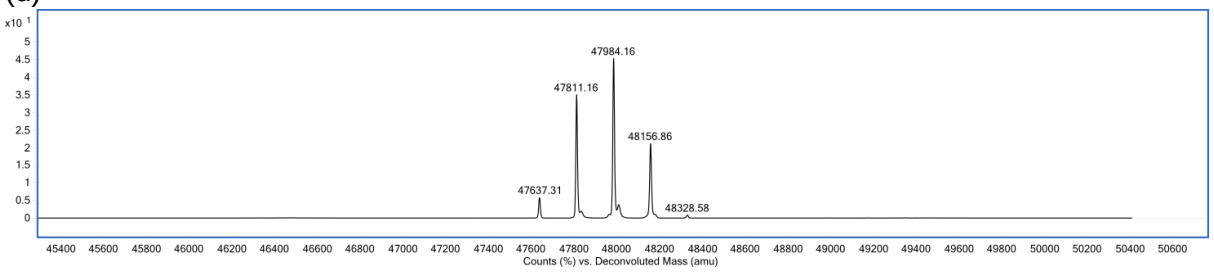
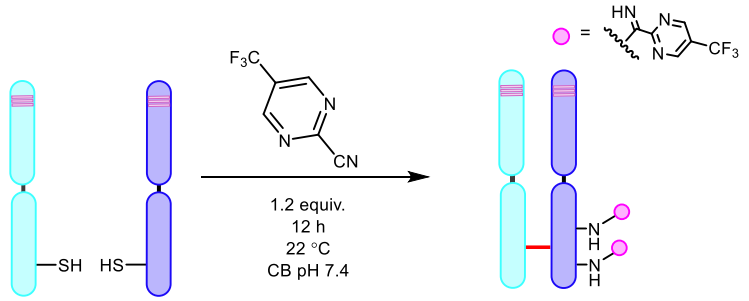


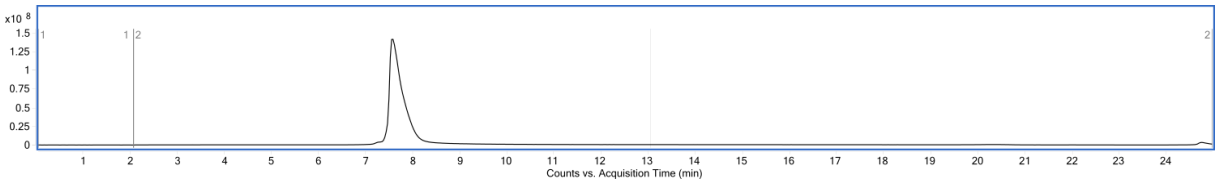
Figure S17: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47637 Da (native, 0 additions), 47810 Da (1 addition), 47983 Da (2 additions), 48156 Da (3 additions), 48329 Da (4 additions). Observed: 47637 Da, 47811 Da, 47984 Da, 48157 Da, 48329 Da.

Low equivalent reactions

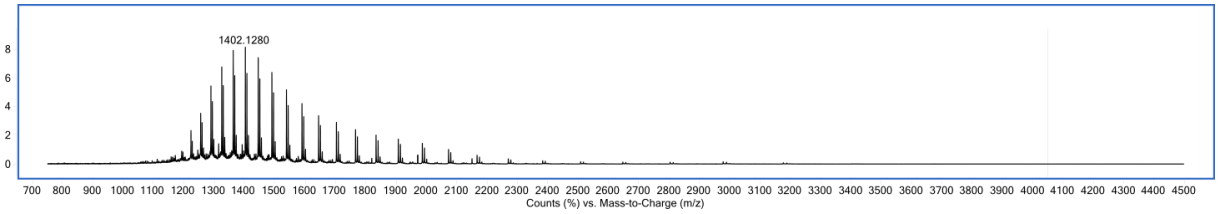


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5-(trifluoromethyl)pyrimidine-2-carbonitrile was then added (0.36 µL, 10 mM in DMF, 1.2 equiv.) and incubated at 22 °C for 12 h. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

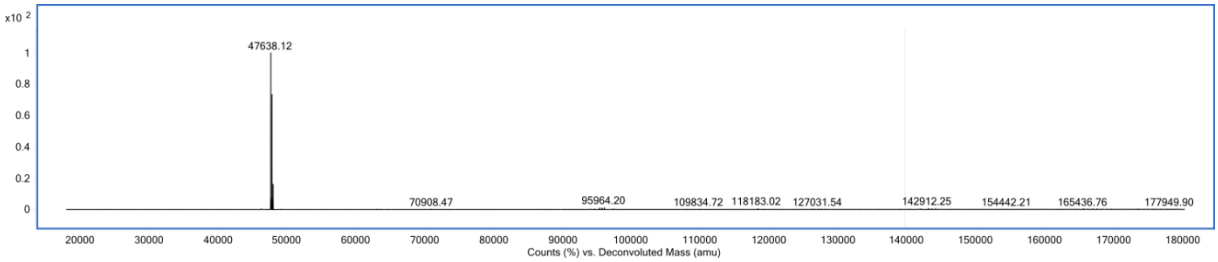
(a



(b)



(c)



(d)

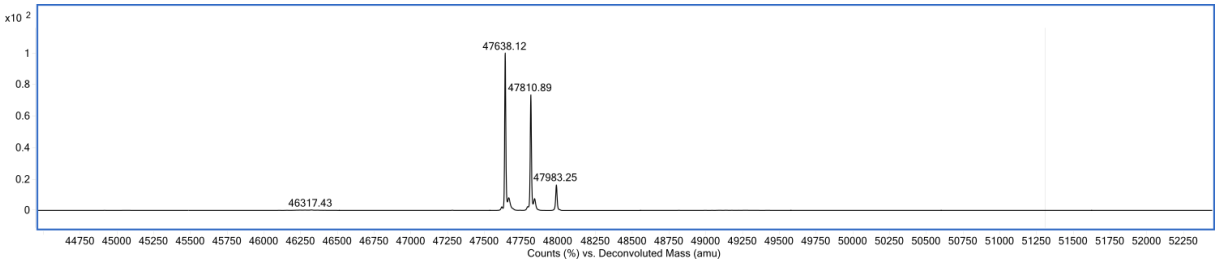
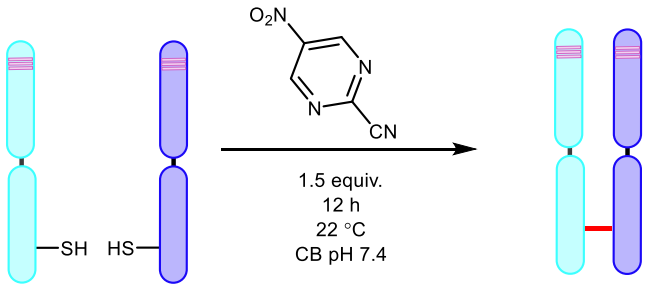


Figure S18: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47637 Da (native, 0 additions), 47810 Da (1 addition), 47983 Da (2 additions). Observed: 47638 Da, 47811 Da, 47983 Da.

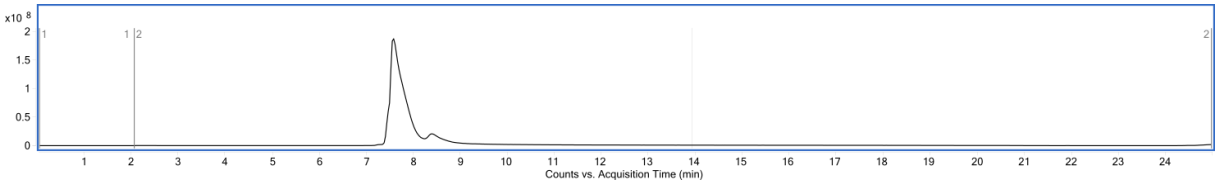
5-Nitropyrimidine-2-carbonitrile **7** reactions on Fab.

Low equivalent reactions

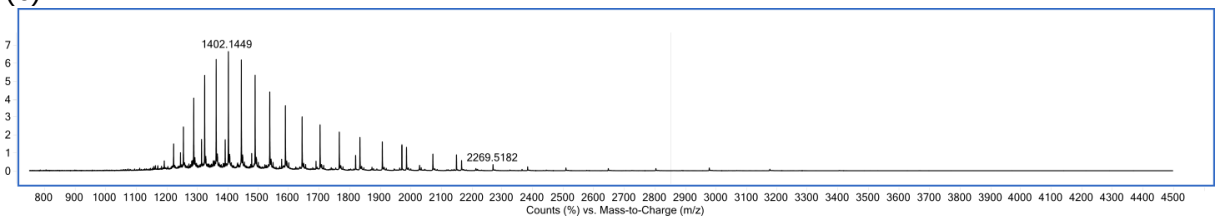


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5-nitropyrimidine-2-carbonitrile was then added (1.5 µL, 3 mM in DMF, 1.5 equiv.) and incubated at 22 °C for 12 h. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

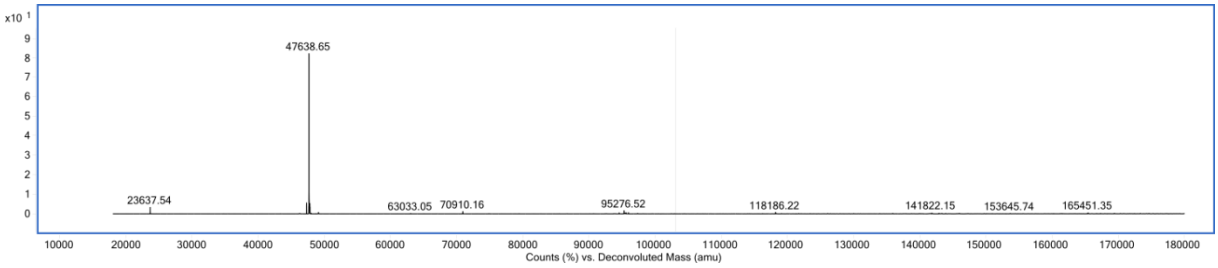
(a)



(b)



(c)



(d)

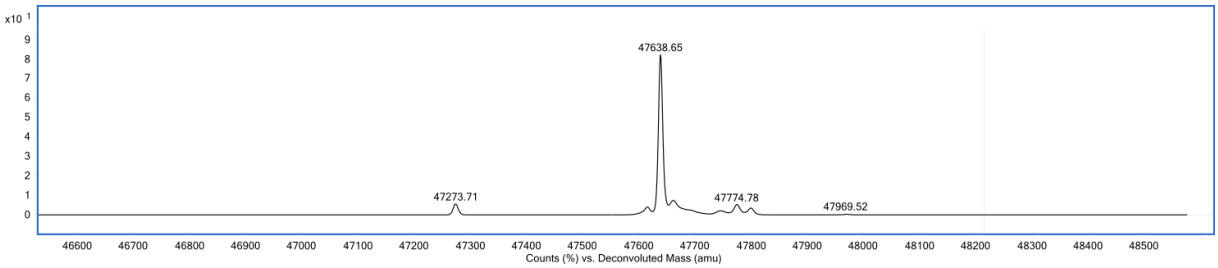
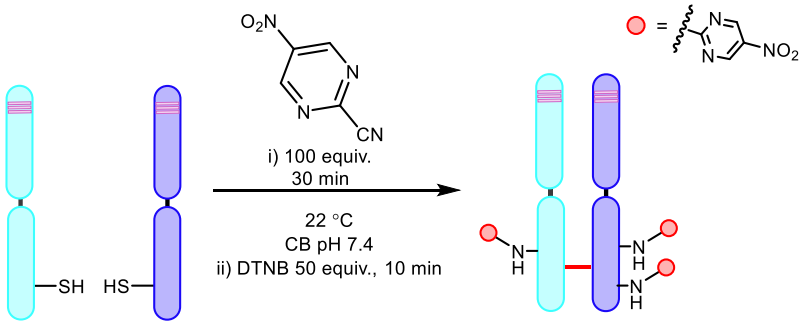


Figure S19: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47788 Da (1 addition). Observed: 47639 Da.

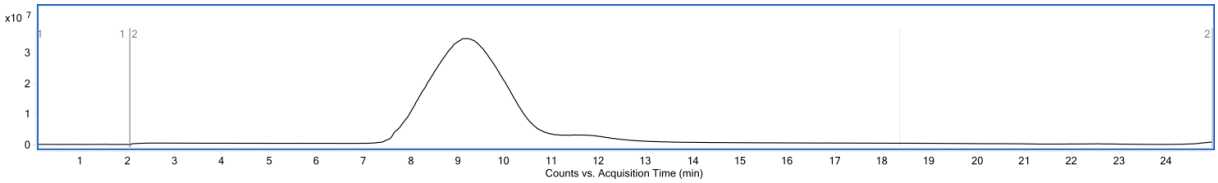
High equivalent reactions



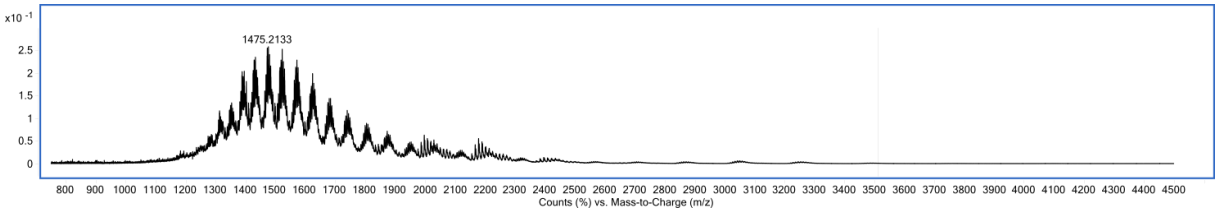
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5-nitropyrimidine-2-carbonitrile was then added (2.0 µL, 150 mM in DMF, 100 equiv.) and incubated at 22 °C for 30 min. Following this DTNB (1.0 µL,

150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

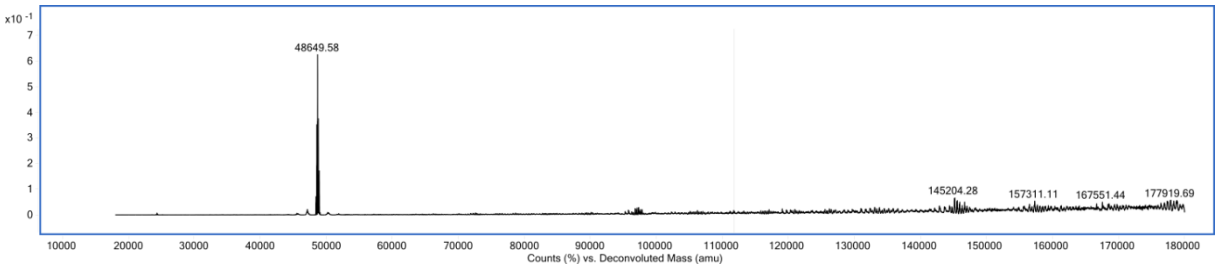
(a)



(b)



(c)



(d)

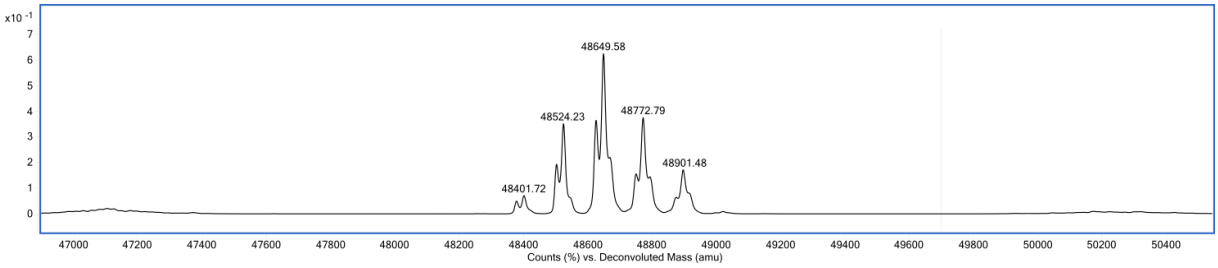
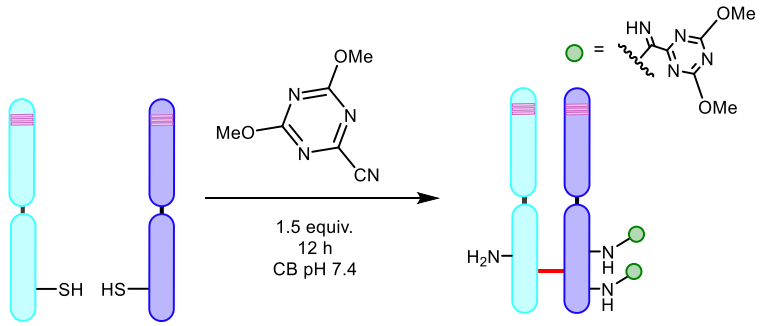


Figure S20: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum.

4,6-Dimethoxy-1,3,5-triazine-2-carbonitrile **10** reactions on Fab.

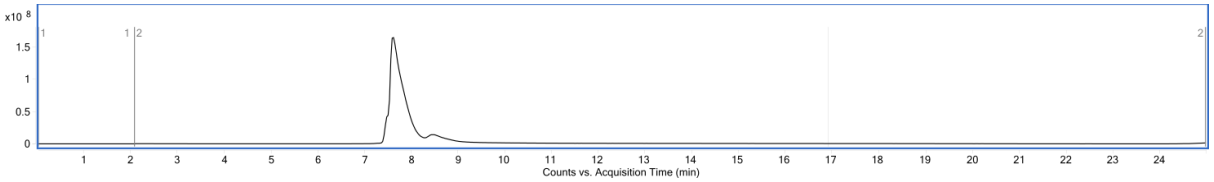
Low equivalent reactions



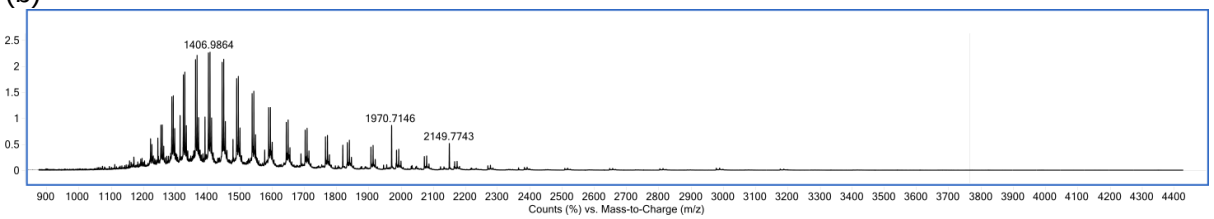
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 4,6-dimethoxy-1,3,5-triazine-2-carbonitrile was then added (1.5 µL, 3 mM in DMF, 1.5 equiv.) and incubated for 12 h. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

22 °C

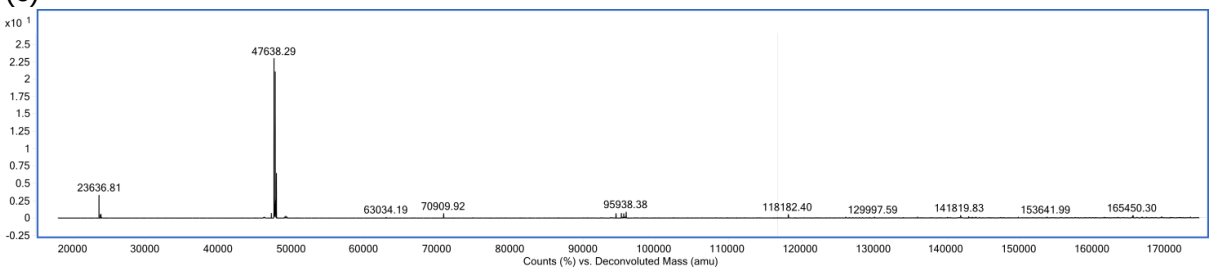
(a)



(b)



(c)



(d)

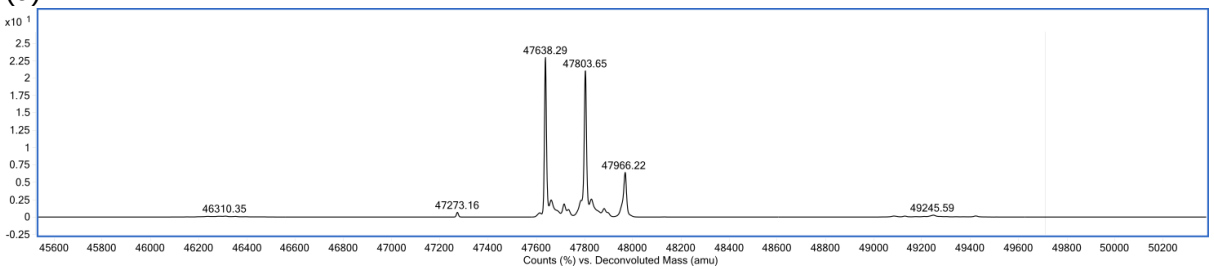
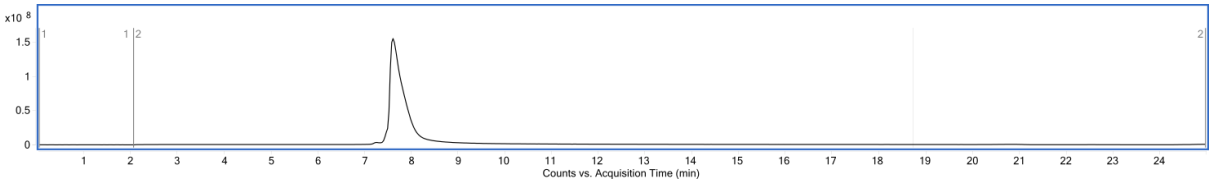


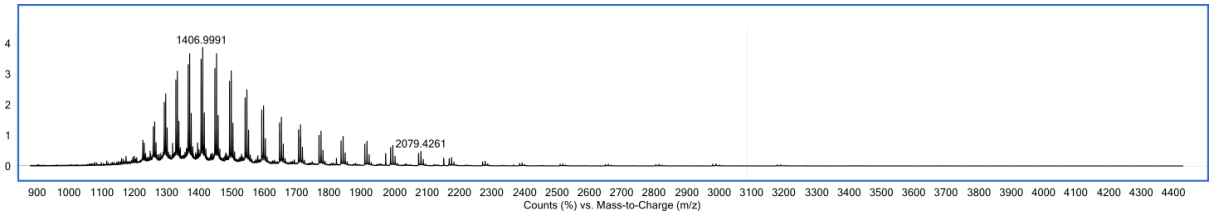
Figure S21: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47804 Da (1 addition), 47970 Da (2 additions). Observed: 47638 Da, 47804 Da, 47966 Da.

37 °C

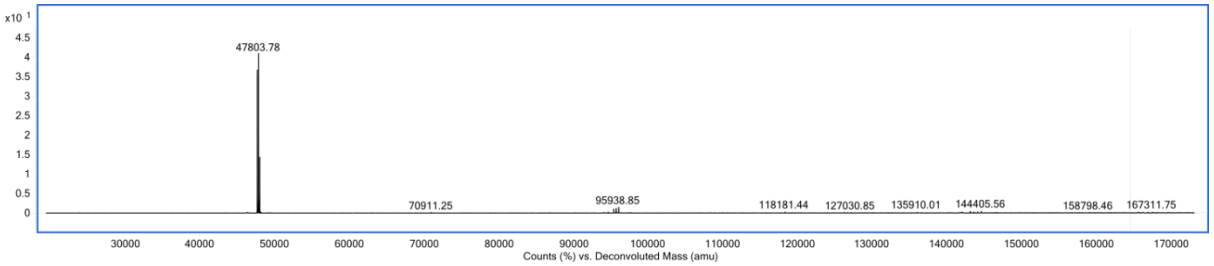
(a)



(b)



(c)



(d)

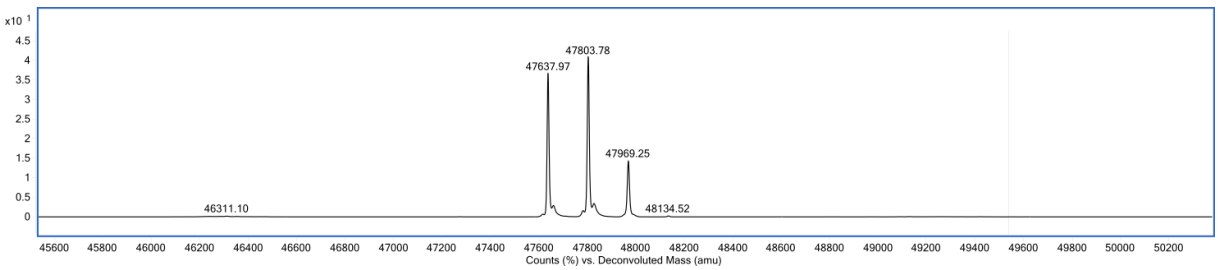
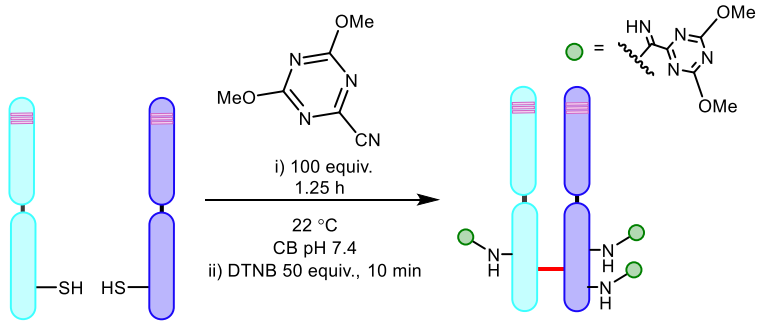


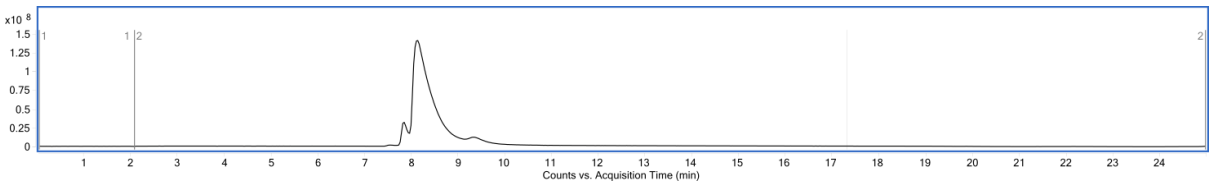
Figure S22: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47804 Da (1 addition), 47970 Da (2 additions). Observed: 47638 Da, 47804 Da, 47969 Da.

High equivalent reactions

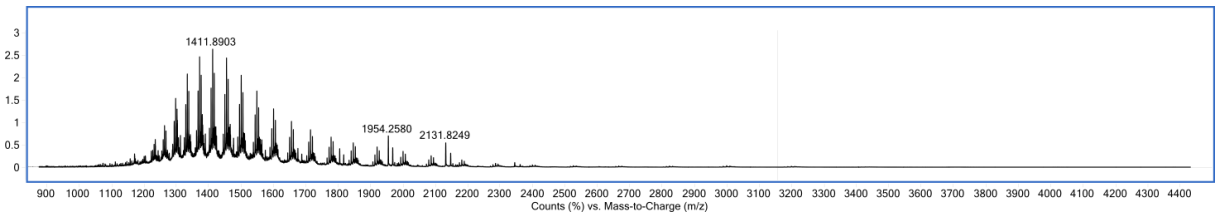


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 4,6-dimethoxy-1,3,5-triazine-2-carbonitrile was then added (2.0 µL, 150 mM in DMF, 100 equiv.) and incubated at 22 °C for 1.25 h. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

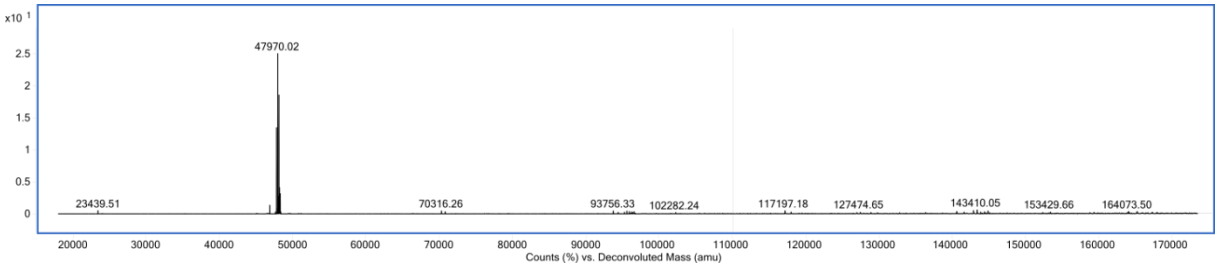
(a)



(b)



(c)



(d)

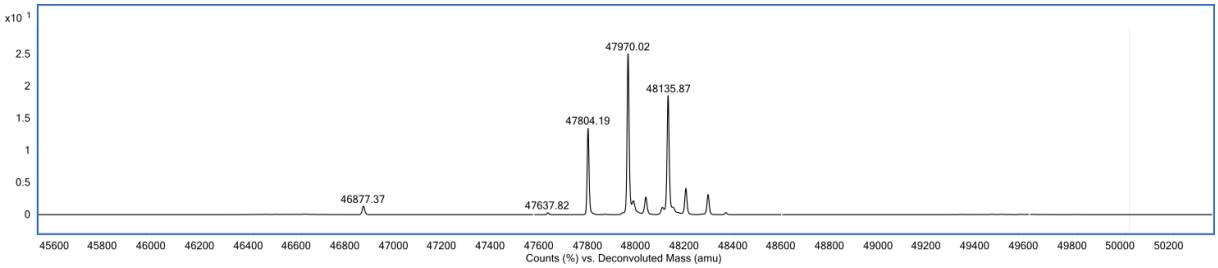
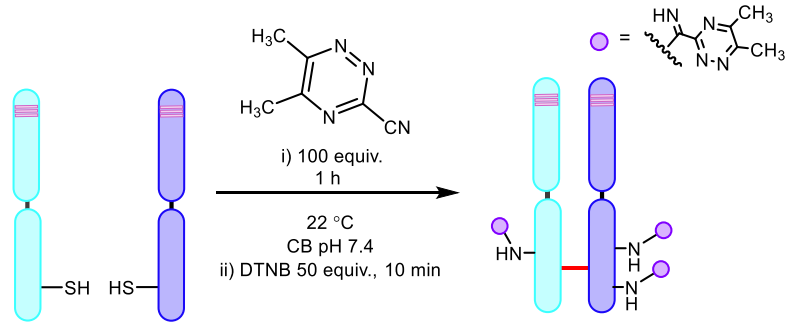


Figure S23: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47804 Da (1 addition), 47970 Da (2 additions), 48136 Da (3 additions). Observed: 47638 Da, 47804 Da, 47970 Da, 48136 Da.

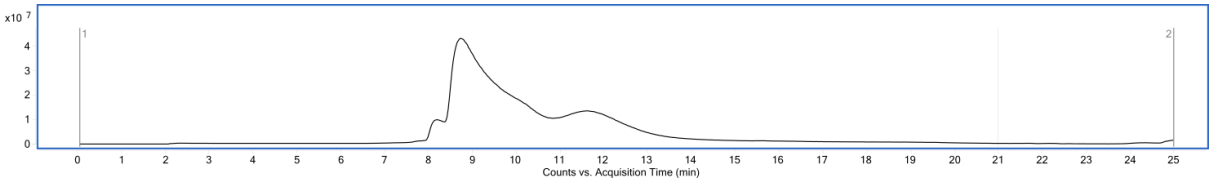
5,6-Dimethyl-1,2,4-triazine-3-carbonitrile **11** reactions on Fab.

High equivalent reactions

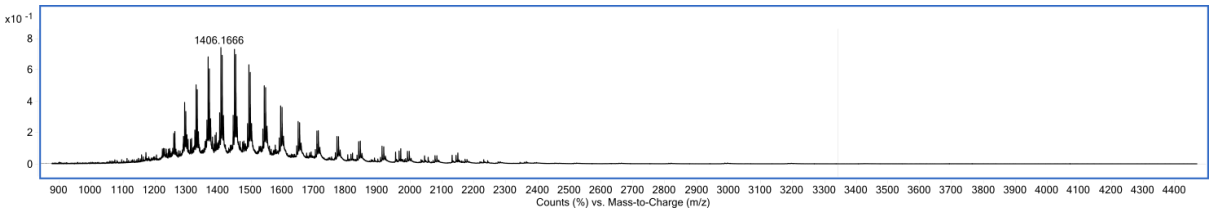


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5,6-dimethyl-1,2,4-triazine-3-carbonitrile was then added (2.0 µL, 150 mM in DMF, 100 equiv.) and incubated at 22 °C for 1 h. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

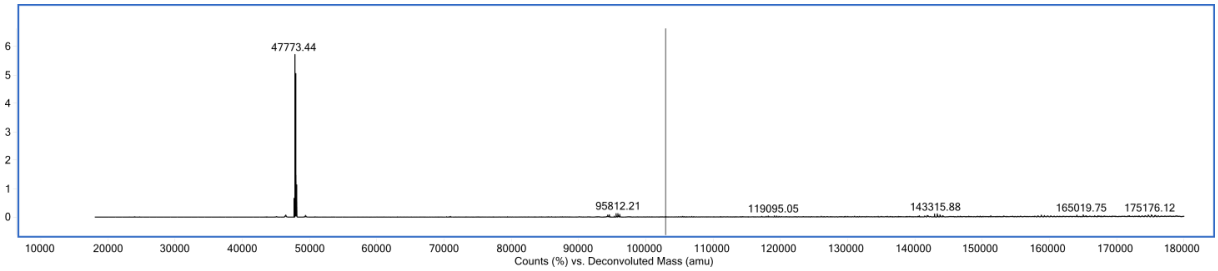
(a)



(b)



(c)



(d)

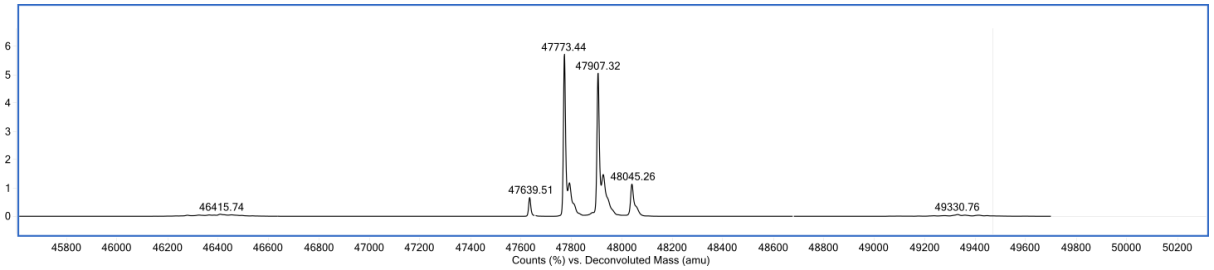
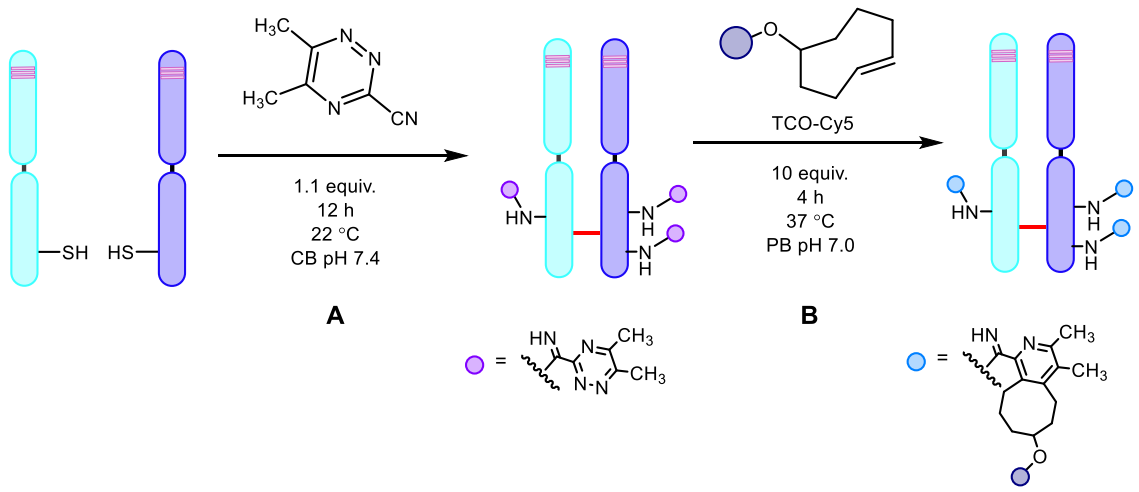


Figure S24: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47639 Da (native, 0 additions), 47773 Da (1 addition), 47907 Da (2 additions), 48041 Da (3 additions). Observed: 47639 Da, 47773 Da, 47907 Da, 48045 Da.

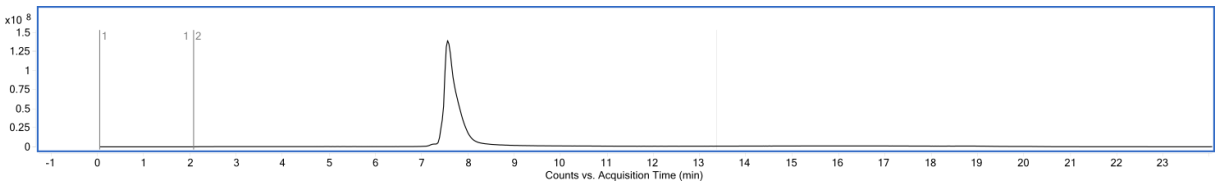
Low equivalent reactions



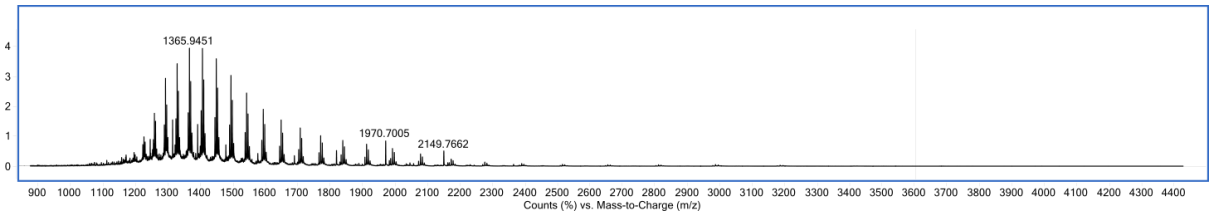
A

Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5,6-dimethyl-1,2,4-triazine-3-carbonitrile was then added (1.1 µL, 3 mM in DMF, 1.5 equiv.) and incubated for 12 h at 22 °C. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}. cm^{-1}.

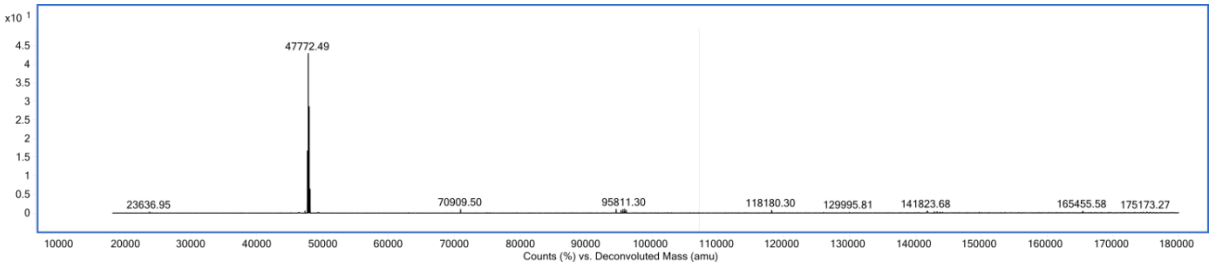
(a)



(b)



(c)



(d)

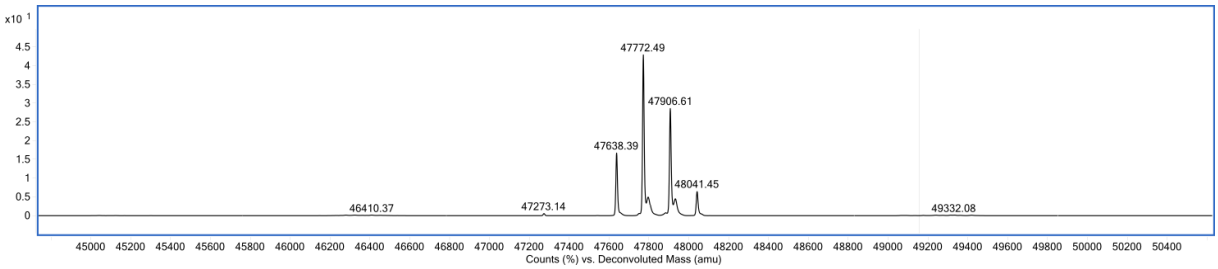
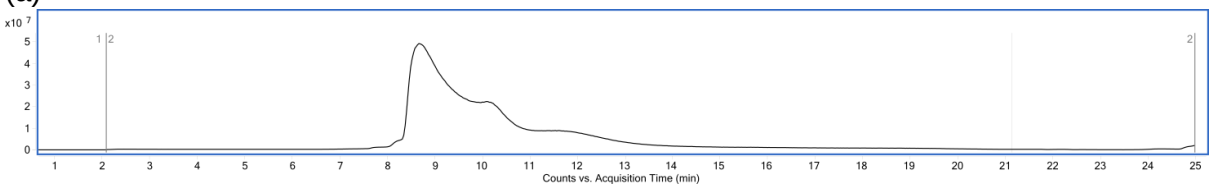


Figure S25: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47773 Da (1 addition), 47907 Da (2 additions), 48041 Da (3 additions). Observed: 47638 Da, 47773 Da, 47907 Da, 48042 Da.

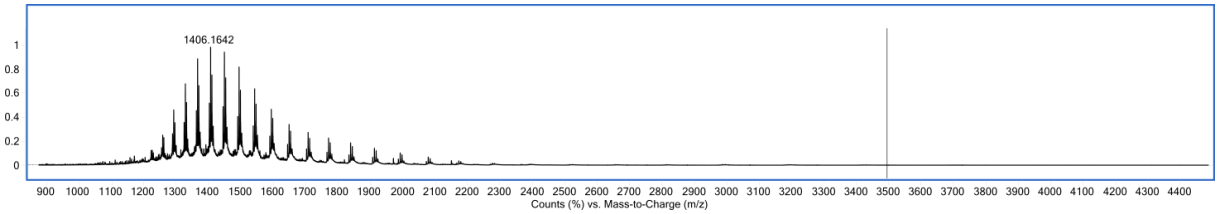
B

Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5,6-dimethyl-1,2,4-triazine-3-carbonitrile was then added (1.1 µL, 3 mM in DMF, 1.5 equiv.) and incubated for 12 h 22 °C. Excess reagent was removed by ultrafiltration (10 kDa MWCO) into PB pH 7.0. To this, was added TCO-Cy5 (20 mM in DMF, 10 equiv.), the reaction was then left in the dark at 37 °C for 4 h, 300 rpm. Lastly, the excess reagent was then removed by ultrafiltration (10 kDa MWCO) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

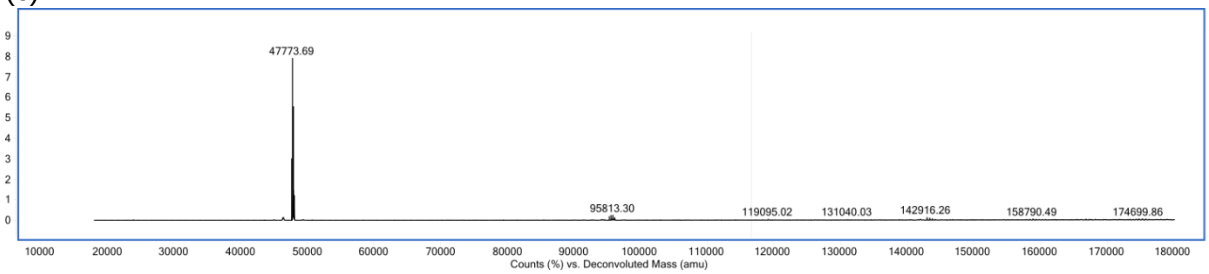
(a)



(b)



(c)



(d)

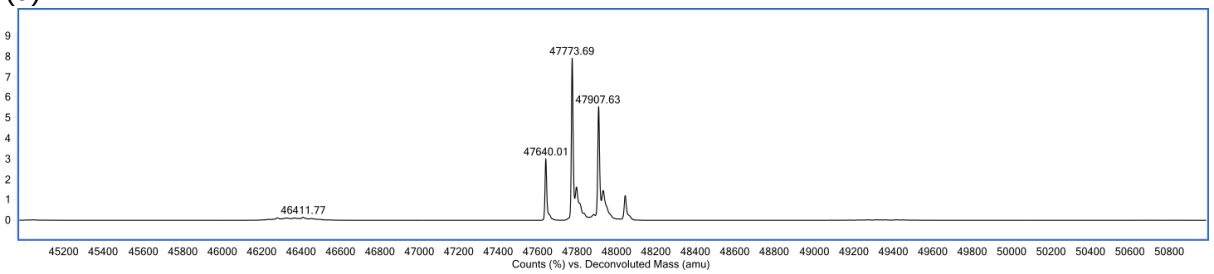
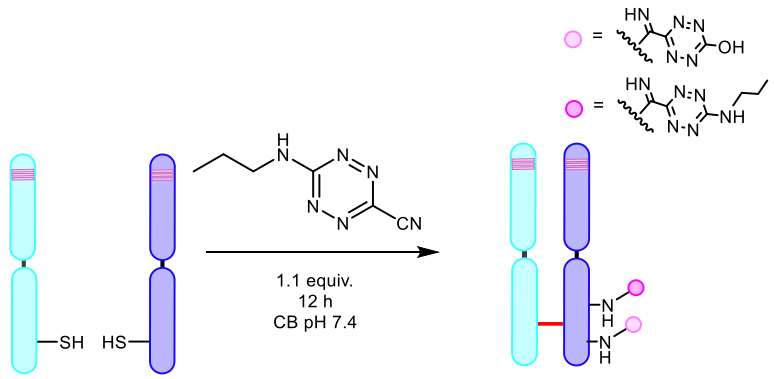


Figure S26: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47639 Da (native, 0 additions), 47773 Da (1 addition), 47907 Da (2 additions), 48041 Da (3 additions). Expected mass of iEDDA conjugate: 48732 Da (1 addition), 49825 Da (2 additions), 50918 Da (3 additions). Observed: 47640 Da, 47773 Da, 47908 Da.

6-(Propylamino)-1,2,4,5-tetrazine-3-carbonitrile **12** reactions on Fab.

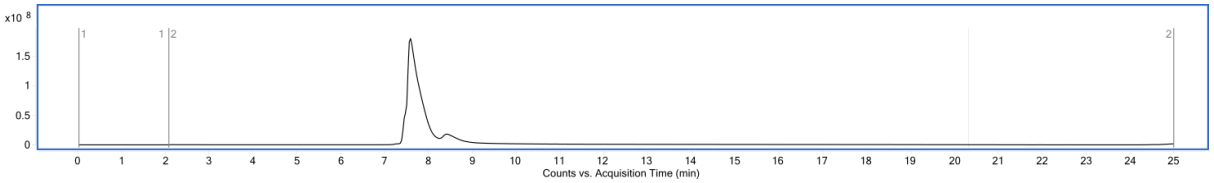
Low equivalent reactions



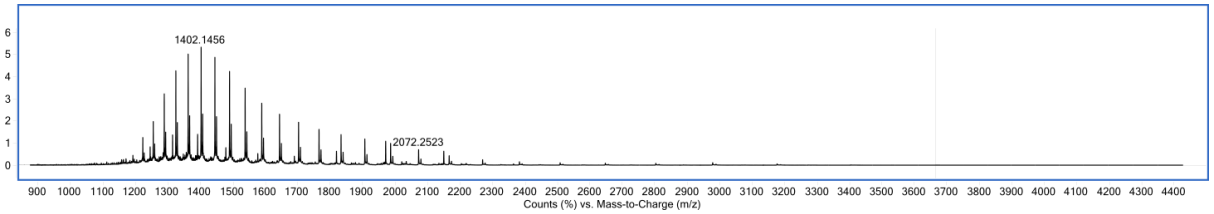
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 6-(propylamino)-1,2,4,5-tetrazine-3-carbonitrile was then added (1.1 µL, 3 mM in DMF, 1.5 equiv.) and incubated for 12 h. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}. cm^{-1}.

22 °C

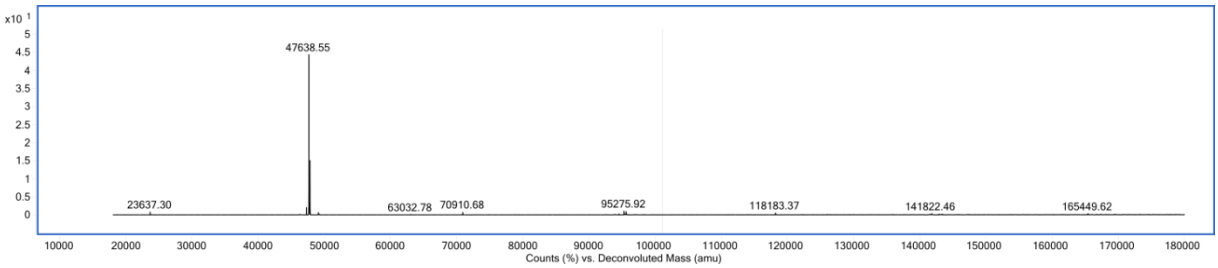
(a)



(b)



(c)



(d)

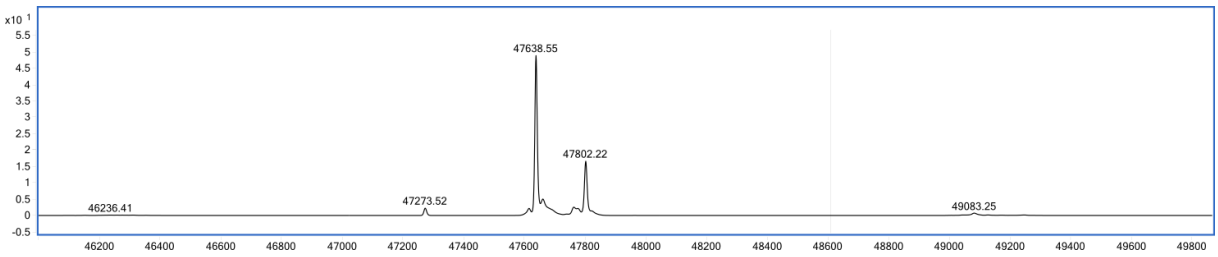
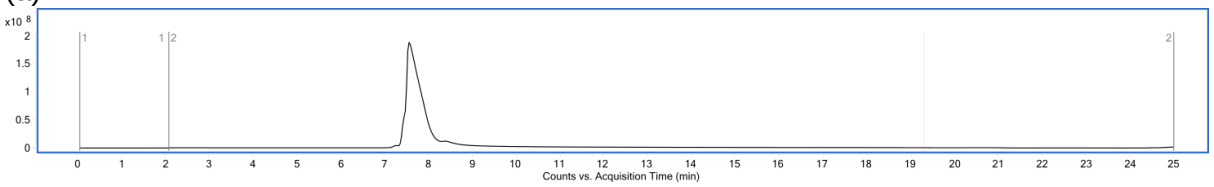


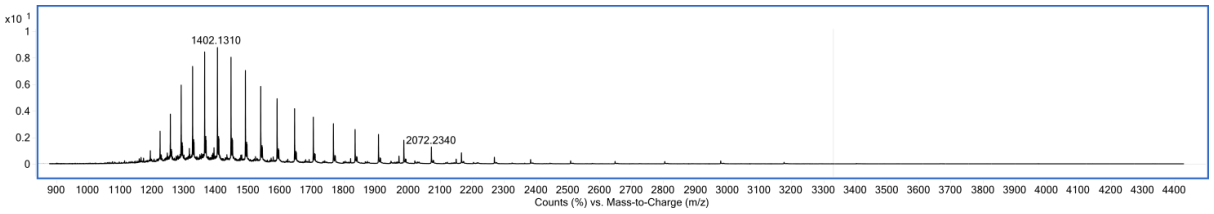
Figure S27: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47802 Da (1 addition). Observed: 47638 Da, 47802 Da.

37 °C

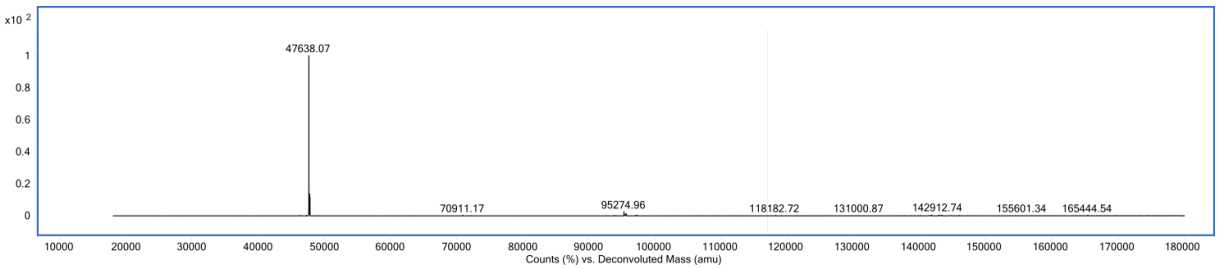
(a)



(b)



(c)



(d)

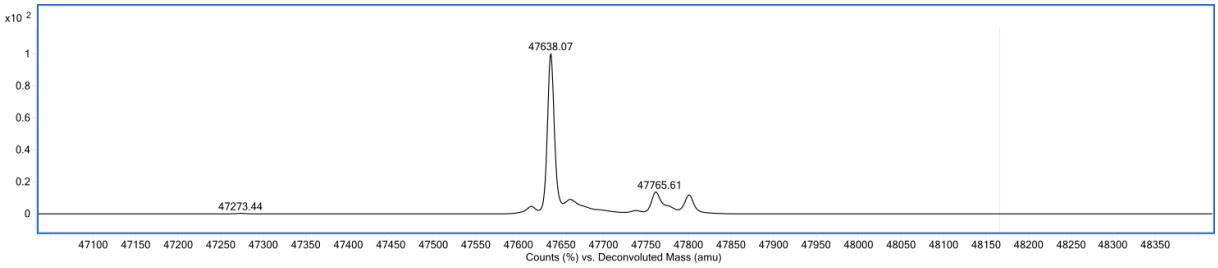
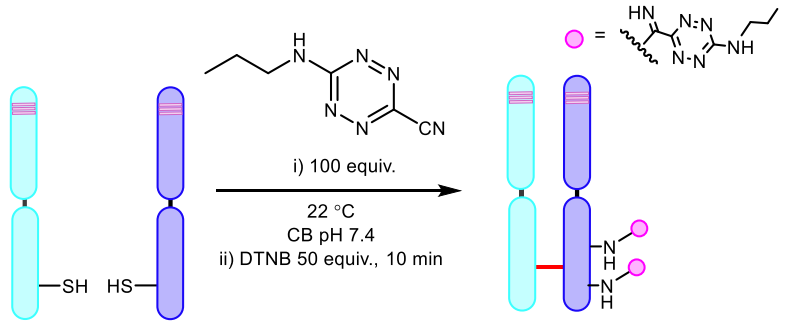


Figure S28: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47761 Da (1 addition-hydrolysed), 47802 Da (1 addition). Observed: 47638 Da, 47766 Da, 47802 Da.

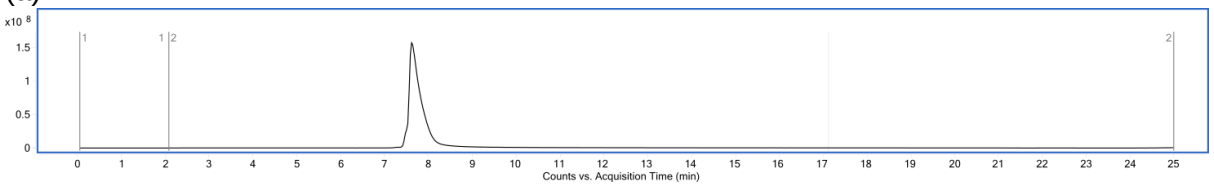
High equivalent reactions



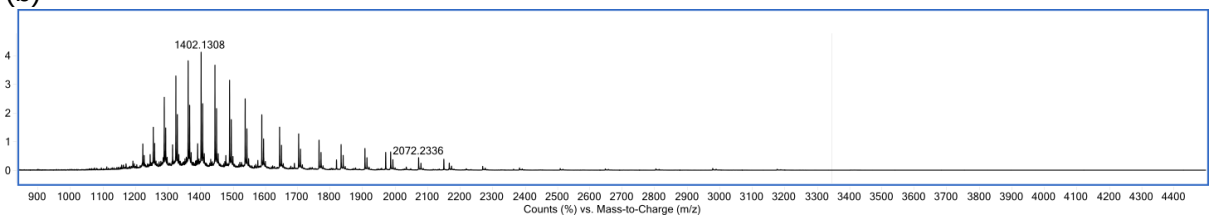
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 6-(propylamino)-1,2,4,5-tetrazine-3-carbonitrile was then added (2.0 µL, 150 mM in DMF, 100 equiv.) and incubated at 22 °. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

30 min

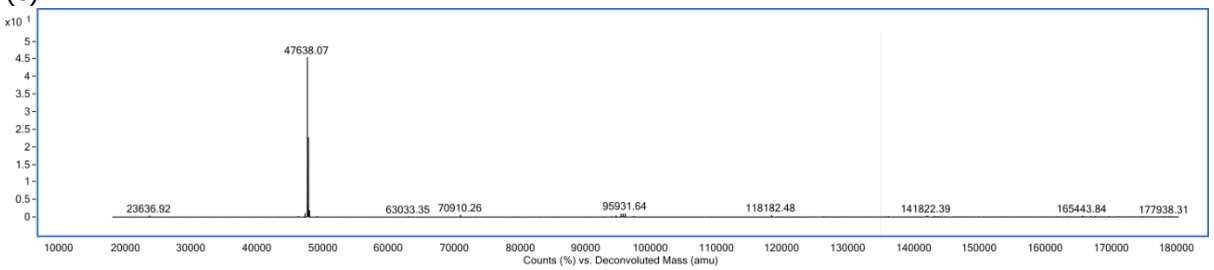
(a)



(b)



(c)



(d)

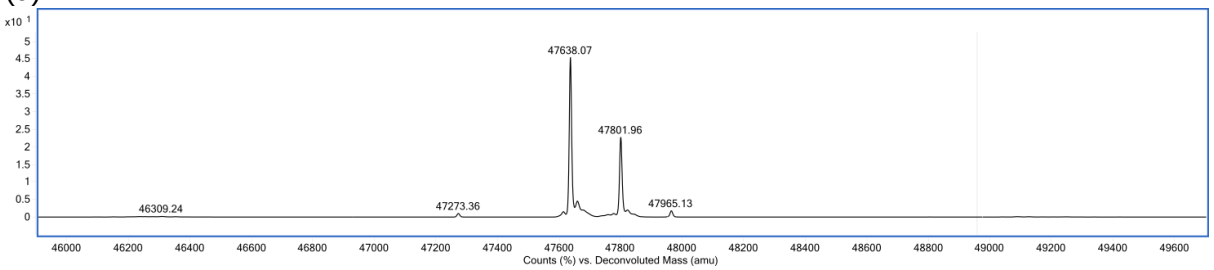
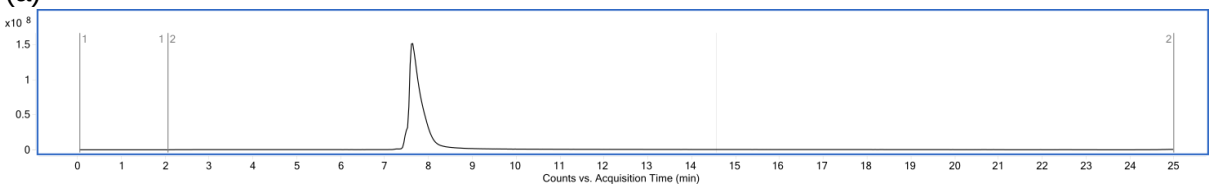


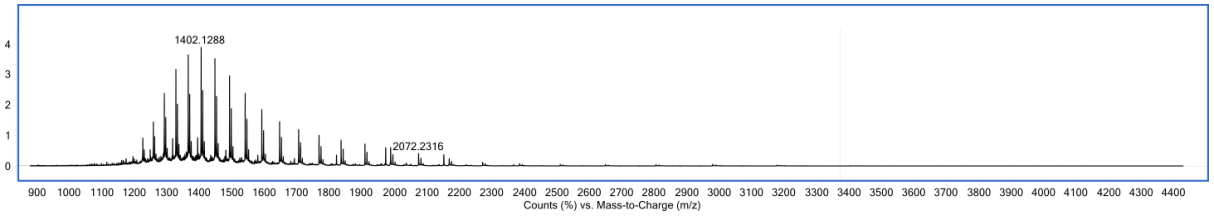
Figure S29: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47802 Da (1 addition), 47964 Da (2 additions). Observed: 47638 Da, 47802 Da, 47965 Da.

60 min

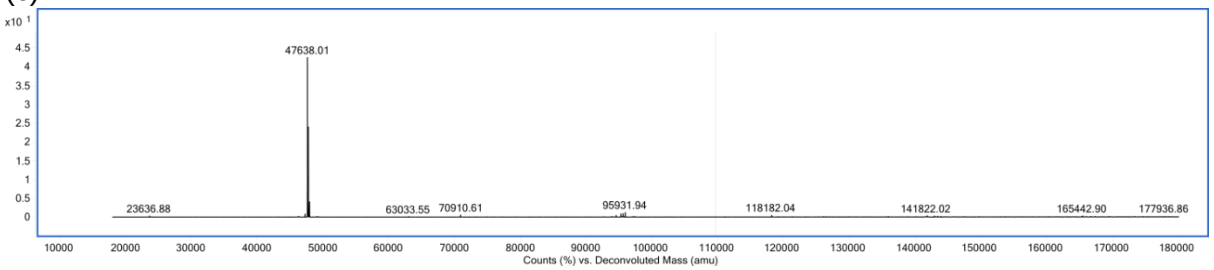
(a)



(b)



(c)



(d)

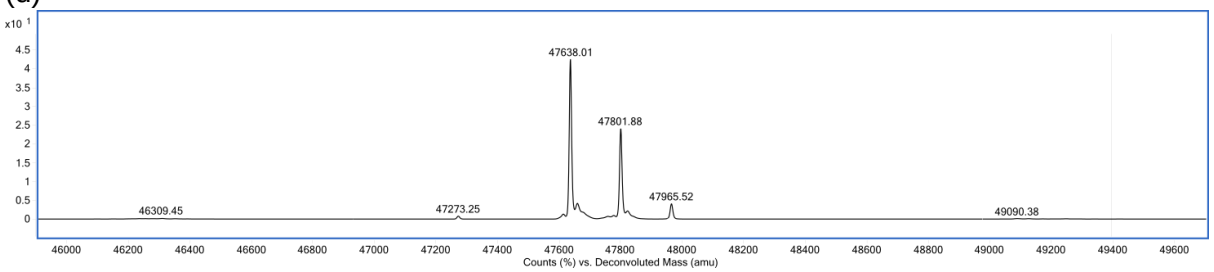
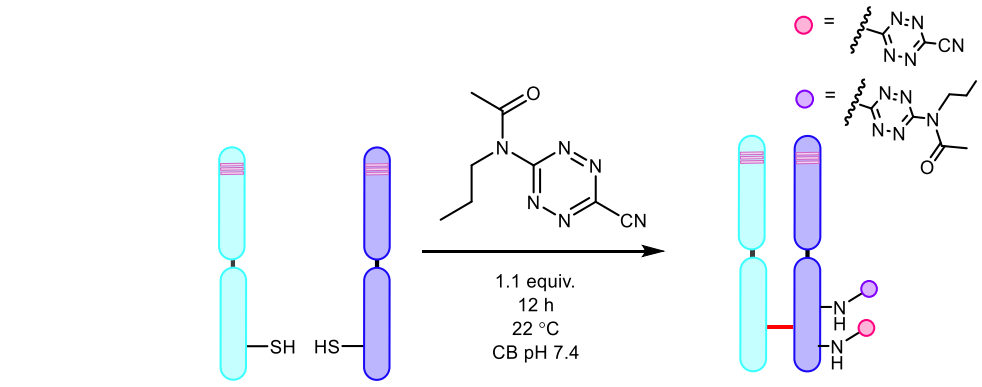


Figure S30: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47802 Da (1 addition), 47964 Da (2 additions). Observed: 47638 Da, 47802 Da, 47966 Da.

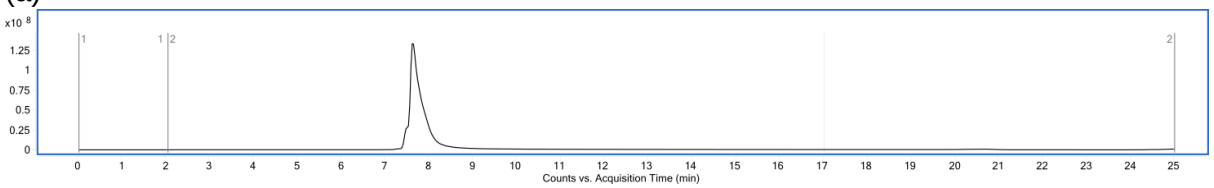
N-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylacetamide **13** reactions on Fab.

Low equivalent reactions

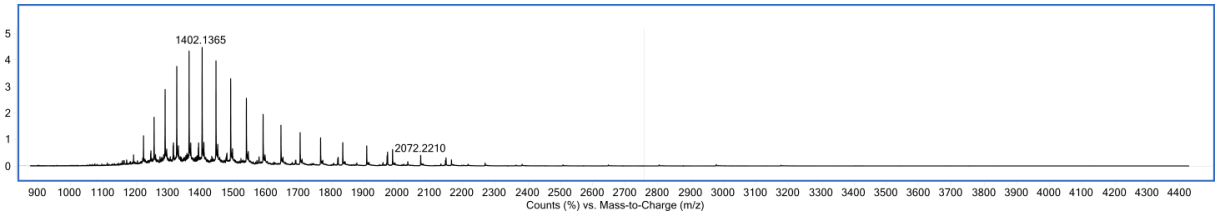


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. *N*-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylacetamide was then added (1.1 µL, 3 mM in DMF, 1.5 equiv.) and incubated for 12 h at 22 °C. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}. cm^{-1}.

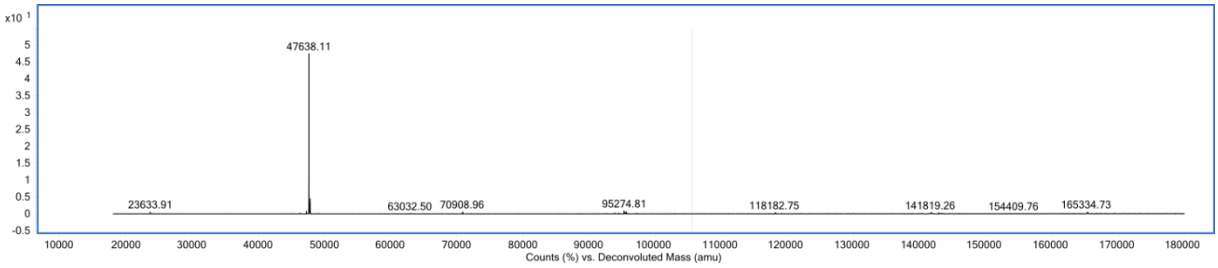
(a)



(b)



(c)



(d)

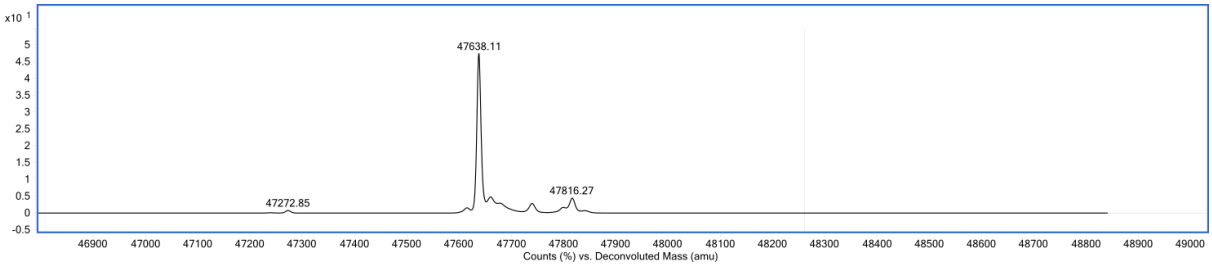
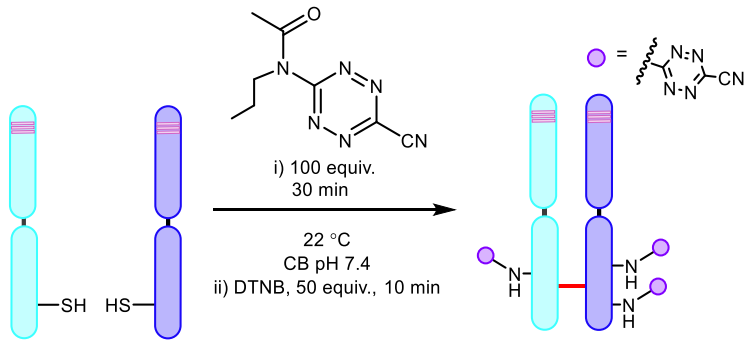


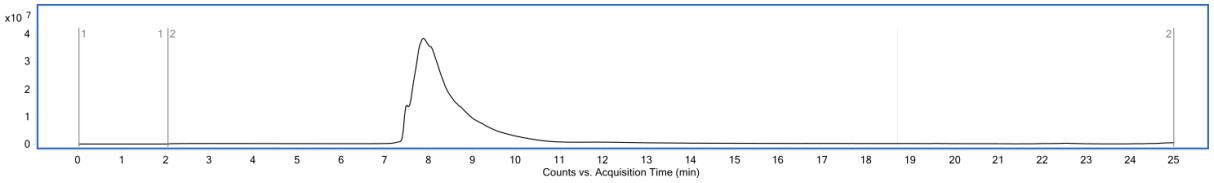
Figure 31: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47744 Da (Amide S_NAr), 47918 Da (Nitrile S_NAr). Observed: 47638 Da, 47740 Da, 47816 Da.

High equivalent reactions

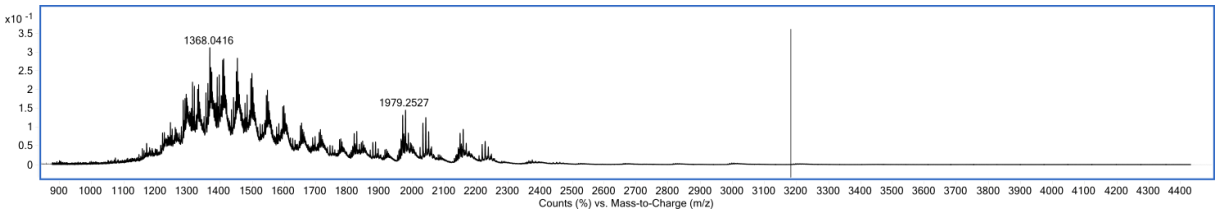


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. N-(6-cyano-1,2,4,5-tetrazin-3-yl)-N-propylacetamide was then added (2.0 µL, 150 mM in DMF, 100 equiv.) and incubated at 22 °C for 30 min. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

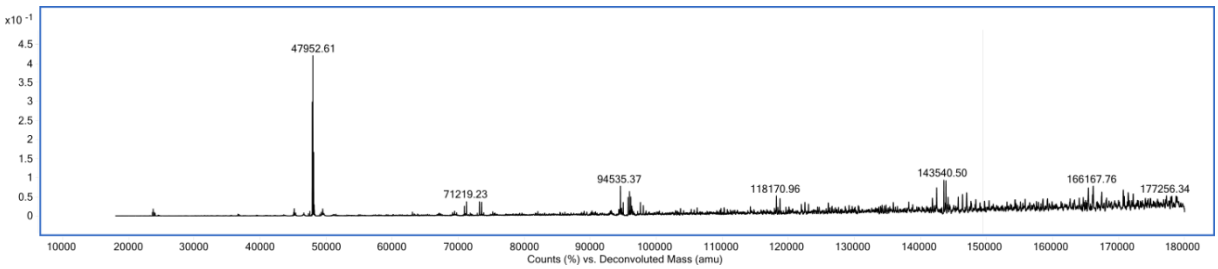
(a)



(b)



(c)



(d)

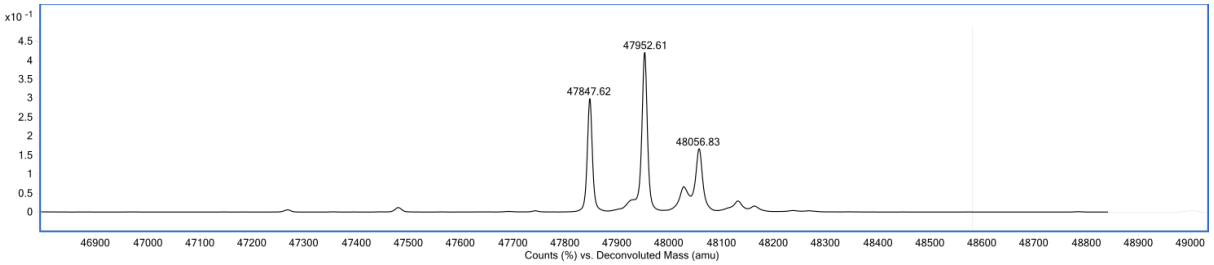
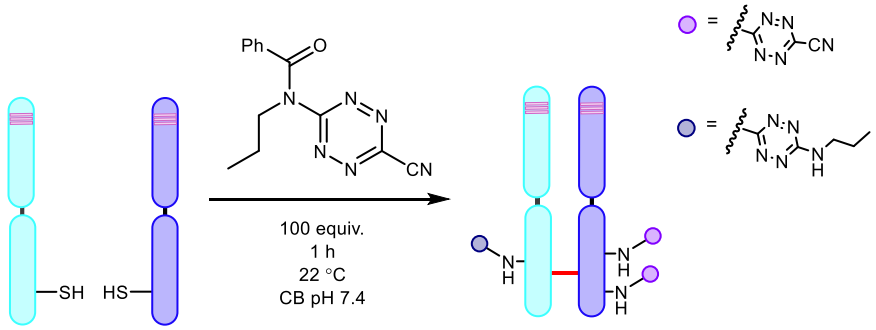


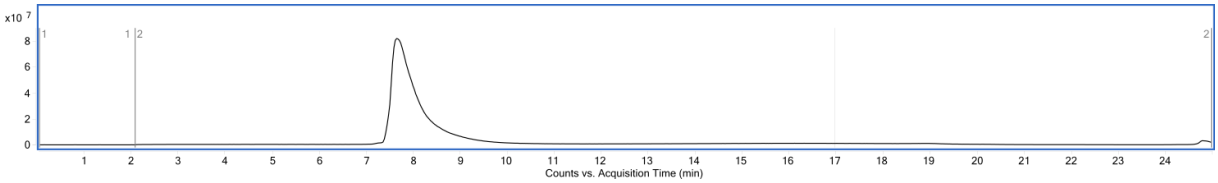
Figure 32: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47744 Da (1 S_NAr addition), 47850 Da (2 S_NAr additions), 47956 Da (3 S_NAr additions), 48062 Da (4 S_NAr additions). Observed: 47848 Da, 47953 Da, 48057 Da.

N-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylbenzamide **14** reactions on Fab.

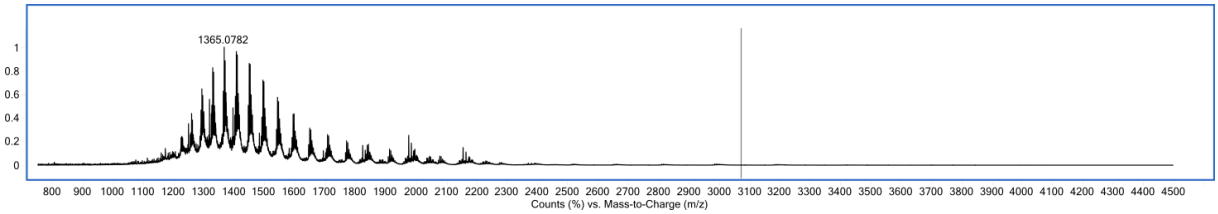


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. *N*-(6-cyano-1,2,4,5-tetrazin-3-yl)-*N*-propylbenzamide was then added (2.0 µL, 150 mM in DMF, 100 equiv.) and incubated at 22 °C for 1 h. The sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

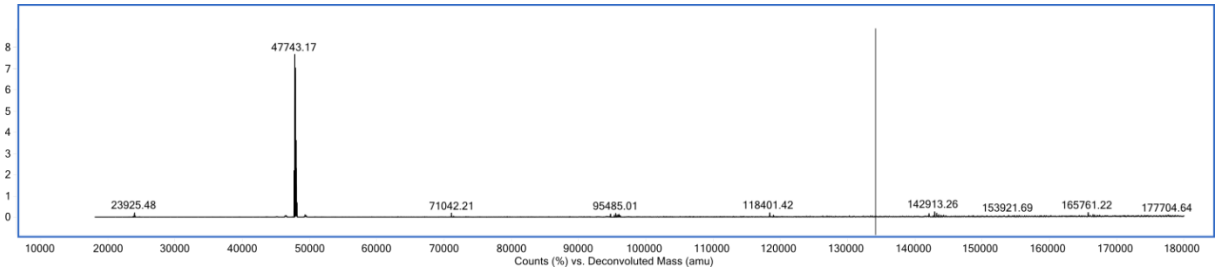
(a)



(b)



(c)



(d)

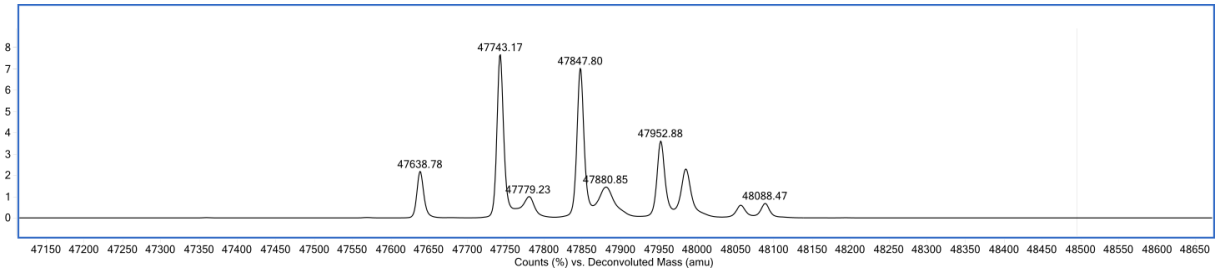
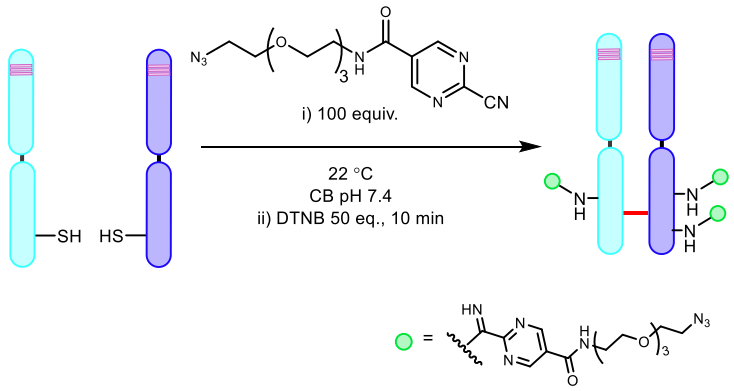


Figure 33: LCMS analysis of Fab CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47639 Da (native, 0 additions), 47745 Da (1 amide S_NAr addition), 47777 Da (1 nitrile S_NAr addition + 1 amide hydrolysis), 47850 Da (2 amide S_NAr additions), 47882 Da (1 amide S_NAr additions + 1 nitrile S_NAr addition + 1 amide hydrolysis), 47956 Da (3 amide S_NAr additions), 47989 Da (2 amide S_NAr additions + 1 nitrile S_NAr addition + 1 amide hydrolysis), 48063 Da (4 amide S_NAr additions), 48094 Da (3 amide S_NAr additions + 1 nitrile S_NAr addition + 1 amide hydrolysis). Observed: 47639 Da, 47743 Da, 47779 Da, 47848 Da, 47881 Da, 47953 Da, 48089 Da.

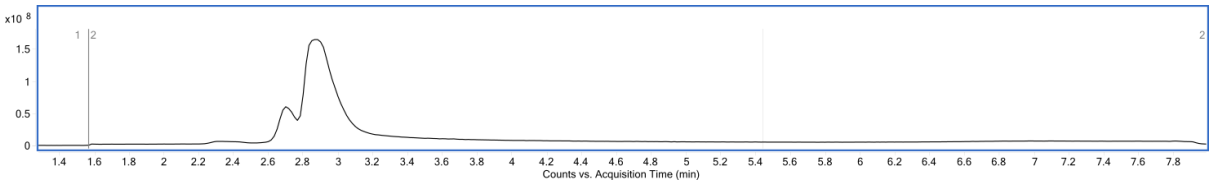
N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide **5** reactions on Fab.



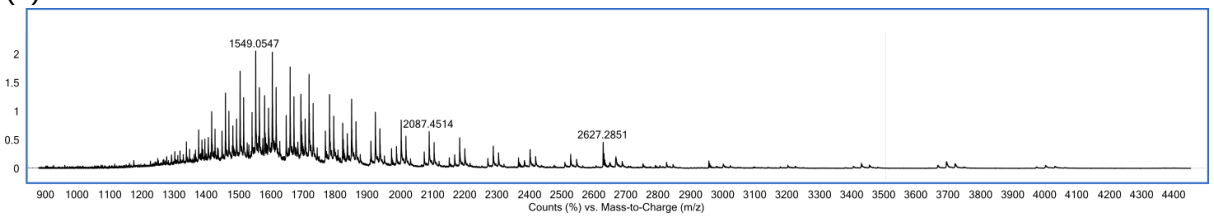
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. *N*-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide (2.0 µL, 150 mM in DMF, 100 equiv.) was added and incubated at 22 °C. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

20 min

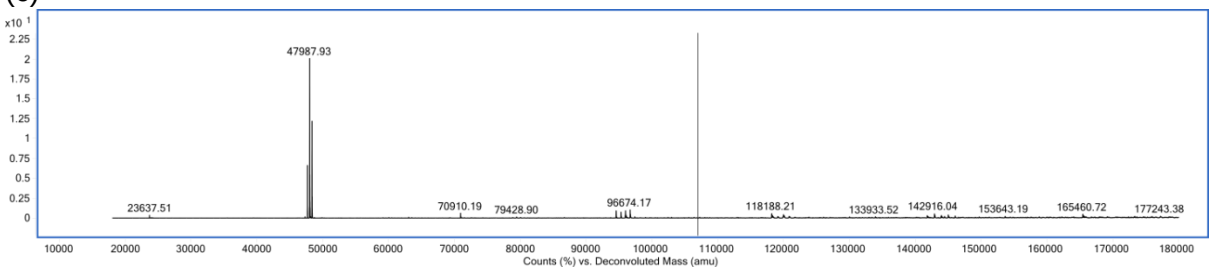
(a)



(b)



(c)



(d)

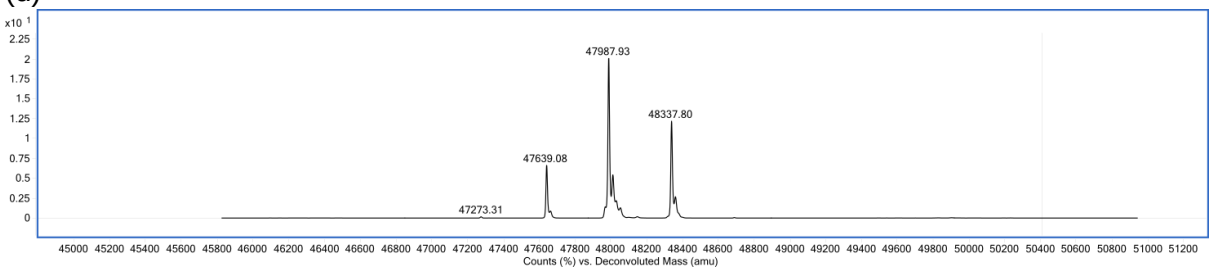
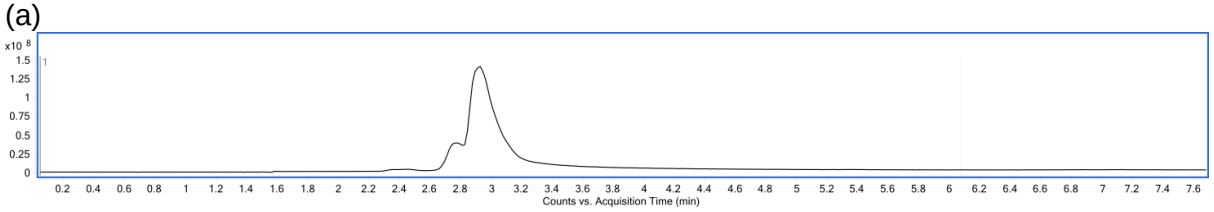


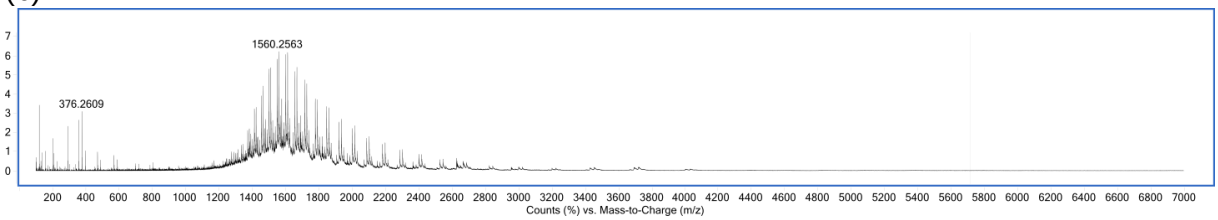
Figure S34: LCMS analysis of conjugate **19**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47639 Da (native, 0 additions), 47987 Da (1 addition), 48336 Da (2 additions). Observed: 47639 Da, 47988 Da, 48338 Da.

30 min

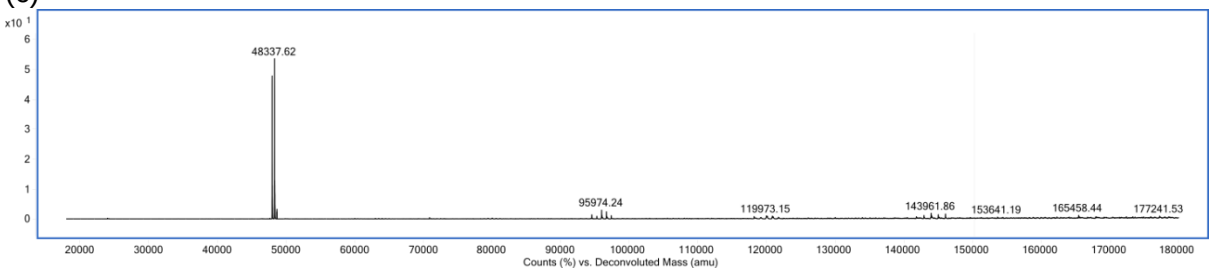
(a)



(b)



(c)



(d)

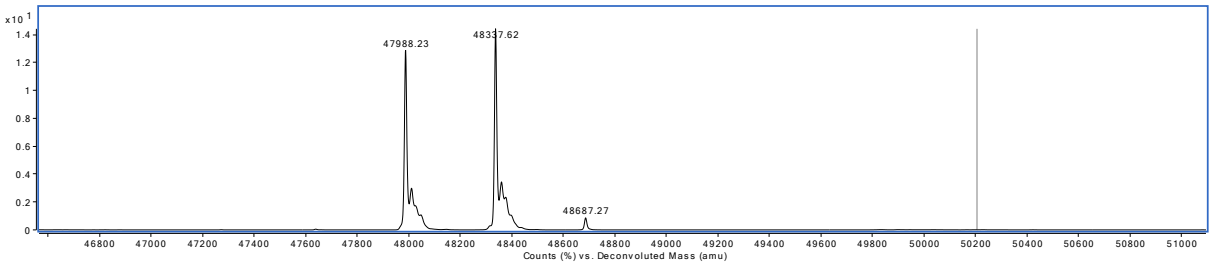
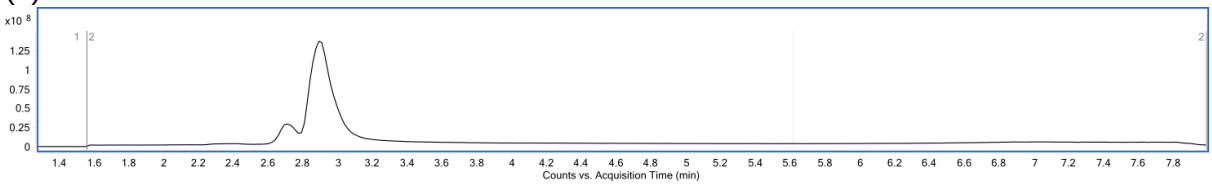


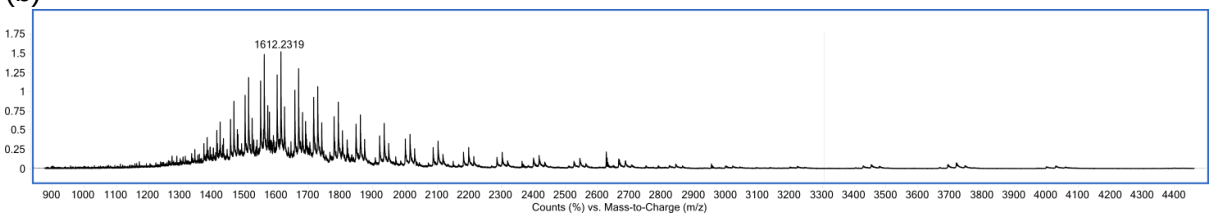
Figure S35: LCMS analysis of conjugate **19**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47987 Da (1 addition), 48336 Da (2 additions), 48686 Da (3 additions). Observed:47988 Da, 48337 Da, 48687 Da.

40 min

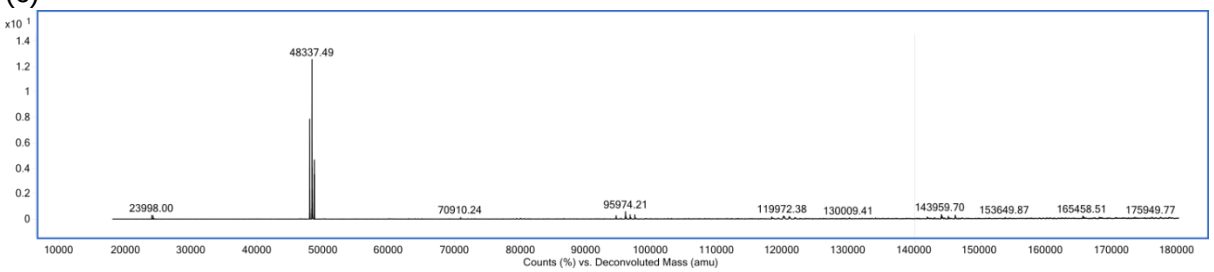
(a)



(b)



(c)



(d)

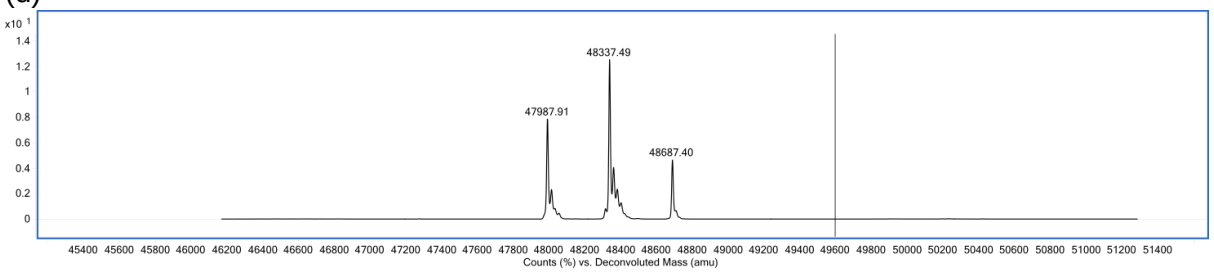
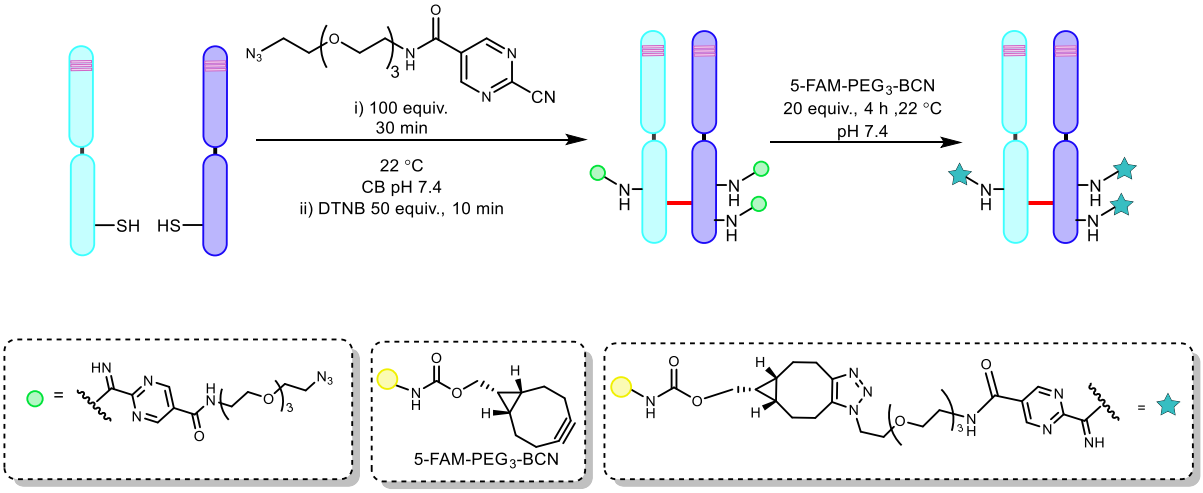


Figure S36: LCMS analysis of conjugate **19**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47987 Da (1 addition), 48336 Da (2 additions), 48686 Da (3 additions). Observed:47988 Da, 48338 Da, 48687 Da.

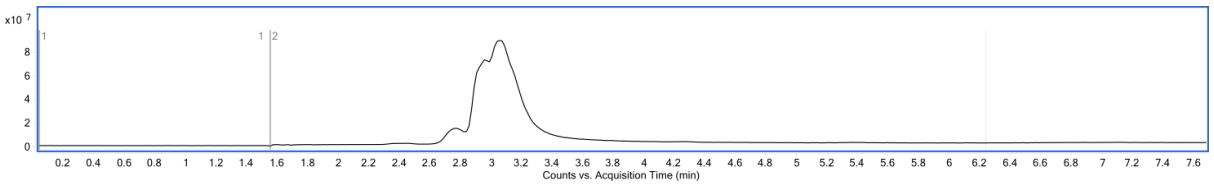
N-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide **5** reactions on Fab followed by SPAAC functionalisation.



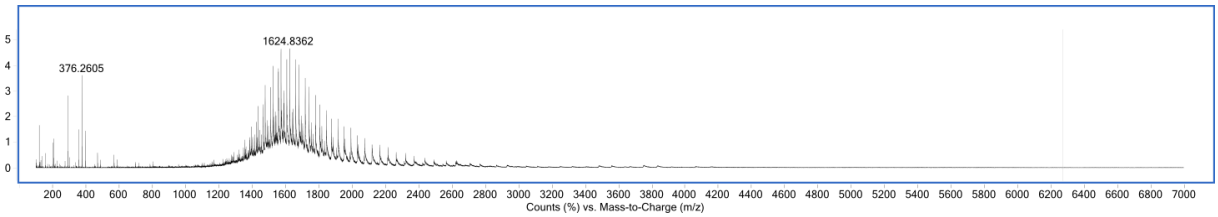
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. *N*-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-2-cyanopyrimidine-5-carboxamide (2.0 µL, 150 mM in DMF, 100 equiv.) was added and

incubated at 22 °C for 30 min. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at RT for 10 min. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into PBS (pH = 7.4). Fluorescein-PEG_3-BCN (20 mM in DMF, 20 equiv.), this was then added, and the reaction was left in the dark at 22 °C for 4 h, 300 rpm. Excess reagent was then removed by ultrafiltration (10 kDa MWCO) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

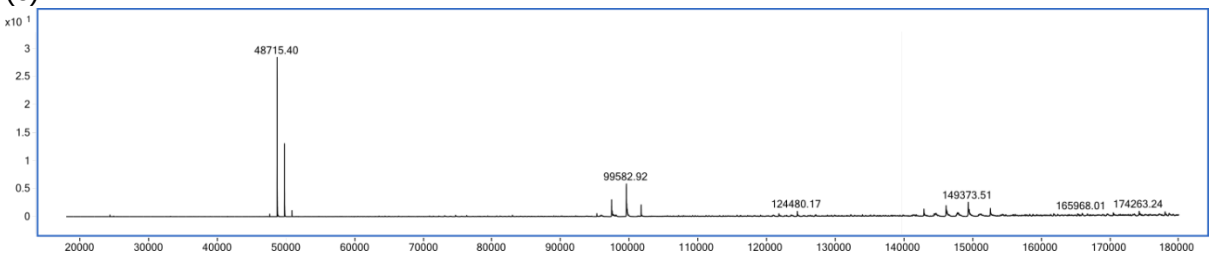
(a)



(b)



(c)



(d)

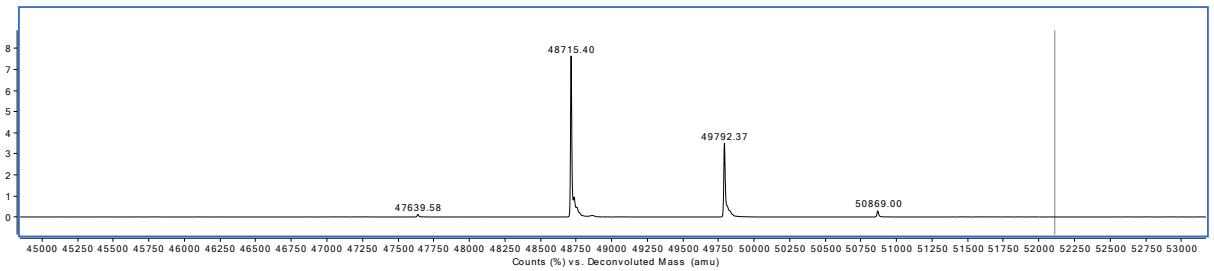
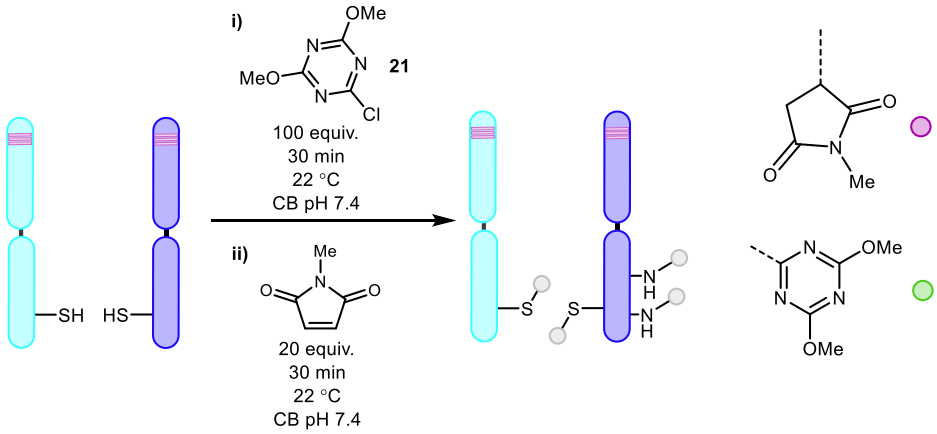


Figure S37: LCMS analysis of conjugate **20**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT-SPAAC conjugate: 47638 Da (native, 0 additions), 48714 Da (1 addition), 49790 Da (2 additions), 50866 Da (3 additions). Observed: 47639 Da, 48715 Da, 49792 Da, 50869 Da. Peaks round 99582 (2 x 49791) are proposed to be artifacts formed during ionisation. The fluorophore-to-antibody ratio (FAR) was determined photometrically by

obtaining the UV/Vis absorption reading of the conjugate **20** and using the following formula, where *Cf* is the correction factor for the absorbance of the fluorescein-PEG$_3$-BCN at 280 nm:

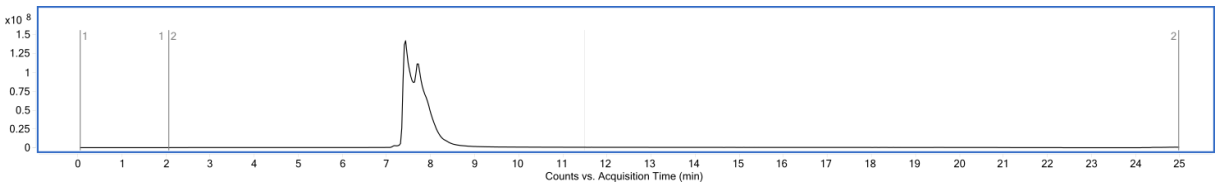
$$FAR = \frac{\frac{Abs_{505}}{\varepsilon_{505}}}{\frac{Abs_{280} - (Cf\ x\ Abs_{505})}{\varepsilon_{280}}} = \frac{\frac{0.434}{83000}}{\frac{0.320 - (0.178\ x\ 0.434)}{68590}} = 1.5$$

2-Chloro-4,6-dimethoxy-1,3,5-triazine **21** reactions on Fab.

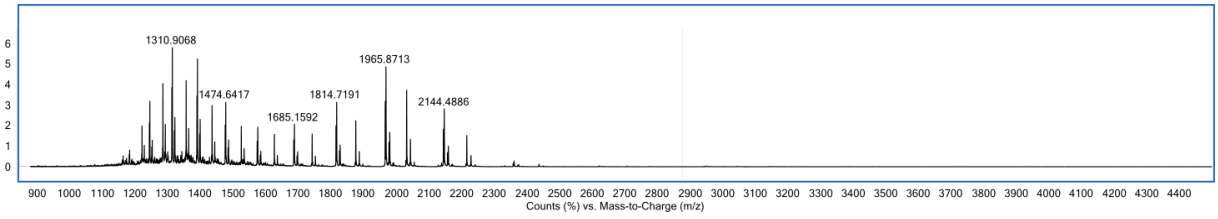


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH$_2$O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 2-chloro-4,6-dimethoxy-1,3,5-triazine (2.0 µL, 150 mM in DMF, 100 equiv.) was added and incubated at 22 °C for 30 min. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500). Following this NMM (20 mM in DMF, 20 equiv.) was added, and the reaction was left at 22°C for 30 min. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε$_{280}$= 68590 M^{-1} cm^{-1}.

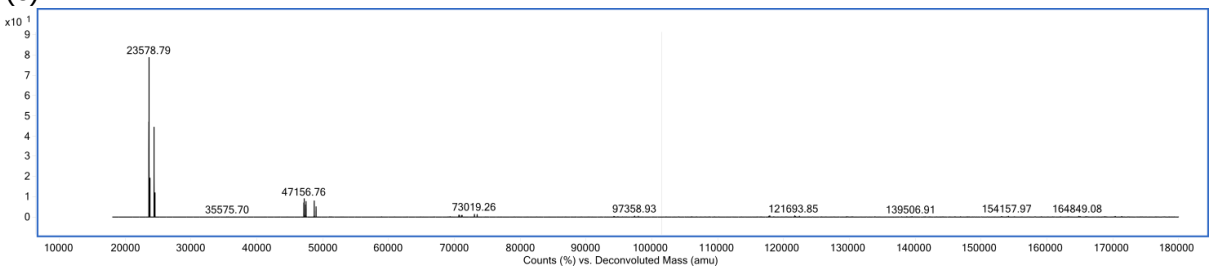
(a)



(b)



(c)



(d)

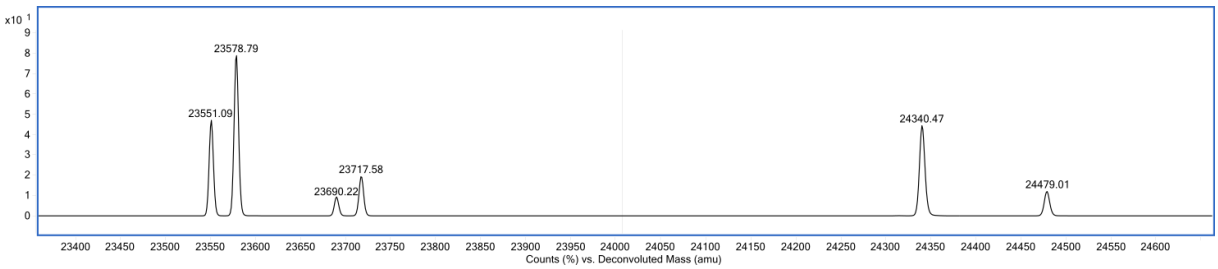
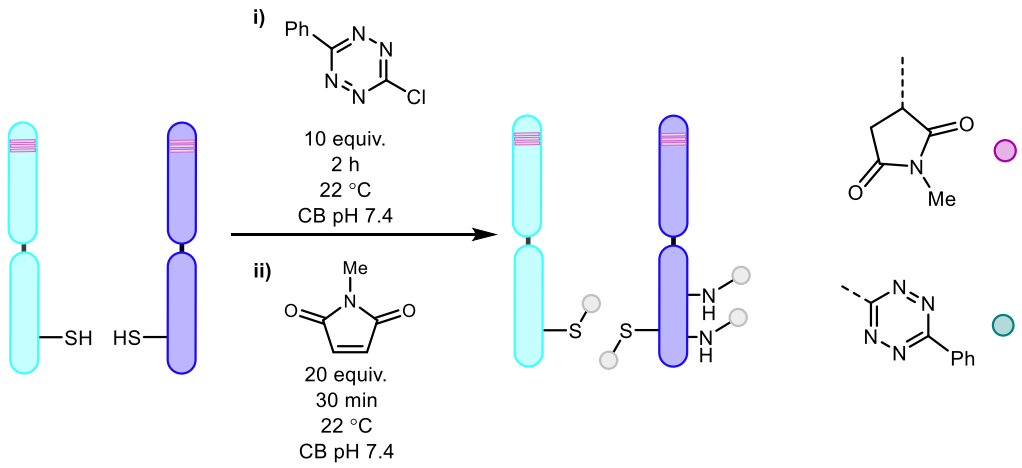


Figure S38: LCMS analysis of Fab conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of S_NAr conjugate: 23550 Da (NMM capped **LC**), 23579 Da (Triazine capped **LC**), 23690 Da (NMM capped **LC** with 1 addition of triazine), 23719 Da (Triazine capped **LC** with 1 addition of triazine), 24340 Da (Triazine capped **HC**), 24480 Da (Triazine capped **HC** with 1 addition of triazine); Observed: **LC** 23551 Da, 23579 Da, 23690 Da, 23718; **HC** 24341 Da, 24479 Da.

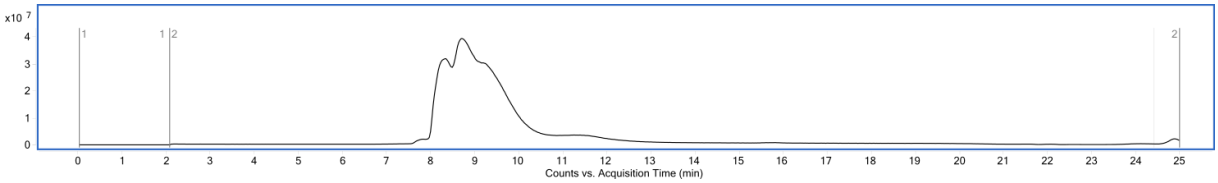
3-Chloro-6-phenyl-1,2,4,5-tetrazine **22** reactions on Fab.



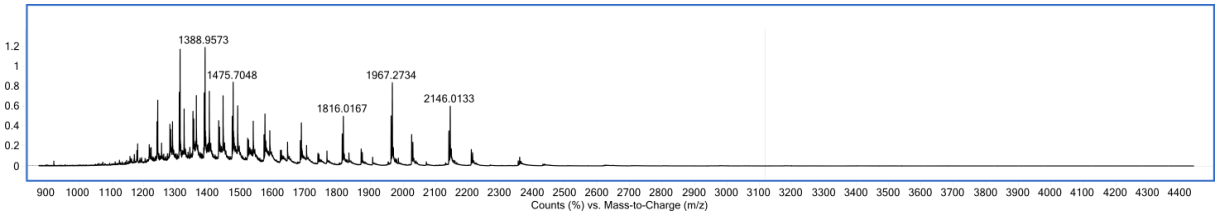
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 3-chloro-6-phenyl-1,2,4,5-tetrazine (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2 h. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500). Following this NMM (20 mM in DMF, 20

equiv.) was added, and the reaction was left at RT for 30 min. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

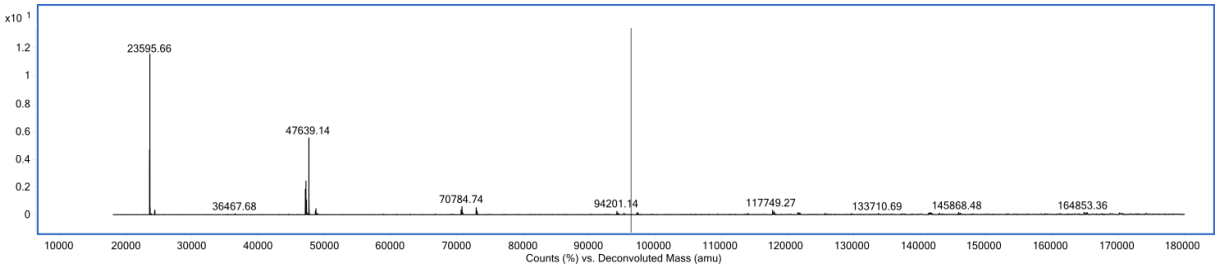
(a)



(b)



(c)



(d)

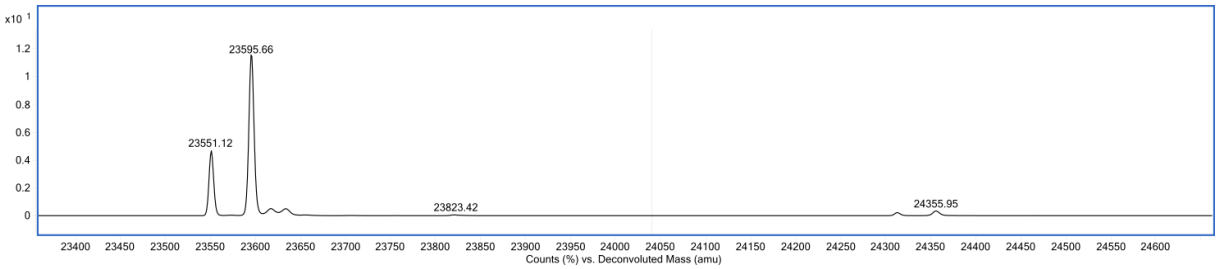
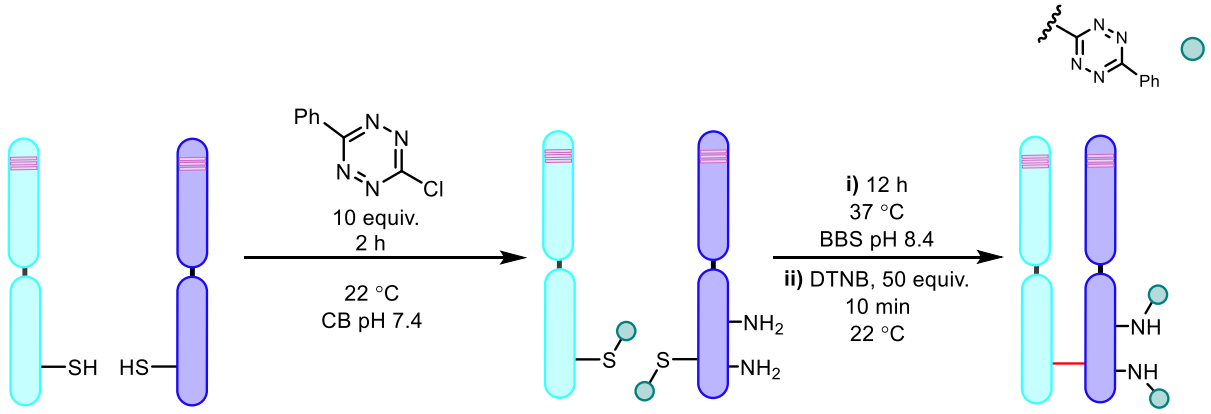


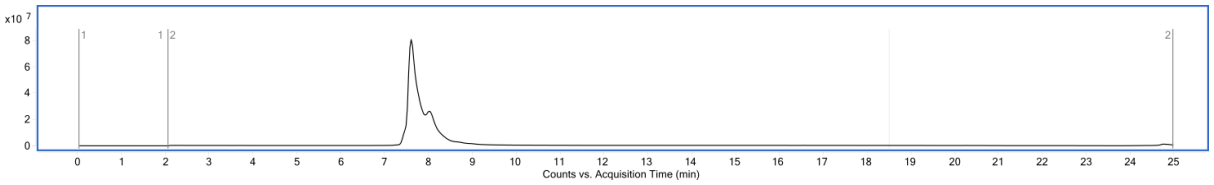
Figure S39: LCMS analysis of Fab conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of S_NAr conjugate: 23550 Da (NMM capped **LC**), 23596 Da (Tetrazine capped **LC**), 24311 Da (NMM capped **LC**), 24357 Da (Triazine capped HC**);** Observed: **LC** 23551 Da, 23596 Da, 23690 Da, 23718 Da; **HC** 24356 Da.

3-Chloro-6-phenyl-1,2,4,5-tetrazine **22** reactions on Fab with transfer.

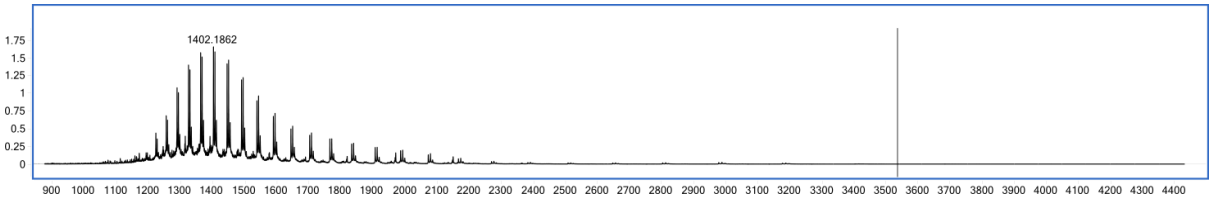


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 3-chloro-6-phenyl-1,2,4,5-tetrazine (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2 h. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into BBS, pH 8.4, the reaction was then incubated at 37 °C for 12 h. Following this DTNB (1.0 µL, 150 mM in EtOH, 50 equiv.) was added and the reaction was left at 22 °C for 10 min. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

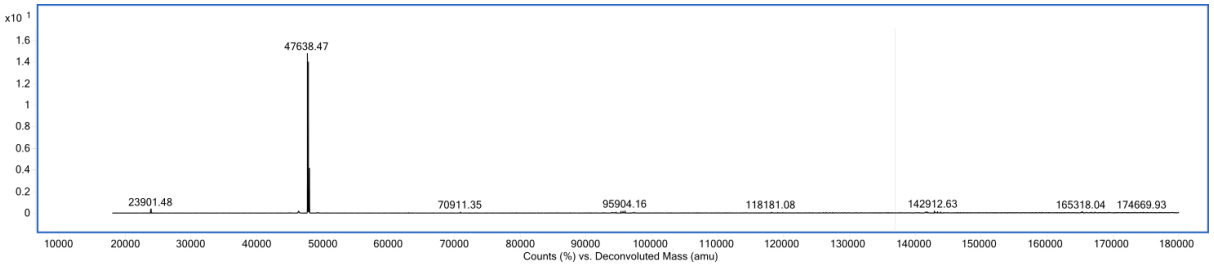
(a)



(b)



(c)



(d)

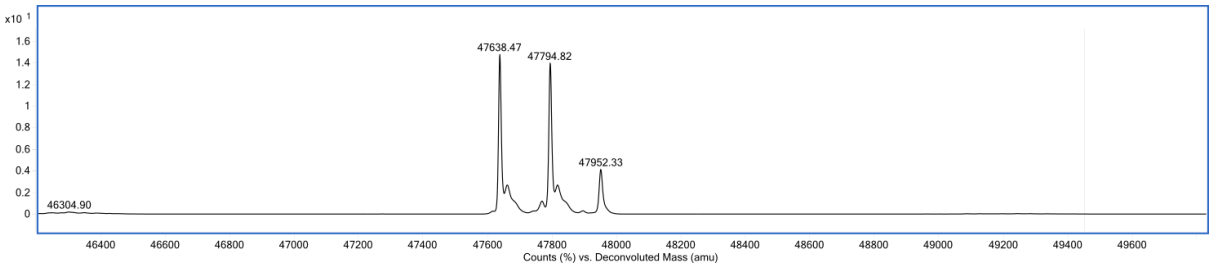
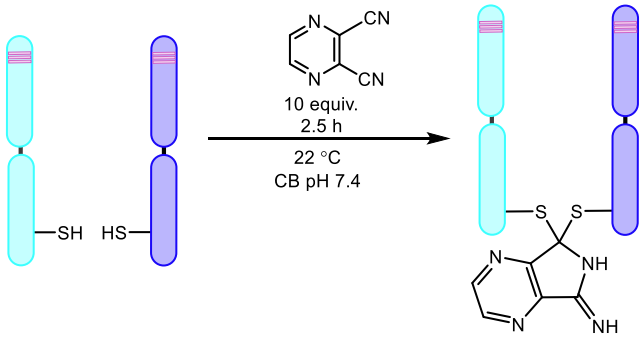


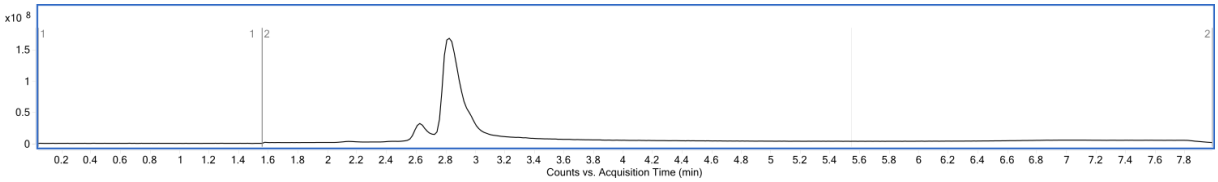
Figure S40: LCMS analysis of Fab S_NAr CLT conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of CLT conjugate: 47638 Da (native, 0 additions), 47795 Da (1 addition), 47952 Da (2 additions). Observed: 47639 Da, 47795 Da, 47952 Da.

Reaction of Fab with pyrazine-2,3-dicarbonitrile

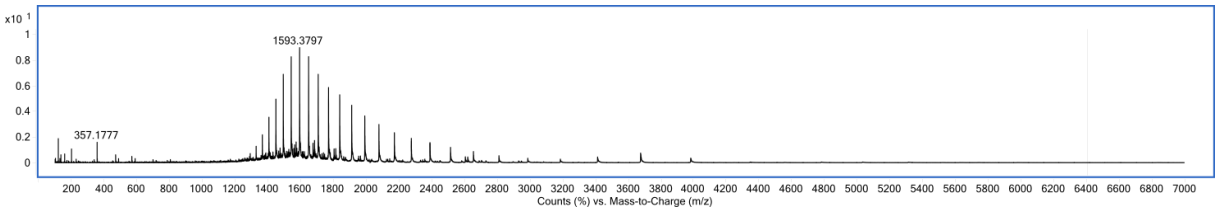


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Pyrazine-2,3-dicarbonitrile **3** (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2.5 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

(a



(b)



(c)

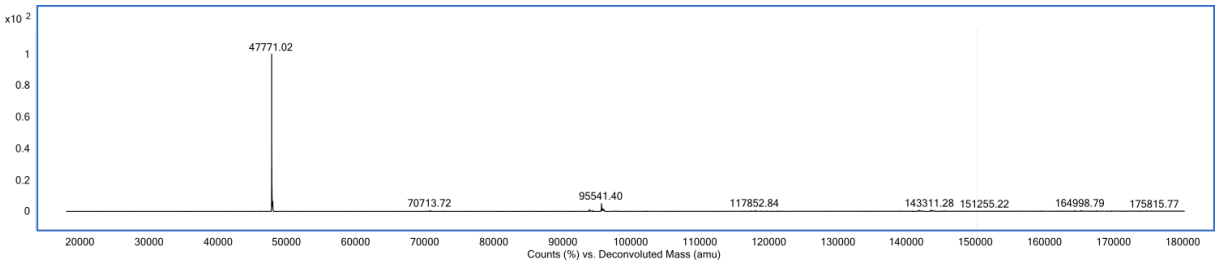
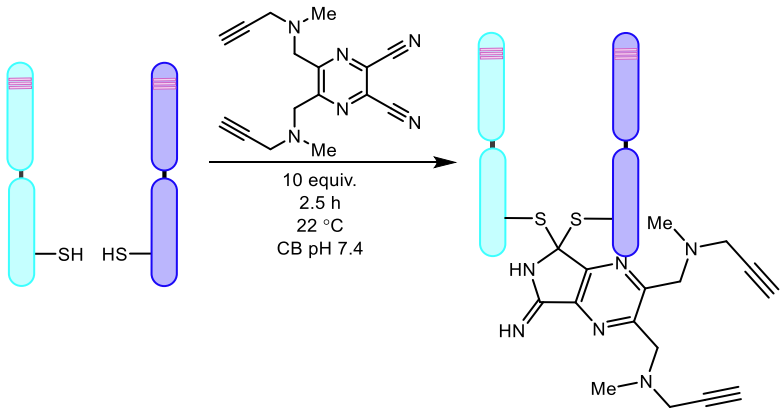


Figure S41: LCMS analysis of conjugate **27**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with pyrazine-2,3-dicarbonitrile: 47768 Da. Observed: 47771 Da.

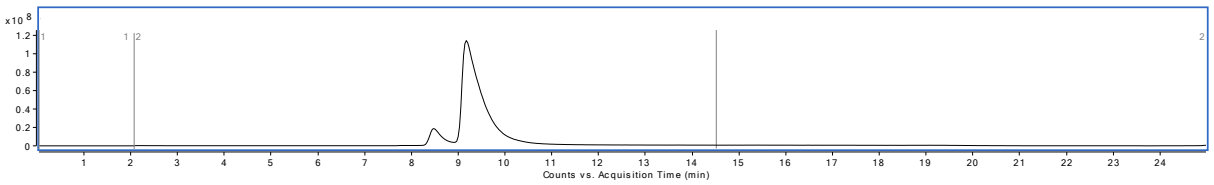
Reaction of Fab with 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile.



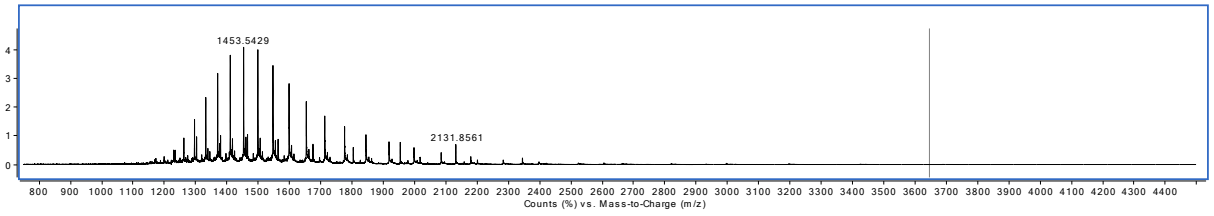
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5,6-Bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-

2,3-dicarbonitrile **9** (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

(a)



(b)



(c)

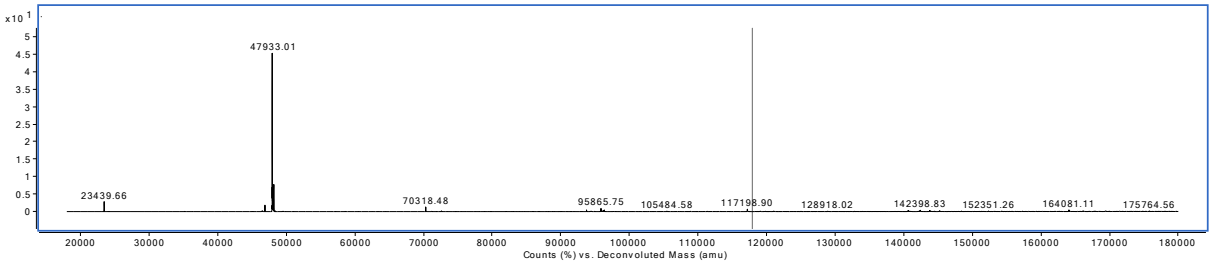
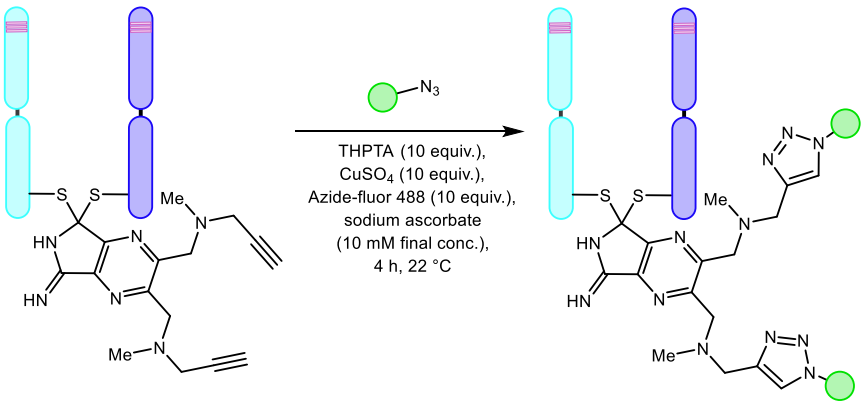


Figure S42: LCMS analysis of conjugate **39**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile: 47930 Da. Observed: 47933 Da.

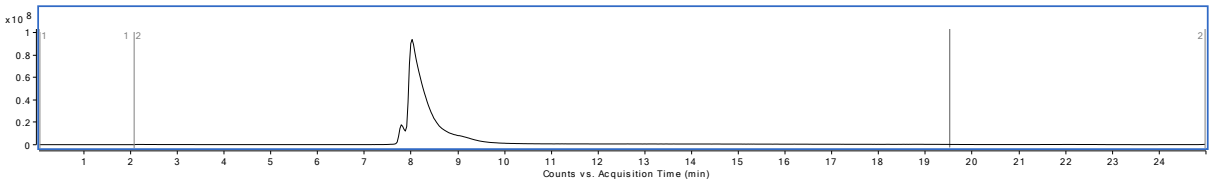
Reaction of Fab with 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile with CuAAC attempts.

Azide fluor-488

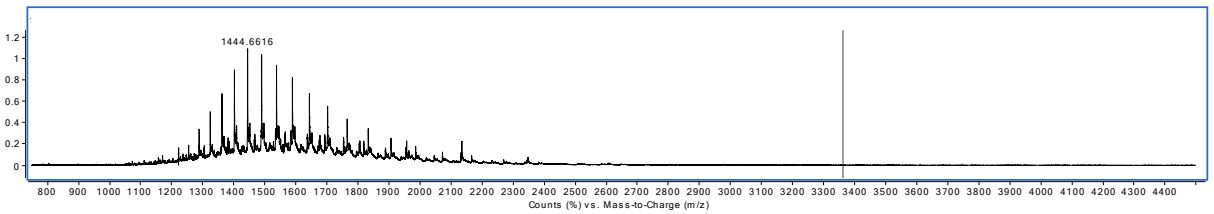


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5,6-Bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile **9** (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2 h. Excess reagent was removed *via* ultrafiltration (7 kDa MWCO, ZebaSpin) into PB (pH 7.0). To conjugate S9 was added THPTA (100 mM in dH_2O, 10 equiv.), $CuSO_4$ (20 mM in dH_2O, 10 equiv.), Azide-fluor 488 (10 mM in DMF, 10 equiv.), and sodium ascorbate (100 mM in dH_2O, final conc. 10 mM), the reaction was left at 22 °C, 300 rpm for 4 h. The excess reagent was then removed using a desalting column (PD Minitrap G-25, GE Healthcare) followed by ultrafiltration (10 kDa MWCO) into conjugate buffer to concentrate the sample. Lastly, the sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

(a)



(b)



(c)

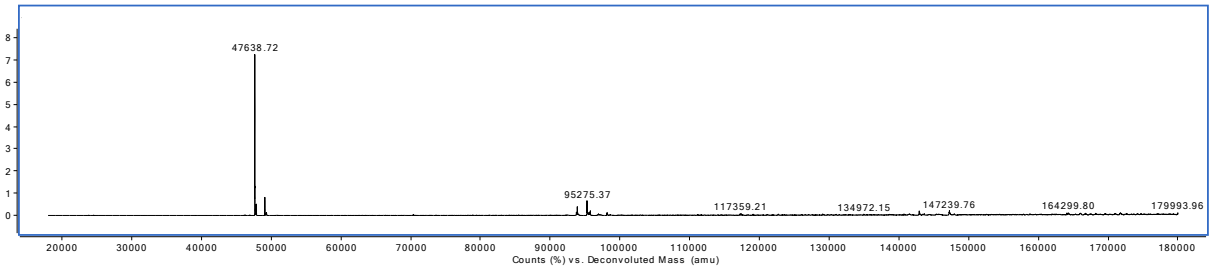
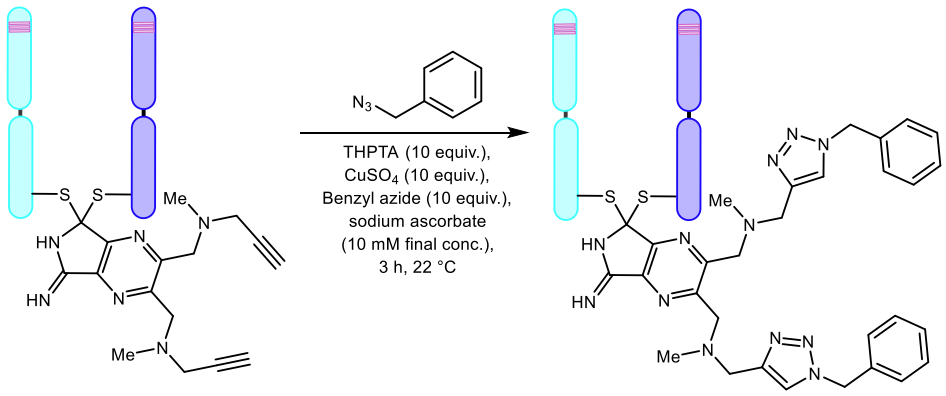


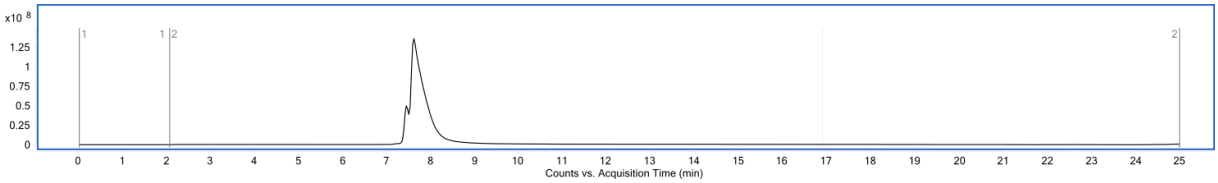
Figure S43: LCMS analysis of Fab conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile after CuAAC: 49079 Da. Observed: 47638 Da, 49083 Da.

Benzyl azide

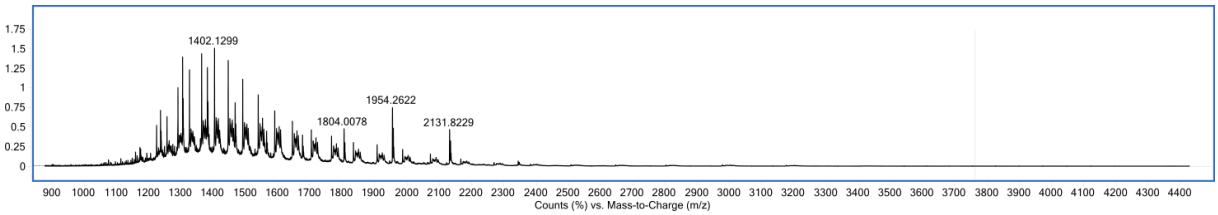


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5,6-Bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile **9** (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2 h. Excess reagent was removed *via* ultrafiltration (7 kDa MWCO, ZebaSpin) into PB (pH 7.0). To conjugate **S9** was added THPTA (100 mM in dH_2O, 10 equiv.), $CuSO_4$ (20 mM in dH_2O, 10 equiv.), benzyl azide (10 mM in DMF, 10 equiv.), and sodium ascorbate (100 mM in dH_2O, final conc. 10 mM), the reaction was left at 22 °C, 300 rpm for 3 h. The excess reagent was then removed using a desalting column (PD Minitrap G-25, GE Healthcare) followed by ultrafiltration (10 kDa MWCO) into conjugate buffer to concentrate the sample. Lastly, the sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

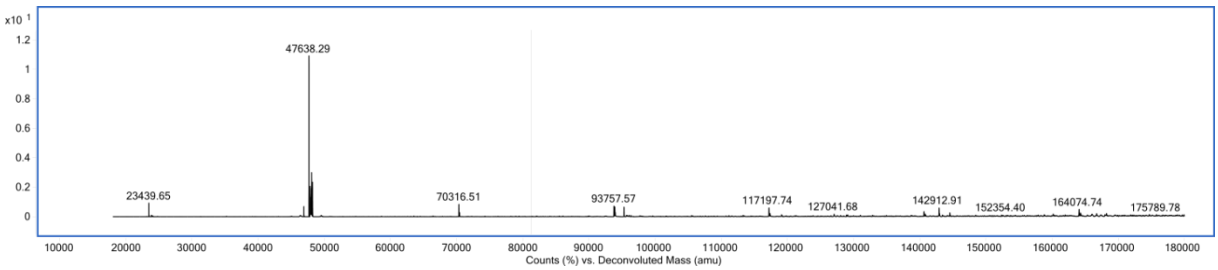
(a)



(b)



(c)



(d)

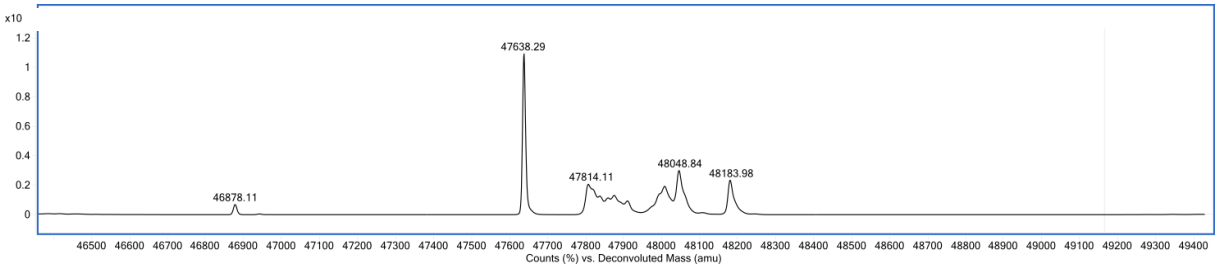
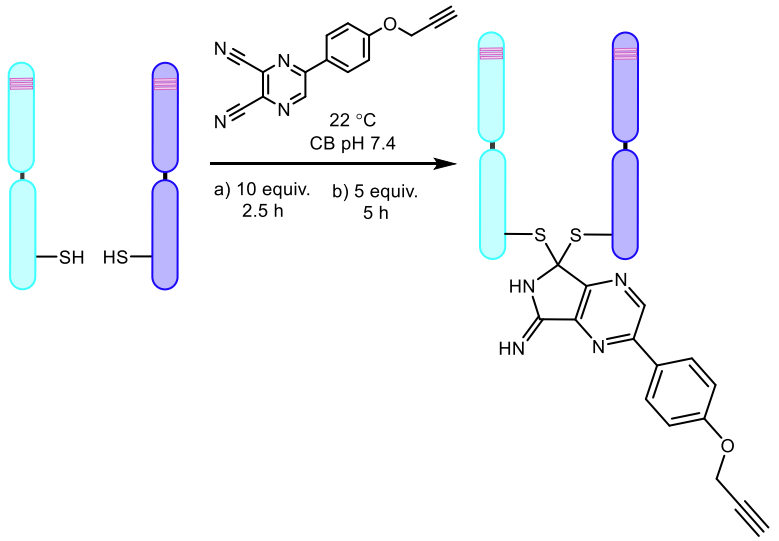


Figure S44: LCMS analysis of Fab conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 5,6-bis((methyl(prop-2-yn-1-yl)amino)methyl)pyrazine-2,3-dicarbonitrile after benzyl azide: 48063 Da (1 clicked species), 48196 Da (2 clicked species). Observed: 47638 Da, 48049 Da, 48184 Da.

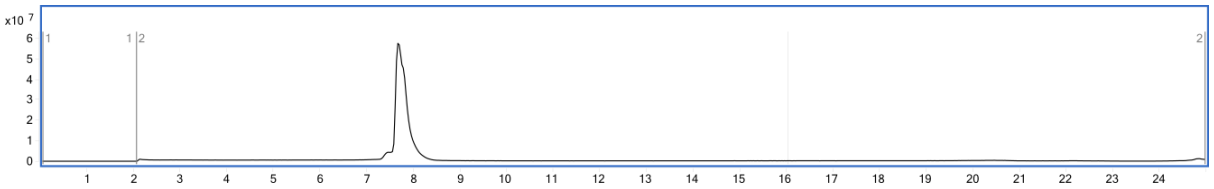
Reaction of Fab with 5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile



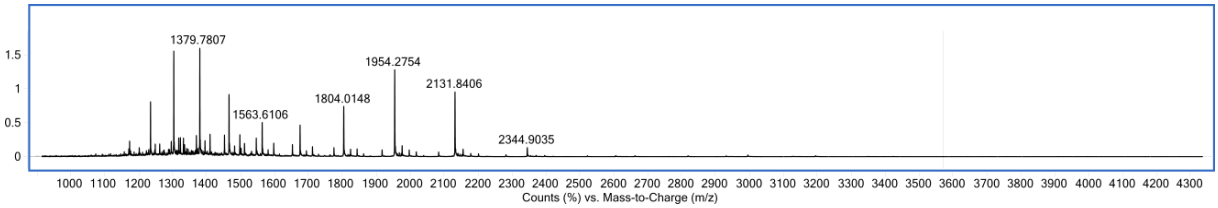
5 equiv.

Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile (1.0 µL, 15 mM in DMF, 5 equiv.) was added and incubated at 22 °C for 5 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

(a)



(b)



(c)

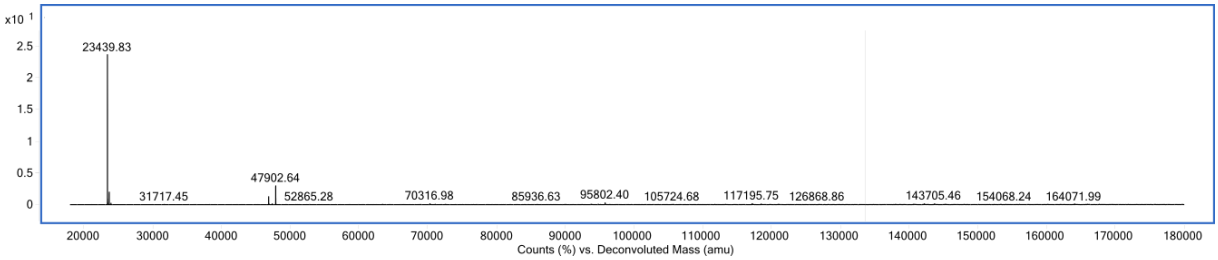
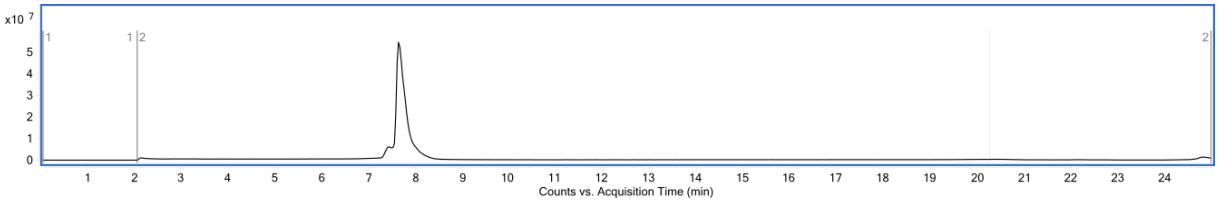


Figure S45: LCMS analysis of Fab conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile: 47898 Da. Observed: 23439 Da, 47902 Da.

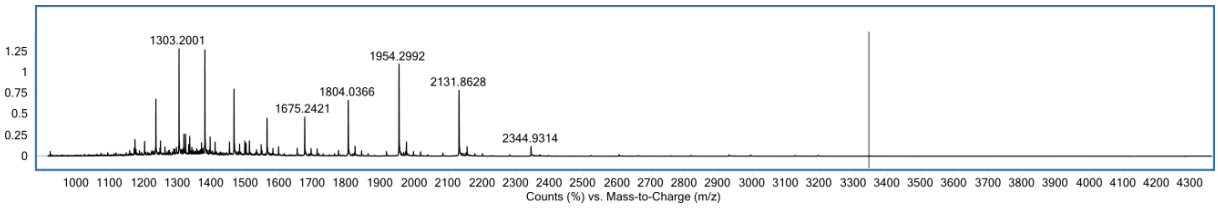
10 equiv.

Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2.5 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

(a)



(b)



(c)

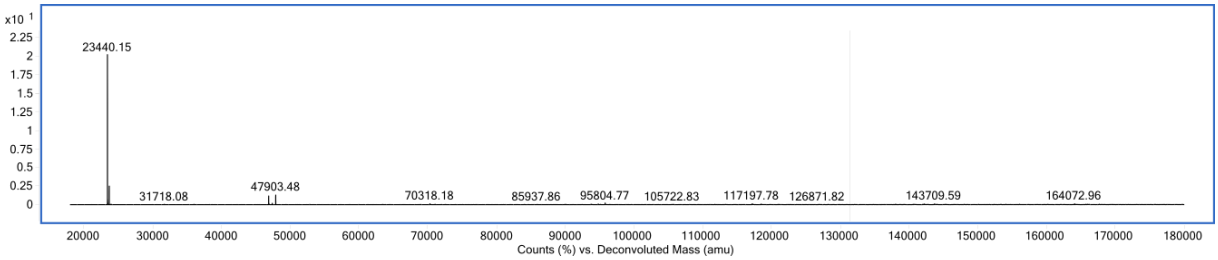
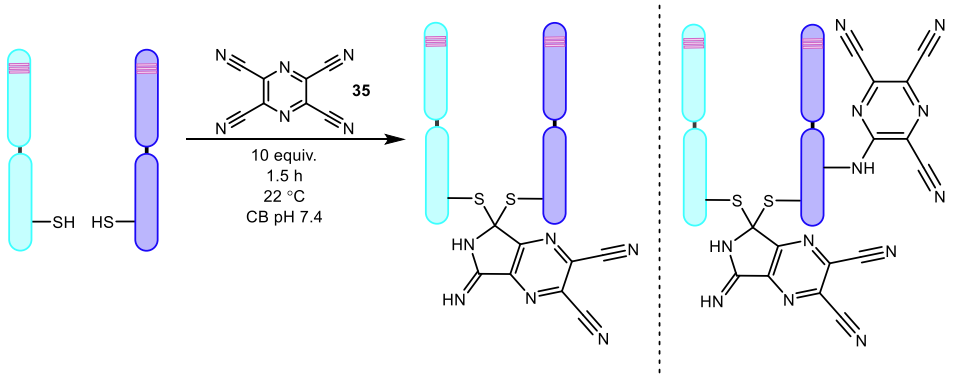


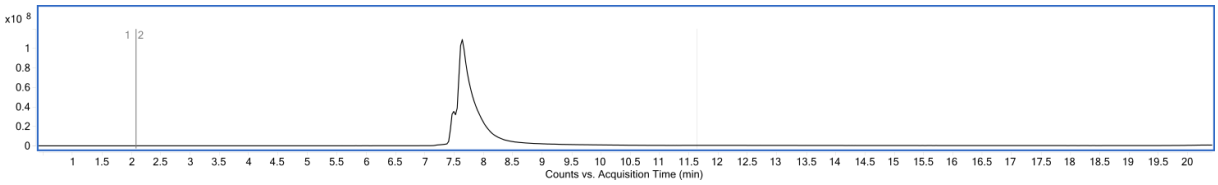
Figure S46: LCMS analysis of Fab conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 5-(4-(prop-2-yn-1-yloxy)phenyl)pyrazine-2,3-dicarbonitrile: 47898. Observed: 23440, 47904.

Reaction of Fab with pyrazine-2,3,5,6-tetracarbonitrile

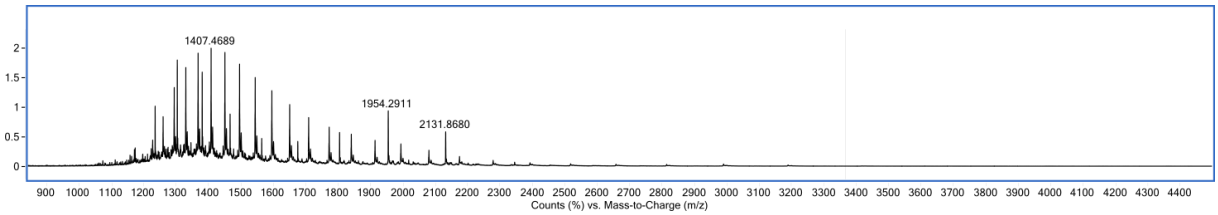


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Pyrazine-2,3,5,6-tetracarbonitrile (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 1.5 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

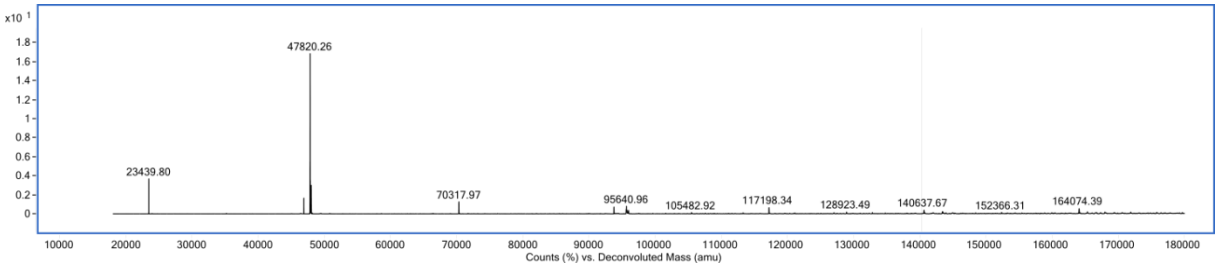
(a)



(b)



(c)



(d)

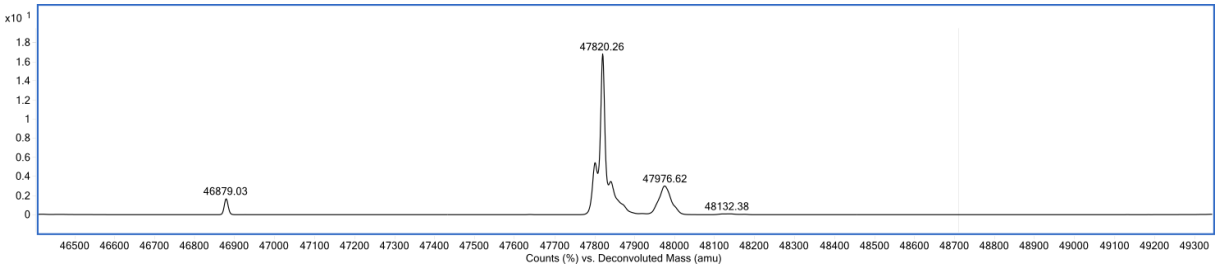
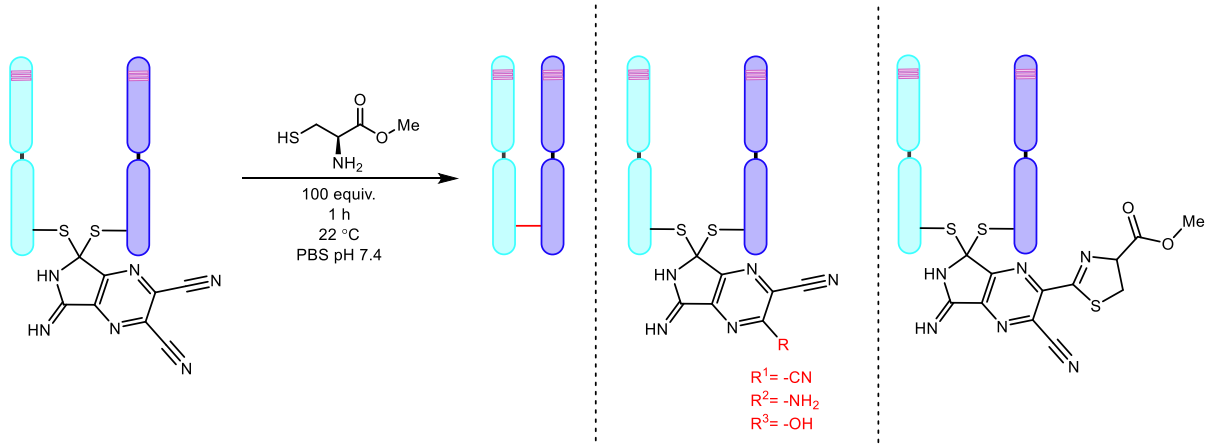


Figure S47: LCMS analysis of conjugate **40**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with pyrazine-2,3,5,6-tetracarbonitrile: 47819 Da, S_NAr side product: 47974 Da. Observed: 23439 Da, 47820 Da, 47977 Da.

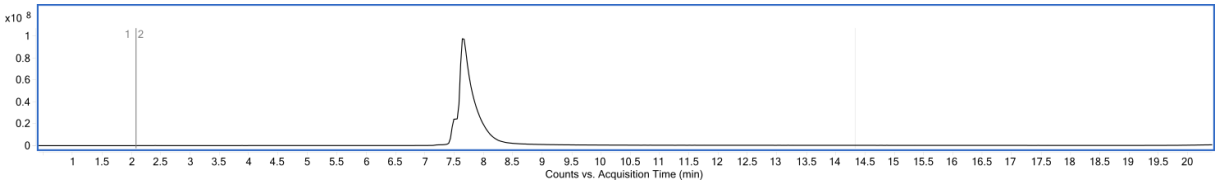
Reaction of Fab with pyrazine-2,3,5,6-tetracarbonitrile with NAT attempt



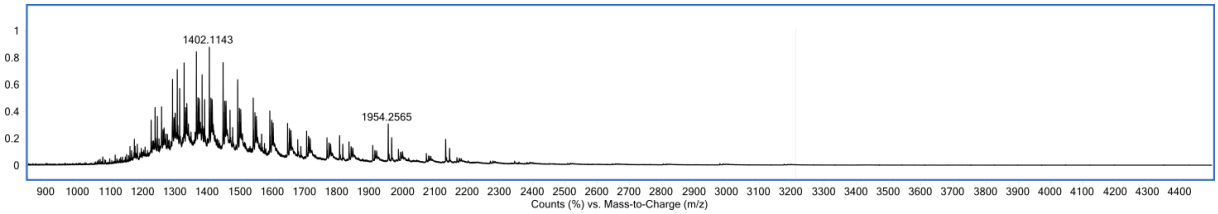
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Pyrazine-2,3,5,6-tetracarbonitrile (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 1.5 h. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) and cysteine methyl ester (150mM in dH_2O, 100 equiv.), was added and the reaction was left at 22 °C, 300 rpm for 1 h. After this time the excess reagent was then removed *via* ultrafiltration (10 kDa MWCO) into HPLC grade

water prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

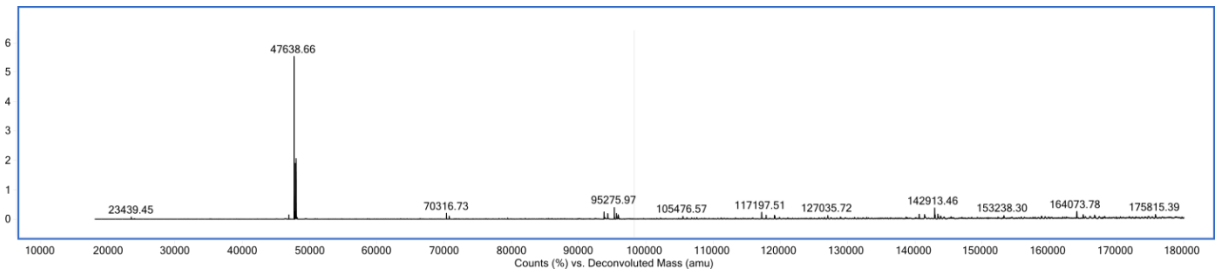
(a)



(b)



(c)



(d)

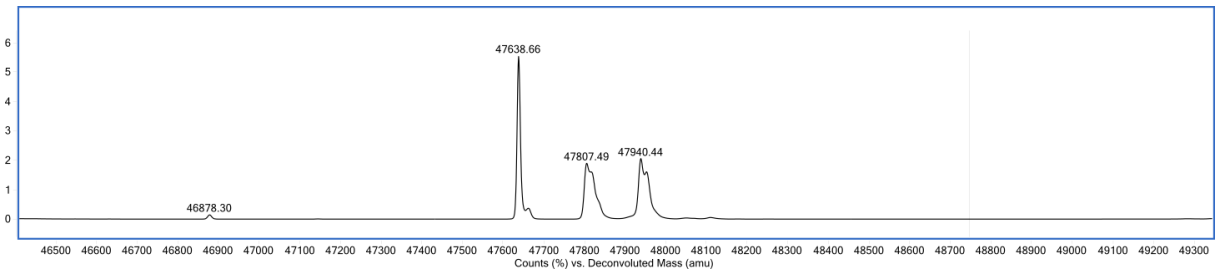
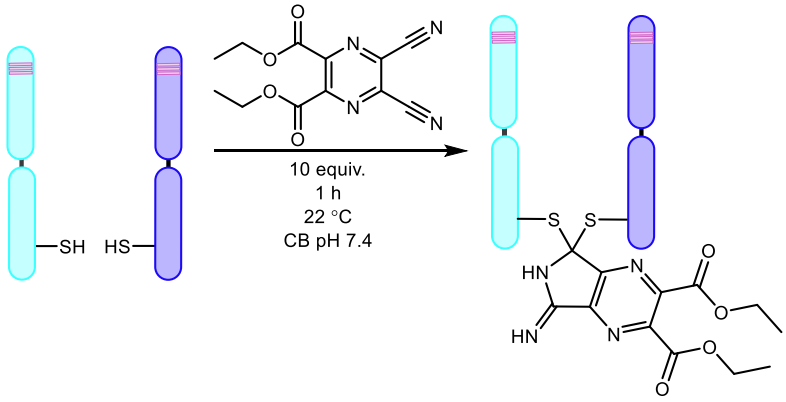


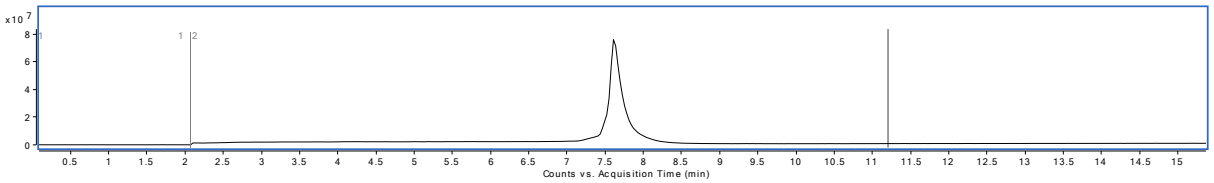
Figure S48: LCMS analysis of NAT functionalisation of conjugate **40**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; d) zoomed in deconvoluted ion series mass spectrum; Expected mass of NAT product: 47937 Da (1 addition), 48055 Da (2 addition). Observed: 47639 Da, 47807 Da, 47940 Da.

Reaction of Fab with diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate

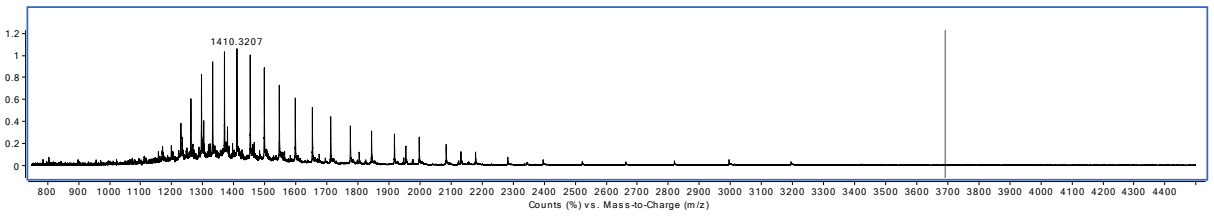


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 1 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

(a)



(b)



(c)

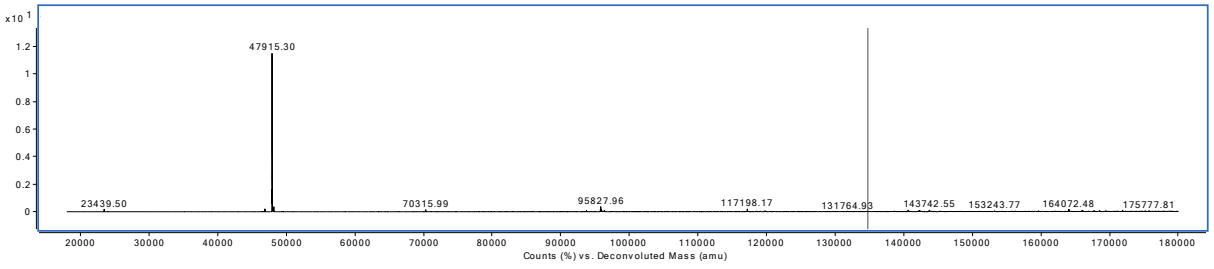
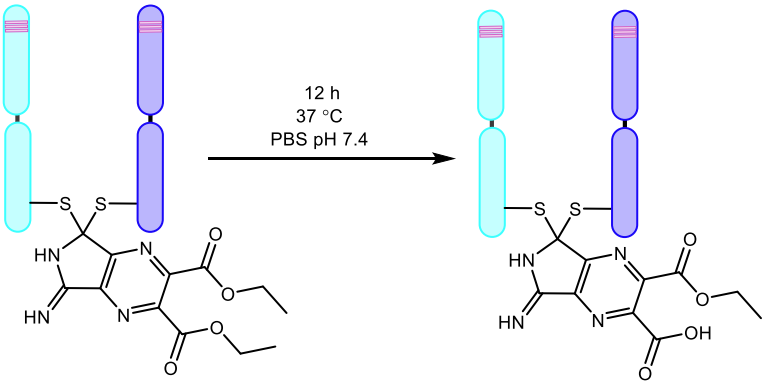


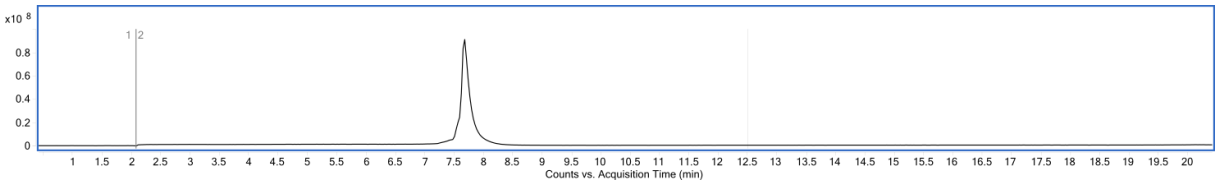
Figure S49: LCMS analysis of conjugate **44**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate: 47912 Da. Observed: 47915 Da.

Diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate bridge hydrolysis study

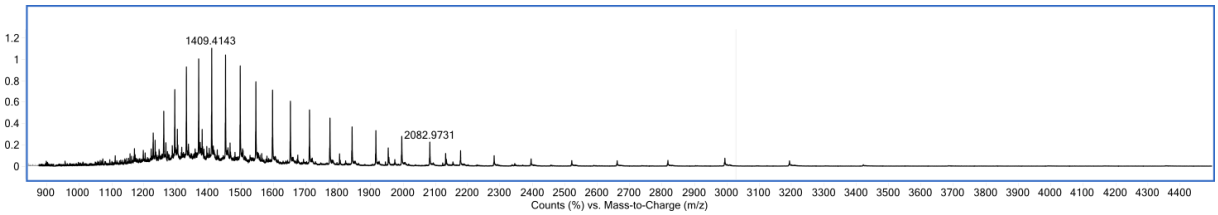


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 1 h. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) into PBS pH 7.4 for 12 h at 37 °C. The sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

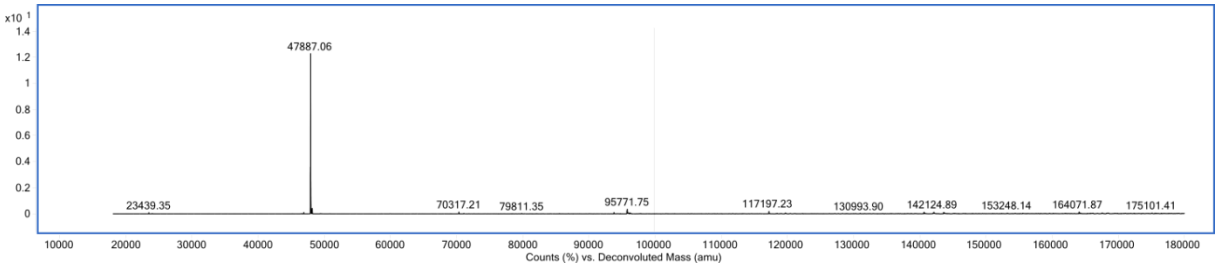
(a)



(b)



(c)



(d)

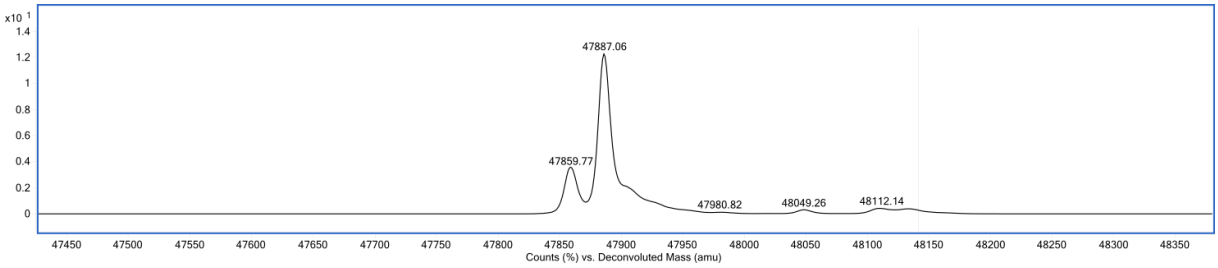
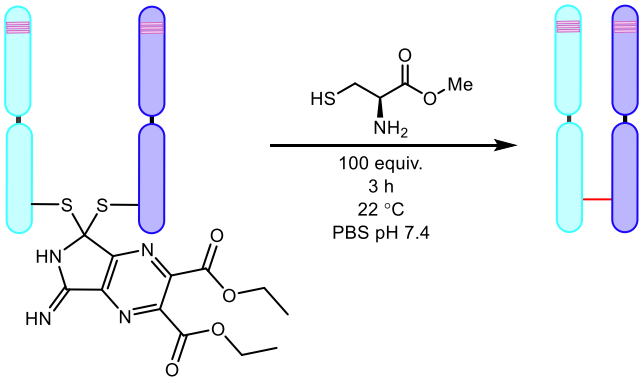


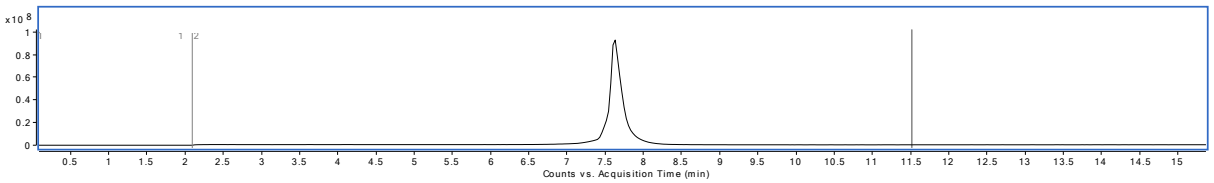
Figure S50: LCMS analysis of hydrolysis study of conjugate **44**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate: 47912 Da, 47885 Da (mono-ester hydrolysis), 47857 Da (di-ester hydrolysis). Observed: 47860 Da, 47887 Da.

Removal of diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate bridge

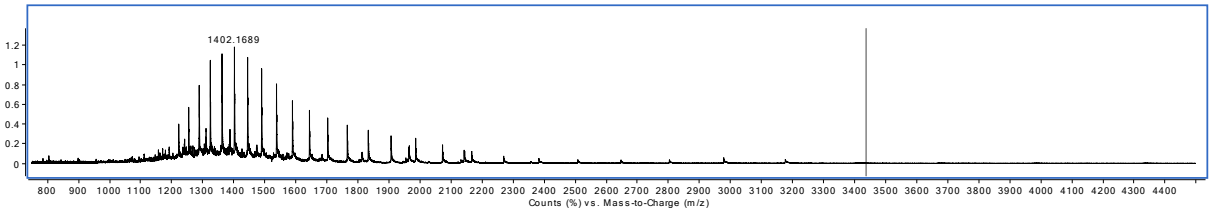


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Diethyl 5,6-dicyanopyrazine-2,3-dicarboxylate (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 1 h. Excess reagent was then removed *via* ultrafiltration (10 kDa MWCO, VivaSpin 500) and cysteine methyl ester (150mM in dH_2O, 100 equiv.), was added and the reaction was left at 22 °C, 300 rpm for 3 h. After this time the excess reagent was then removed *via* ultrafiltration (10 kDa MWCO) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

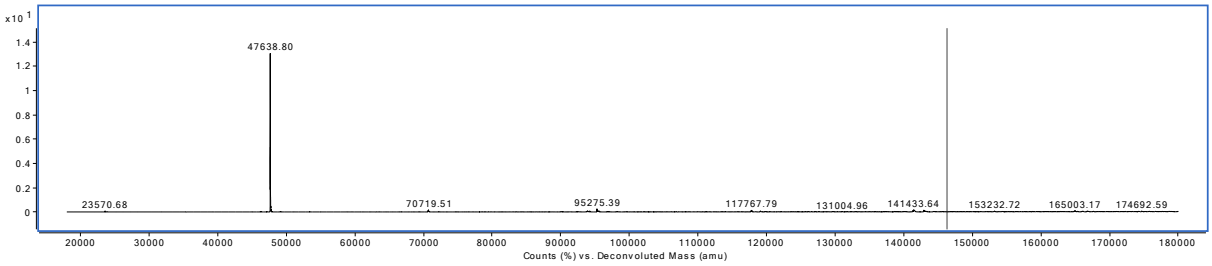
(a)



(b)



(c)



(d)

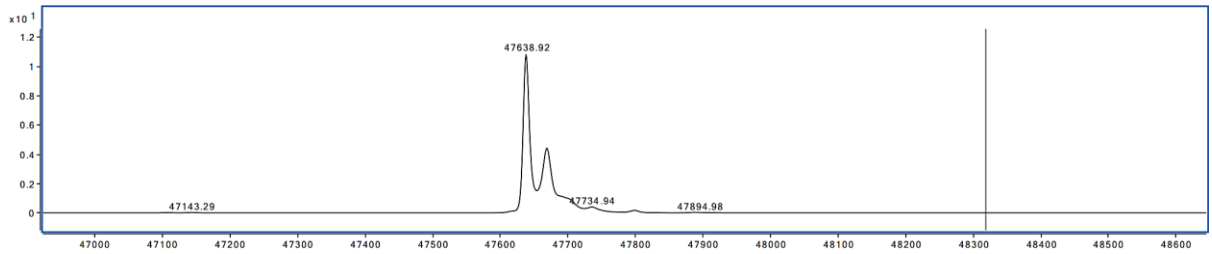
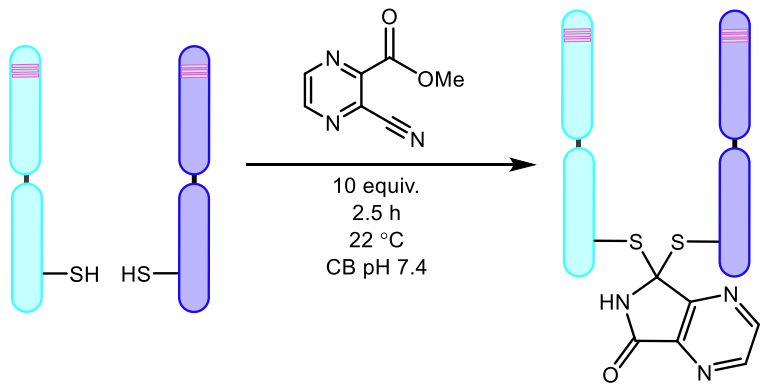


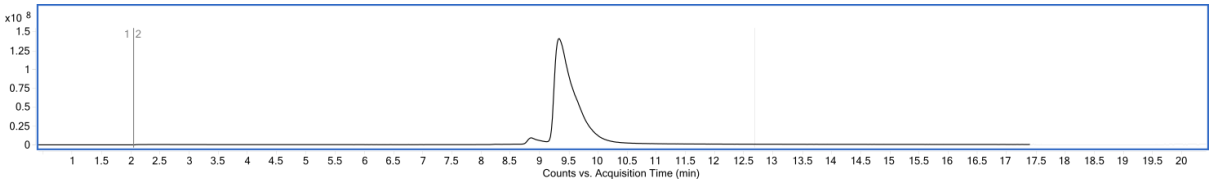
Figure S51: LCMS analysis of conjugate **44** removal; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of bridge removal Fab with cysteine methyl ester: 47638 Da. Observed: 47638 Da. d) zoomed in deconvoluted ion series mass spectrum; Expected mass of bridge removal Fab with cysteine methyl ester: 47638 Da. Observed: 47638 Da, 47668 Da (-OMe adduct due to excess cysteine methyl ester).

Reaction of Fab with methyl 3-cyanopyrazine-2-carboxylate

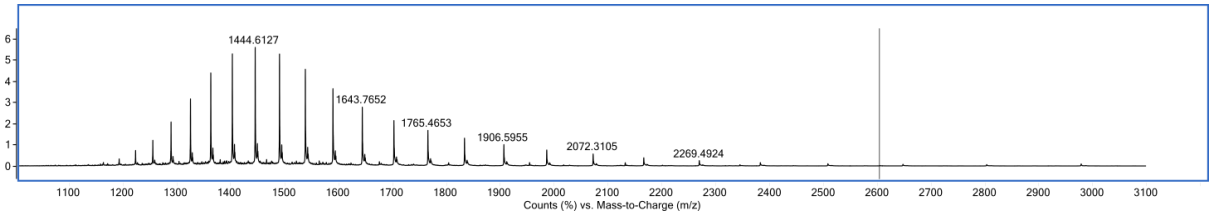


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Methyl 3-cyanopyrazine-2-carboxylate (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2.5 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

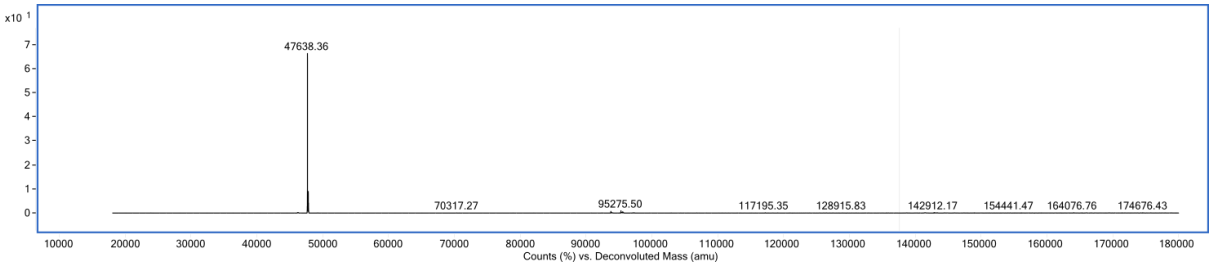
(a)



(b)



(c)



(d)

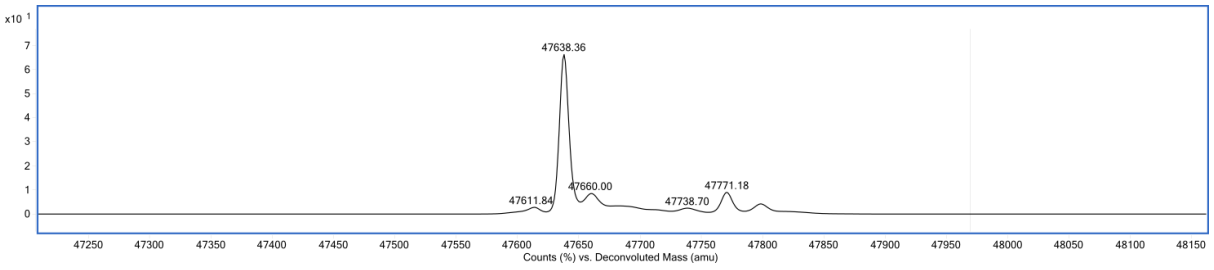
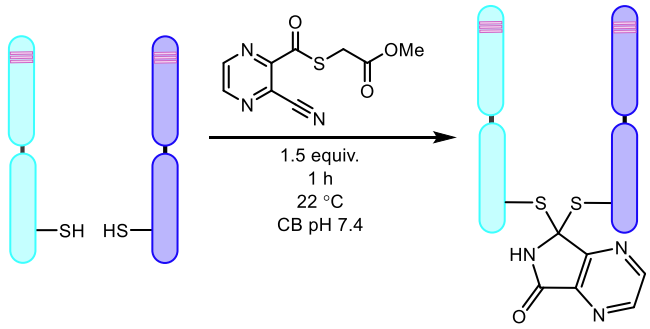


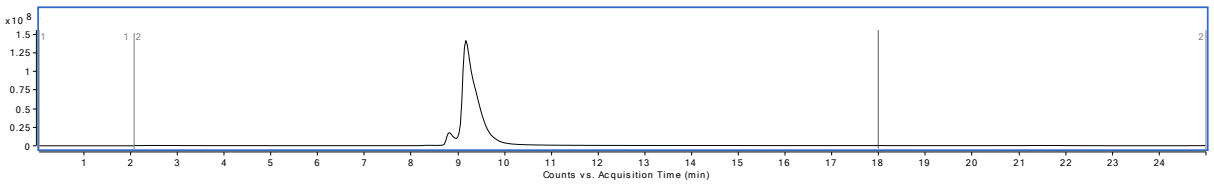
Figure S52: LCMS analysis of Fab conjugate; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate: 47771 Da. Observed: 47638 Da, 47771 Da.

Reaction of Fab with methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate

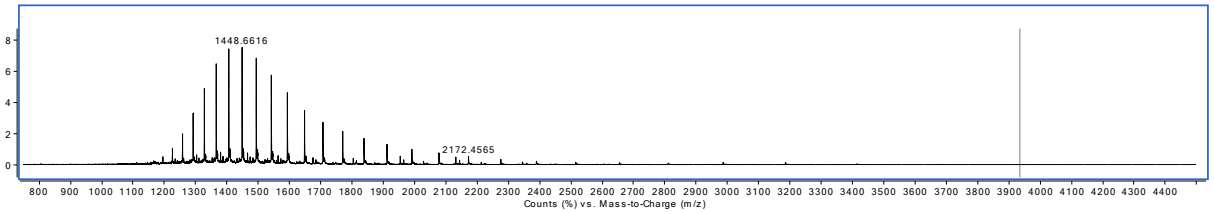


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate **11** (0.9 µL, 5 mM in DMF, 1.5 equiv.) was added and incubated at 22 °C for 1 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

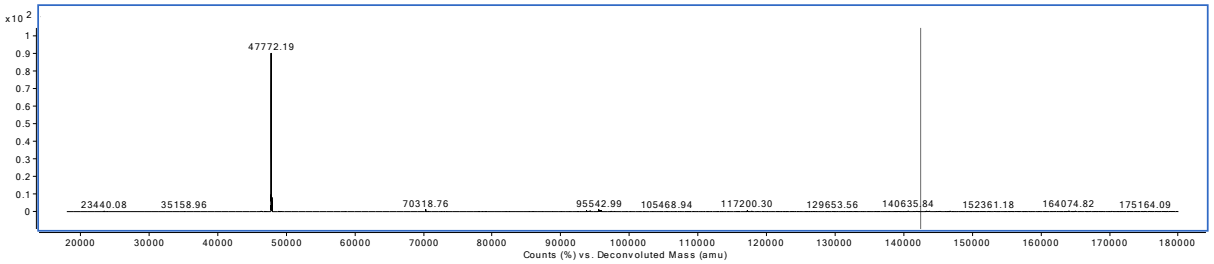
(a)



(b)



(c)



(d)

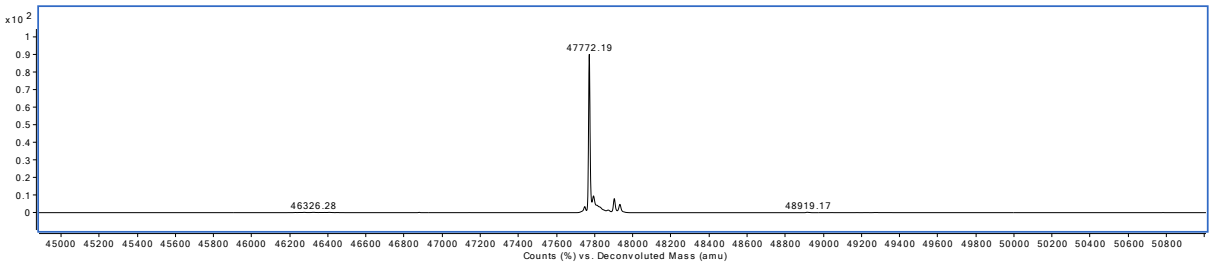
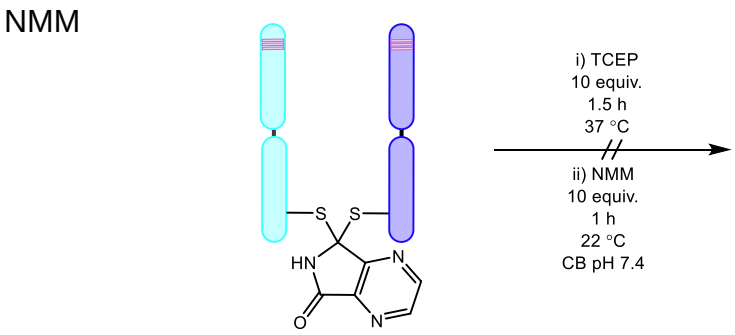


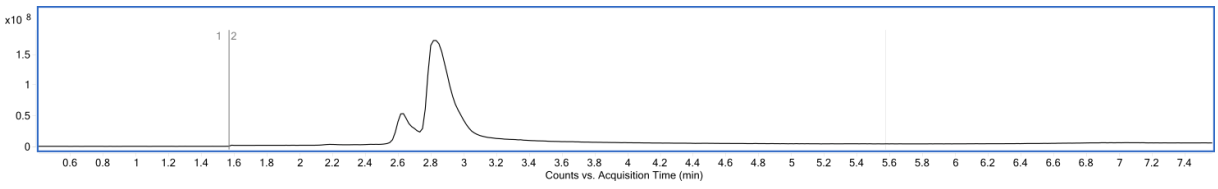
Figure S53: LCMS analysis of conjugate **45**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate: 47771 Da. Observed: 47772 Da.

Methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate bridge stability studies

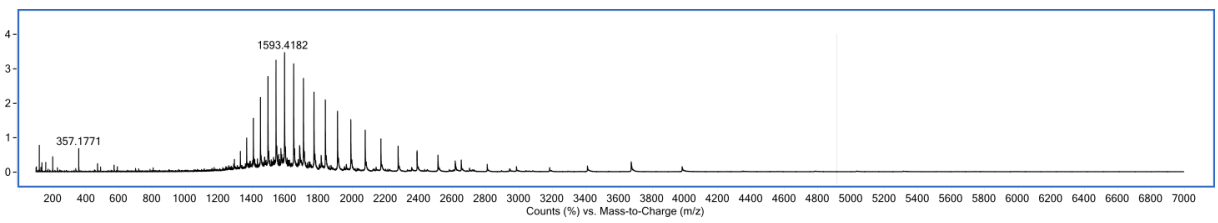


To conjugate **45** was added tris(2-carboxyethyl)phosphine (TCEP) (15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. After this time the excess reagent was removed *via* ultrafiltration (10 kDa MWCO) and NMM (20 mM in DMF, 10 equiv.)) was added for 1 h at 22 °C. The reaction was then purified *via* ultrafiltration (10 kDa MWCO) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

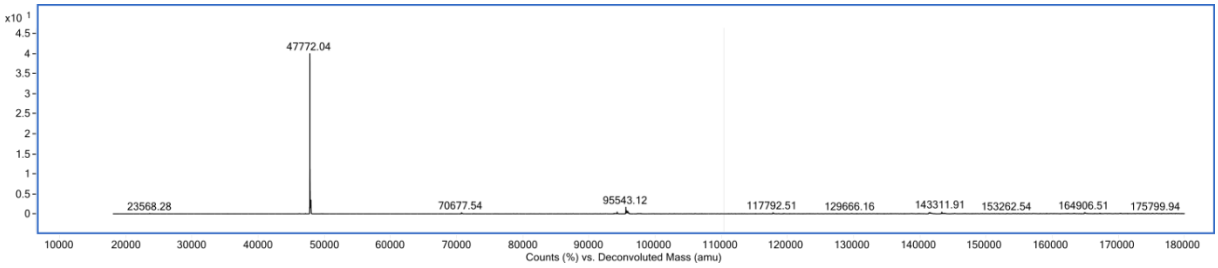
(a)



(b)



(c)



(d)

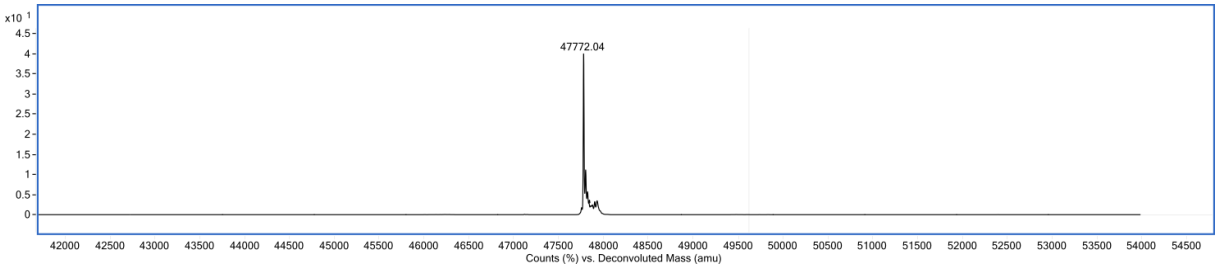
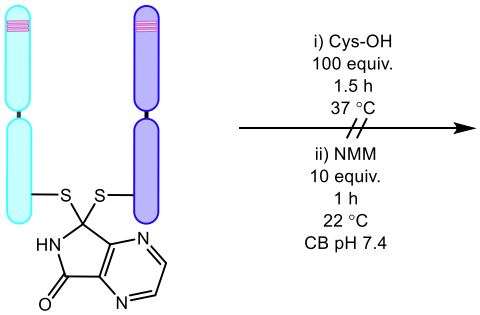


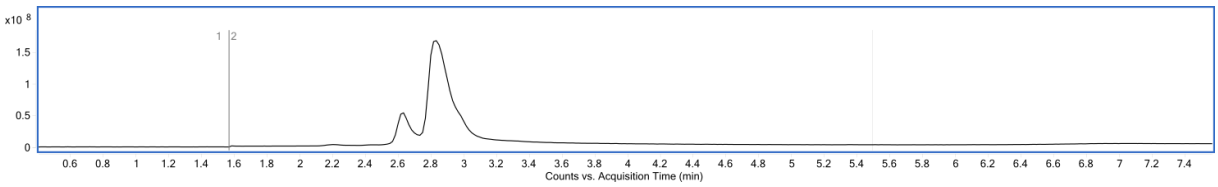
Figure S54: LCMS analysis of conjugate **45** stability; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate: 47771 Da. Observed: 47772 Da.

Cysteine

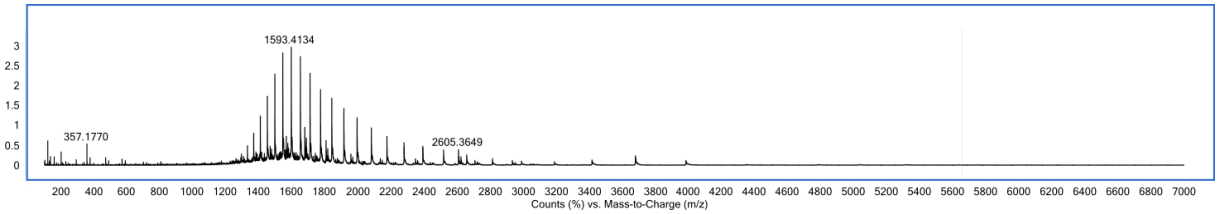


To conjugate **45** was added cysteine (150mM in dH_2O, 100 equiv.), the reaction was left at 37 °C, 300 rpm for 1.5 h. After this time the excess reagent was removed *via* ultrafiltration (10 kDa MWCO) and NMM (20 mM in DMF, 10 equiv.)) was added for 1 h at 22 °C. The reaction was then purified *via* ultrafiltration (10 kDa MWCO) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

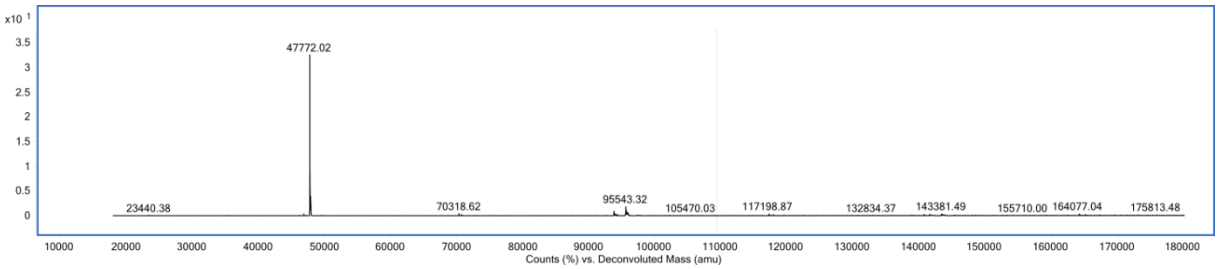
(a)



(b)



(c)



(d)

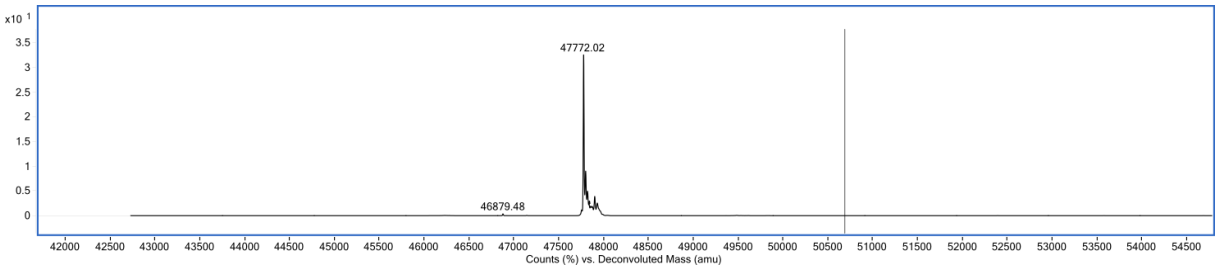
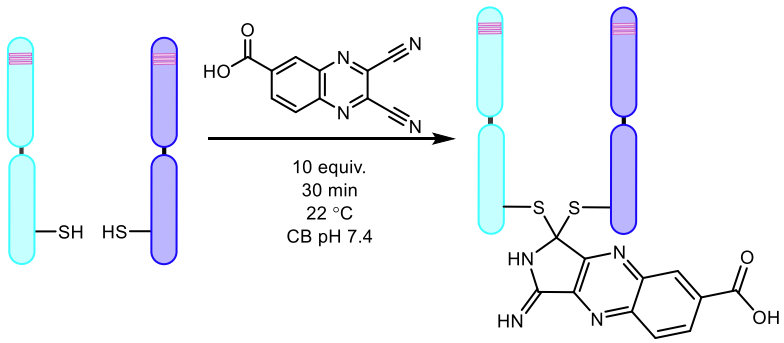


Figure S55: LCMS analysis of conjugate **45** stability; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with methyl 2-((3-cyanopyrazine-2-carbonyl)thio)acetate: 47771 Da. Observed: 47772 Da.

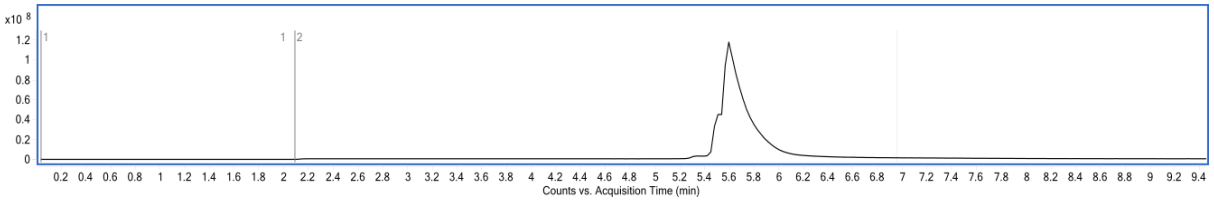
Reaction of Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid



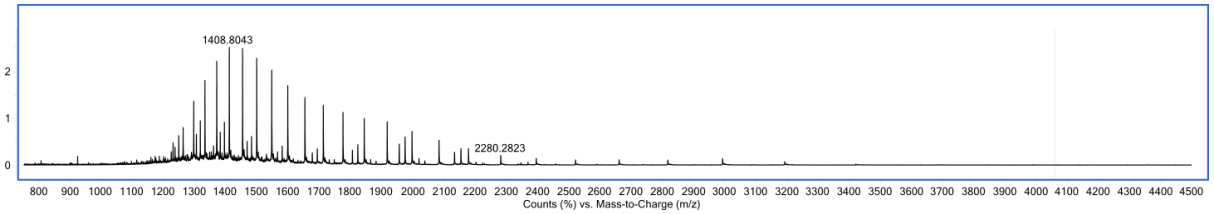
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 2,3-dicyanoquinoxaline-6-carboxylic acid (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 30 min. Lastly, excess reagent was removed, and the sample was desalted *via* ultrafiltration (7 kDa MWCO, ZebaSpin) prior

to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

(a)



(b)



(c)

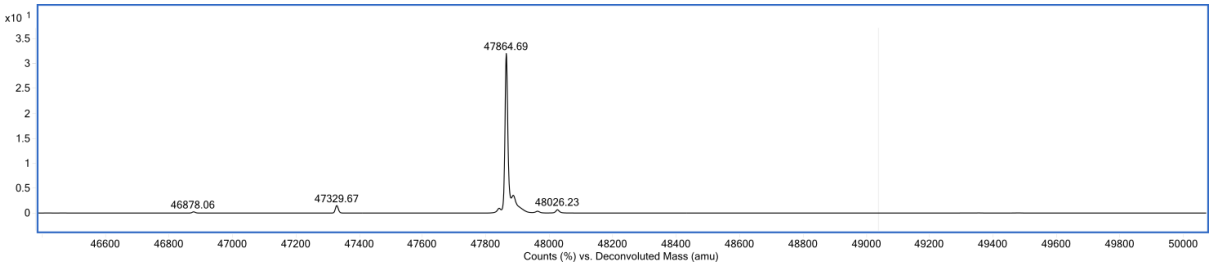
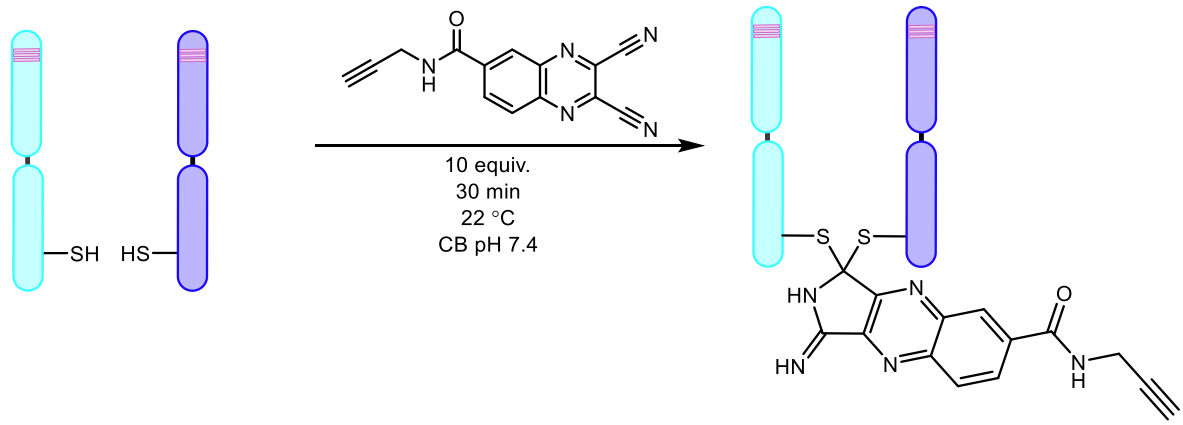


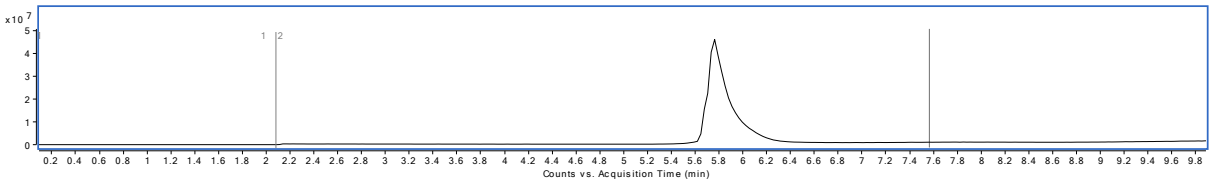
Figure S56: LCMS analysis of conjugate **55**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid :47863 Da. Observed: 47865 Da.

Reaction of Fab with 2,3-dicyano-N-(prop-2-yn-1-yl)quinoxaline-6-carboxamide.

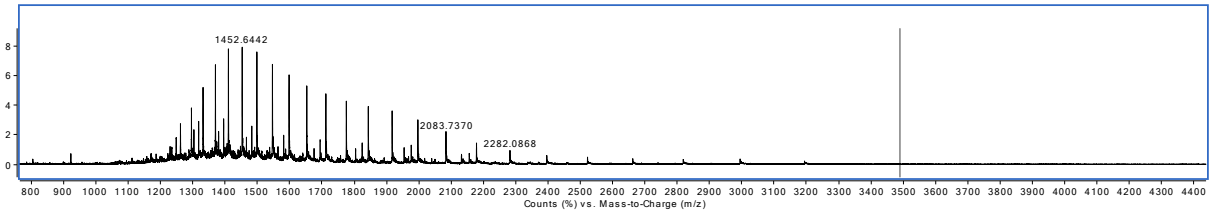


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 2,3-dicyano-N-(prop-2-yn-1-yl)quinoxaline-6-carboxamide (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 30 min. Lastly, excess reagent was removed, and the sample was desalted *via* ultrafiltration (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

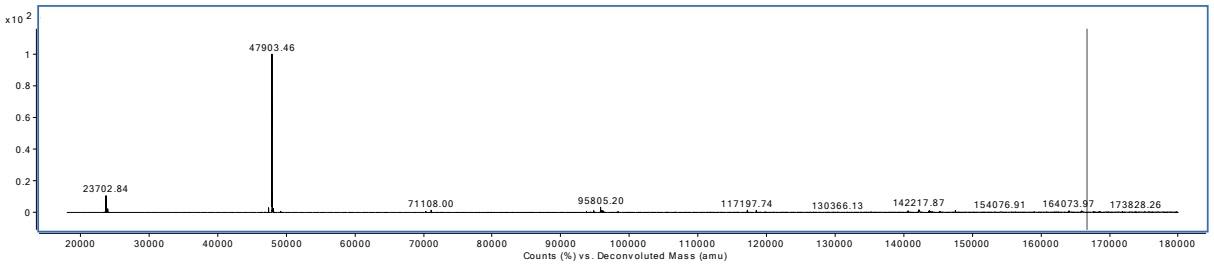
(a)



(b)



(c)



(d)

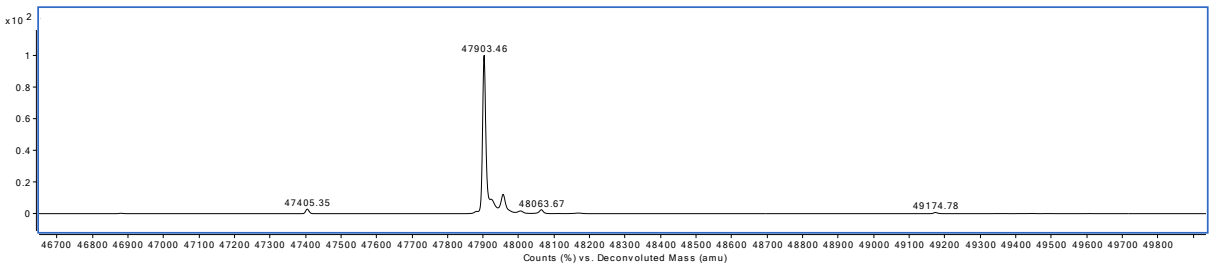
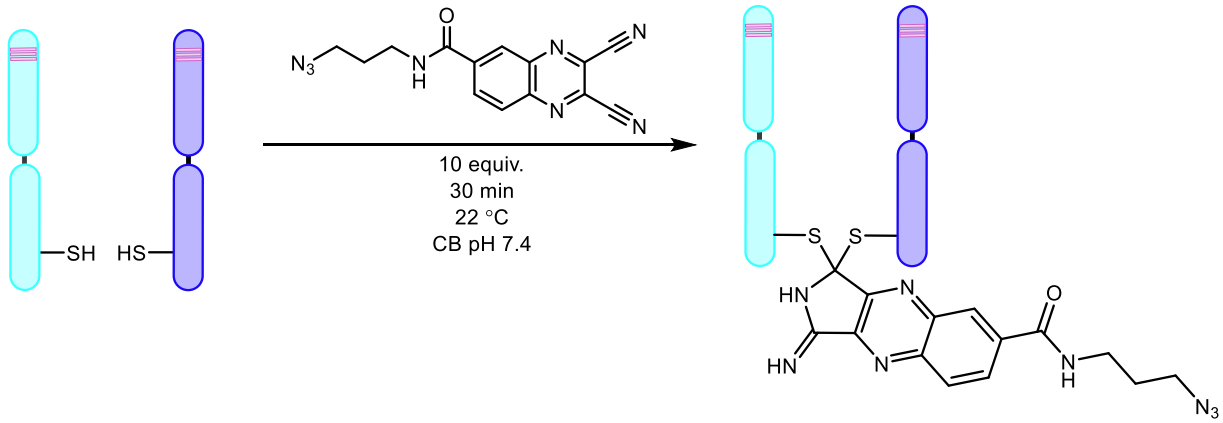


Figure S57: LCMS analysis of conjugate **50**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass

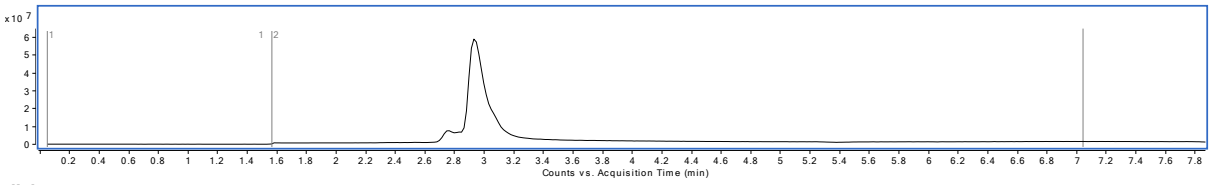
spectrum; Expected mass of re-bridged Fab with 2,3-dicyano-N-(prop-2-yn-1-yl)quinoxaline-6-carboxamide:47899 Da. Observed: 47903 Da.

Reaction of Fab with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide

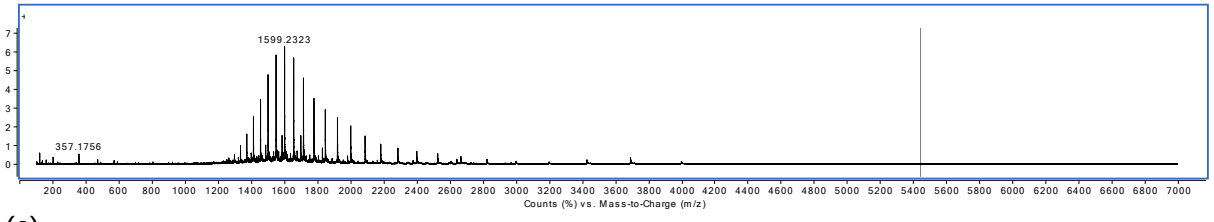


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide **6** (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 30 min. Lastly, excess reagent was removed, and the sample was desalted *via* ultrafiltration (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

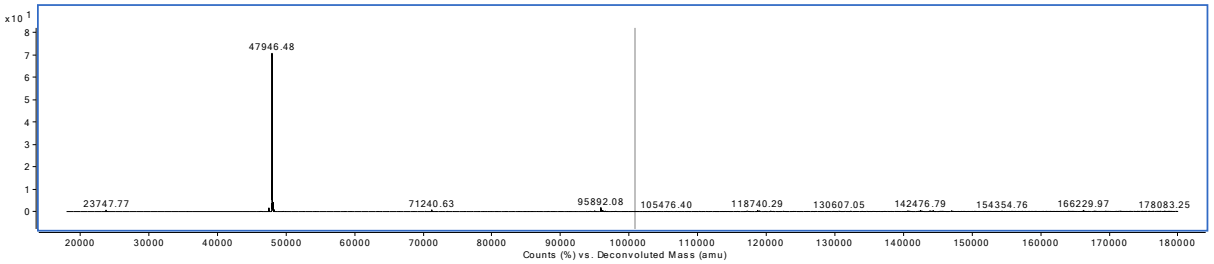
(a)



(b)



(c)



(d)

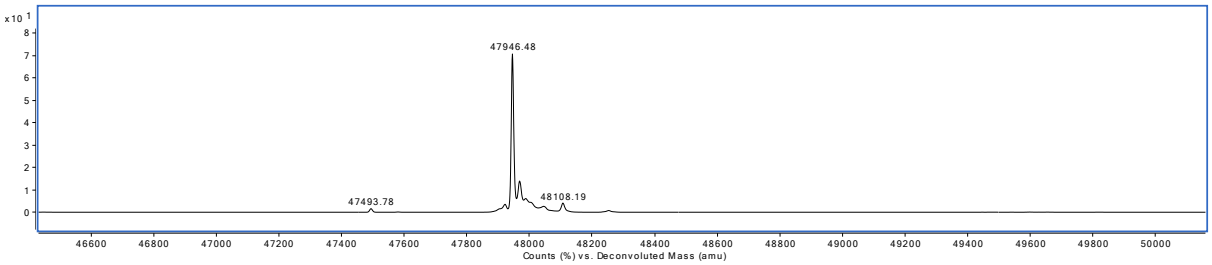
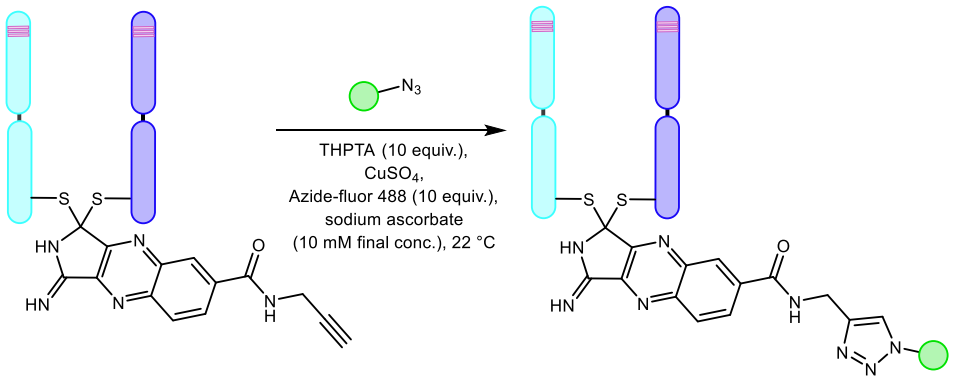


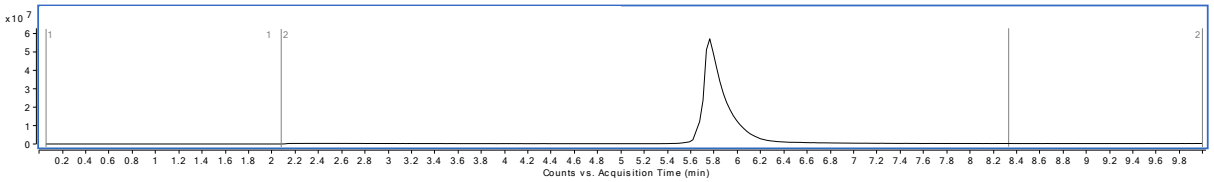
Figure S58: LCMS analysis of conjugate **51**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide: 47944 Da. Observed: 47947 Da.

Reaction of Fab with 2,3-dicyano-N-(prop-2-yn-1-yl)quinoxaline-6-carboxamide with CuAAC.

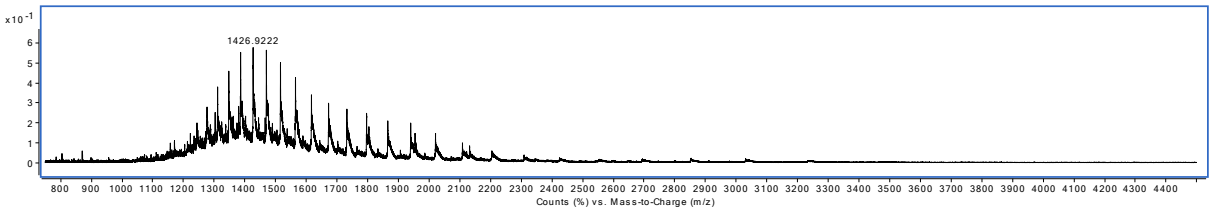


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with TCEP (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. 2,3-Dicyano-*N*-(prop-2-yn-1-yl)quinoxaline-6-carboxamide **S4** (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2 h. Excess reagent was removed *via* ultrafiltration (7 kDa MWCO, ZebaSpin) into PB (pH 7.0). To conjugate **S13** was added THPTA (100 mM in dH_2O, 10 equiv.), $CuSO_4$ (20 mM in dH_2O, 5 equiv.), Azide-fluor 488 (10 mM in DMF, 10 equiv.), and sodium ascorbate (100 mM in dH_2O, final conc. 10 mM), the reaction was left at 22 °C, 300 rpm for 4 h. The excess reagent was then removed using a desalting column (PD Minitrap G-25, GE Healthcare) followed by ultrafiltration (10 kDa MWCO) into conjugate buffer to concentrate the sample. Lastly, the sample was desalted into HPLC grade water (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

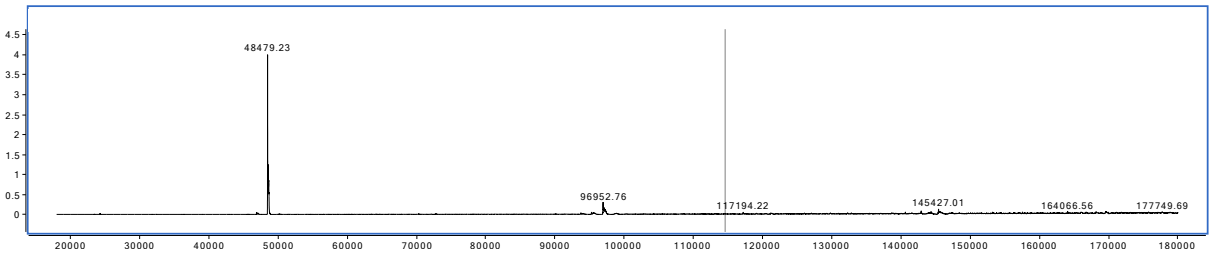
(a)



(b)



(c)



(d)

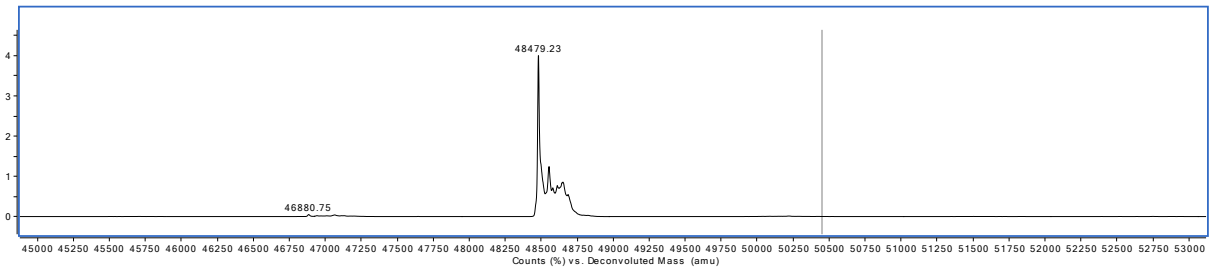
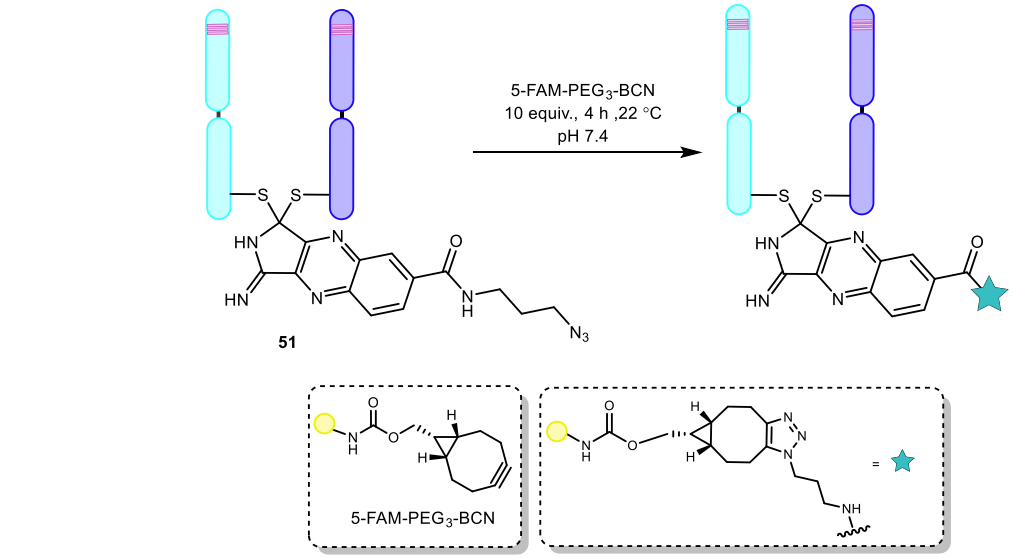


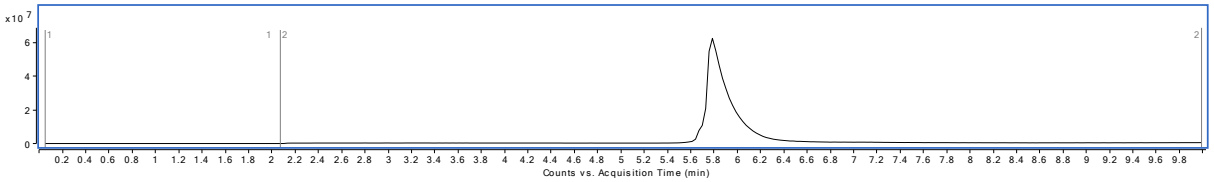
Figure S59: LCMS analysis of conjugate **50** CuAAC; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyano-N-(prop-2-yn-1-yl)quinoxaline-6-carboxamide after CuAAC: 48473 Da. Observed: 48479 Da.

Reaction of Fab with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide and SPAAC

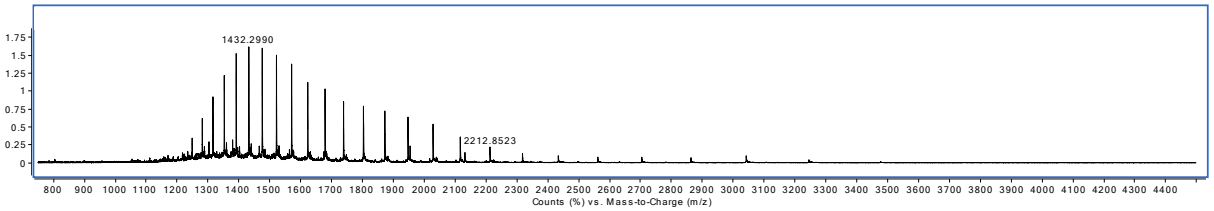


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide **48** was then added (2.0 µL, 150 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 30 min. Excess reagent was then removed *via* ultrafiltration (7 kDa MWCO, ZebaSpin) into PBS (pH = 7.4). To conjugate **51** was added fluorescein-PEG_3-BCN (20 mM in DMF, 10 equiv.), this was then left in the dark at 22 °C for 4 h, 300 rpm. Excess reagent was then removed by ultrafiltration (10 kDa MWCO) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε_{280}= 68590 M^{-1} cm^{-1}.

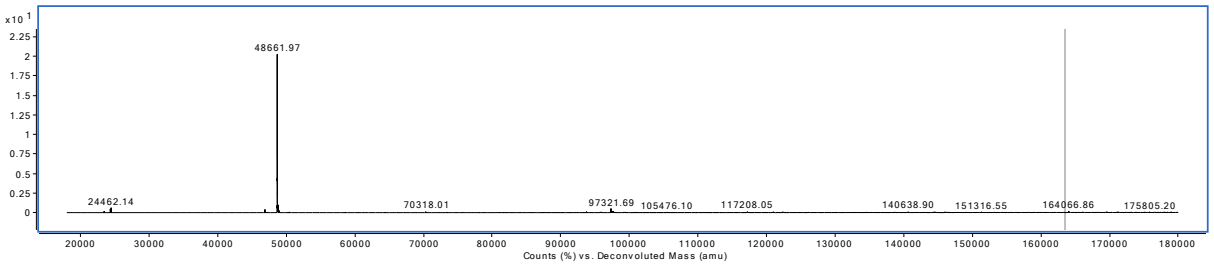
(a)



(b)



(c)



(d)

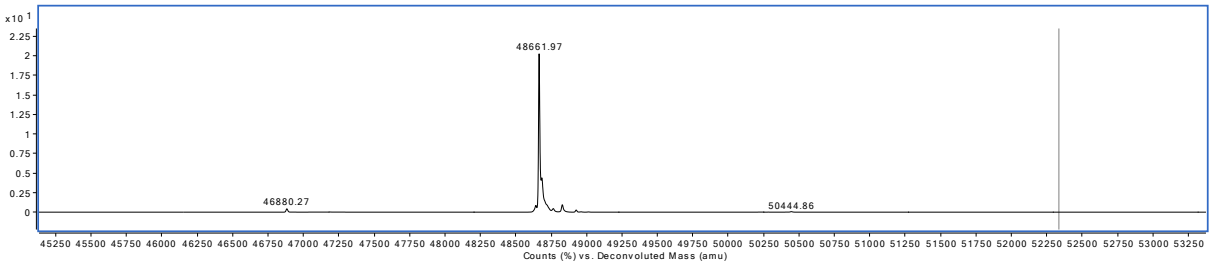
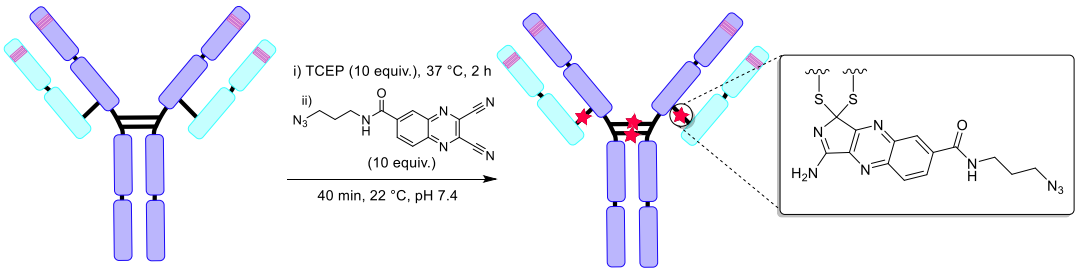


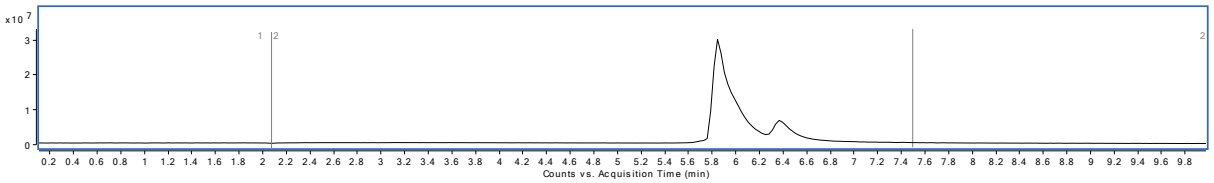
Figure S60: LCMS analysis of conjugate **52**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide after SPAAC: 48670 Da. Observed: 48662 Da.

Re-bridging of full antibody with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide

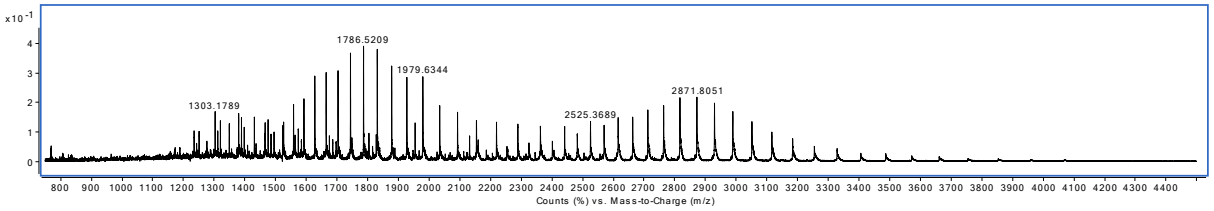


Trastuzumab full antibody (80 µL, 25 µM, 3.75 mg/mL) in conjugation buffer was reduced with TCEP (1.3 µL, 15 mM in dH_2O, 10 equiv.). The mixture was incubated at 37 °C for 2 h, 300 rpm. Following this *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide (0.92 µL, 21.8 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 40 min. The excess reagent was removed *via* ultrafiltration (10 kDa MWCO) into 50 mM ammonium acetate pH 6.9 (7 kDa MWCO, ZebaSpin) and adjusted to 5 µM with ammonium acetate (pH 6.9). To this (40 µL) was added 0.67 µL of PNGase F and the reaction was left at 37 °C for 16 h. Concentration was determined photometrically using ε_{280}=215380 M^{-1} cm^{-1}.

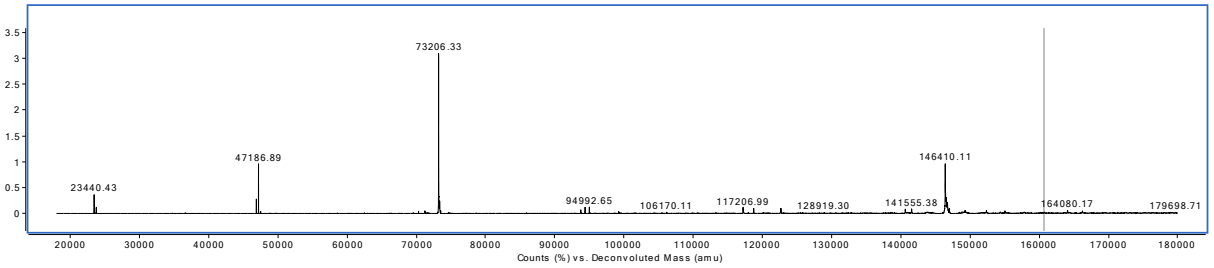
(a)



(b)



(c)



(d)

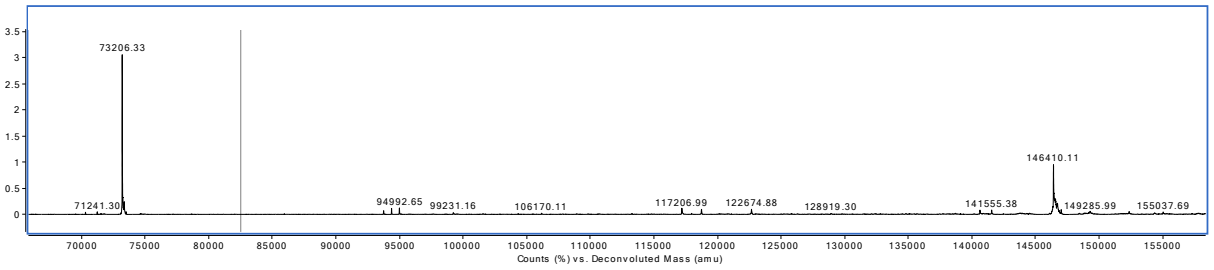
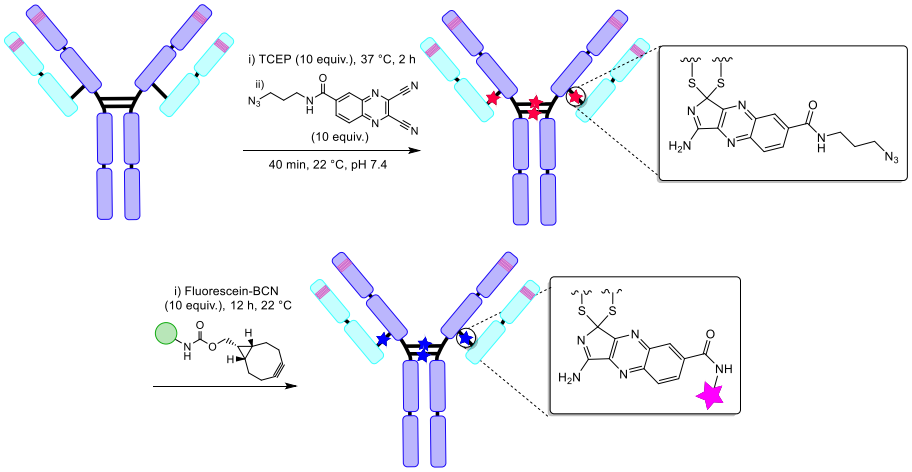


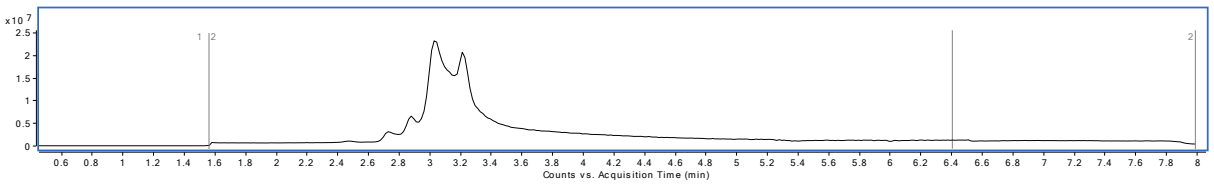
Figure S61: LCMS analysis of conjugate **53**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged trastuzumab with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide: **HHLL** 146403 Da, **HL** 73206 Da. Observed: **HHLL** 146410 Da, **HL** 73206 Da, **LC** 23440 Da.

Re-bridging of full antibody with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide and SPAAC

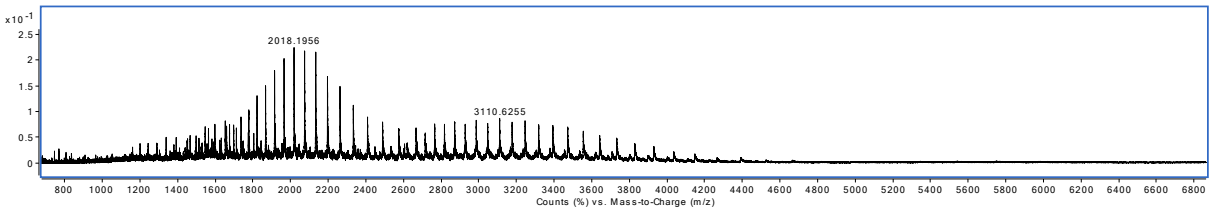


Trastuzumab full antibody (80 µL, 25 µM, 3.75 mg/mL) in conjugation buffer was reduced with TCEP (1.3 µL, 15 mM in dH_2O, 10 equiv.). The mixture was incubated at 37 °C for 2 h, 300 rpm. Following this *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide (0.92 µL, 21.8 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 40 min. After this time, excess reagent was removed *via* ultrafiltration (7 kDa MWCO, ZebaSpin) into PBS (pH = 7.4). To the resulting conjugate was added fluorescein-PEG_3-BCN (20 mM in DMF, 20 equiv.), this was then left in the dark at 22 °C for 12 h, 300 rpm. The excess reagent was removed *via* ultrafiltration (10 kDa MWCO) into 50 mM ammonium acetate pH 6.9 (7 kDa MWCO, ZebaSpin) and adjusted to 5 µM with ammonium acetate (pH 6.9). To this (40 µL) was added 0.67 µL of PNGase F and the reaction was left at 37 °C for 16 h. Concentration was determined photometrically using ε_{280}=215380 M^{-1} cm^{-1}.

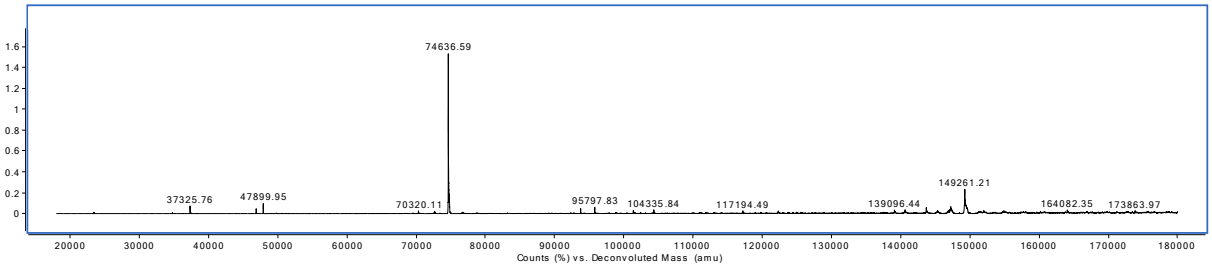
(a)



(b)



(c)



(d)

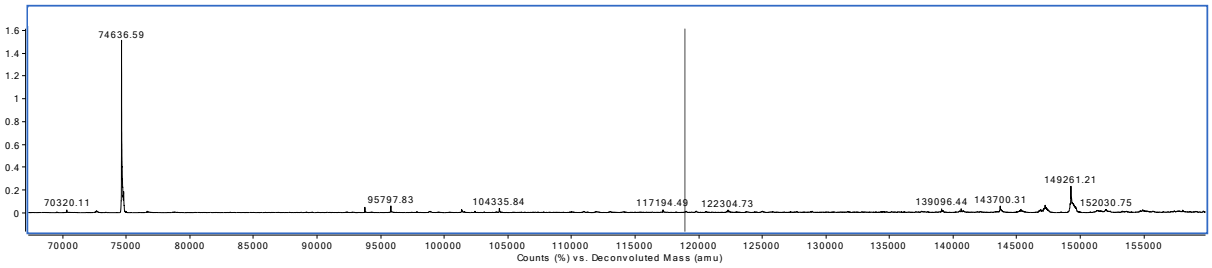
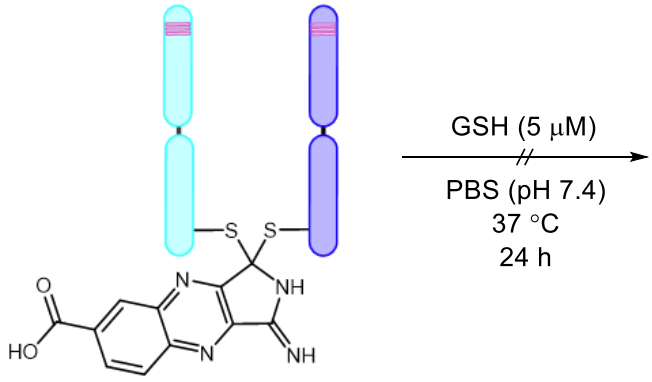


Figure S62: LCMS analysis of conjugate **54** a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged trastuzumab with *N*-(3-azidopropyl)-2,3-dicyanoquinoxaline-6-carboxamide after SPAAC: **HHLL** 149306 Da, **HL** 74658 Da. Observed: **HHLL** 149261 Da, **HL** 74636 Da.

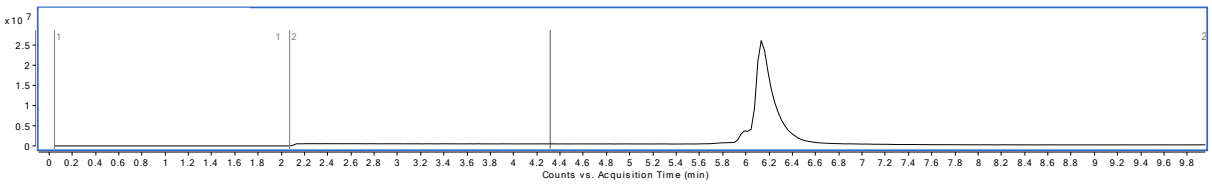
GSH stability studies – Bridge



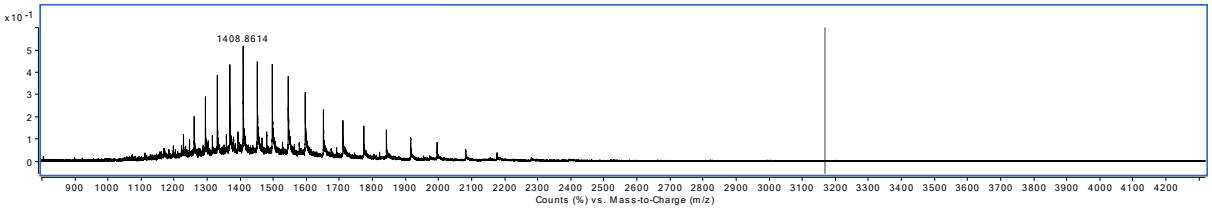
Fab conjugate **55** (50 µL, 7 µM) was buffer exchanged using desalting columns (7000 MWCO, ZebaSpin, Thermo Scientific) into GSH containing buffer (5 µM GSH, in PBS pH 7.4) and maintained at 37 °C for 24 h. Excess reagents were removed using desalting columns (7000 MWCO, ZebaSpin, Thermo Scientific) prior to LCMS analysis.

0 hours

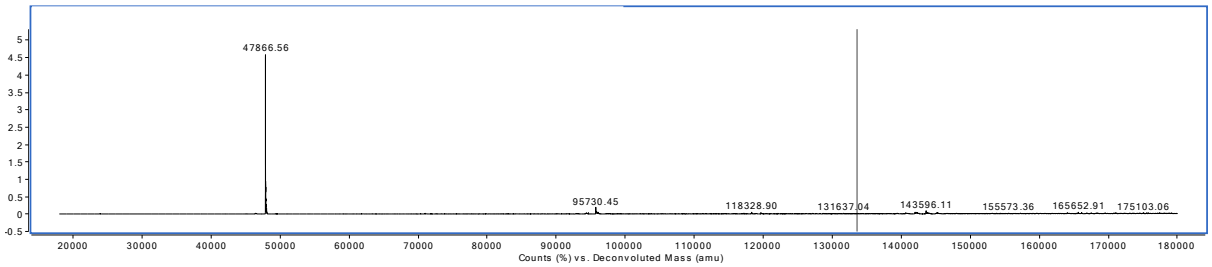
(a)



(b)



(c)



(d)

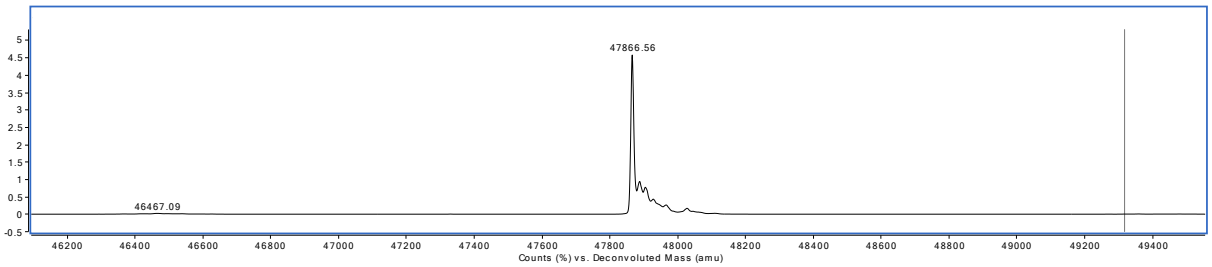
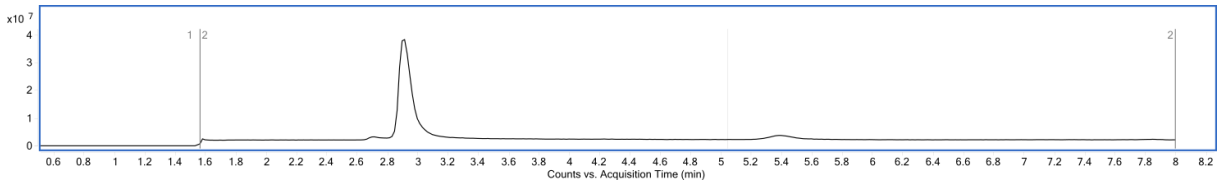


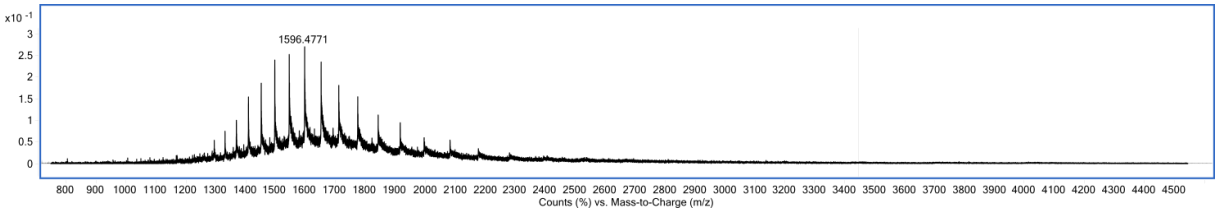
Figure S63: LCMS analysis of conjugate **55**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47866 Da.

6 hours

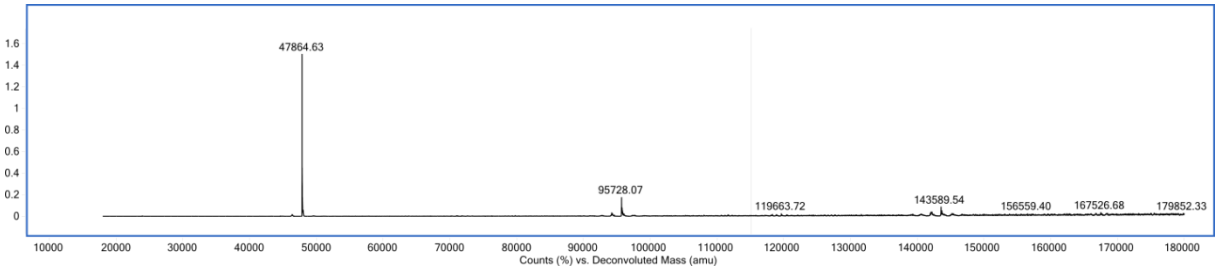
(a)



(b)



(c)



(d)

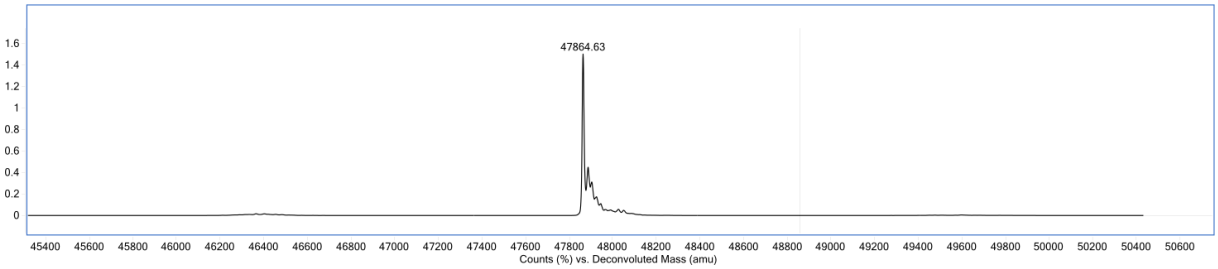
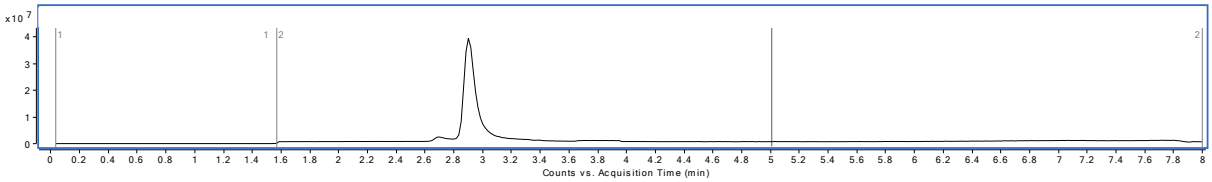


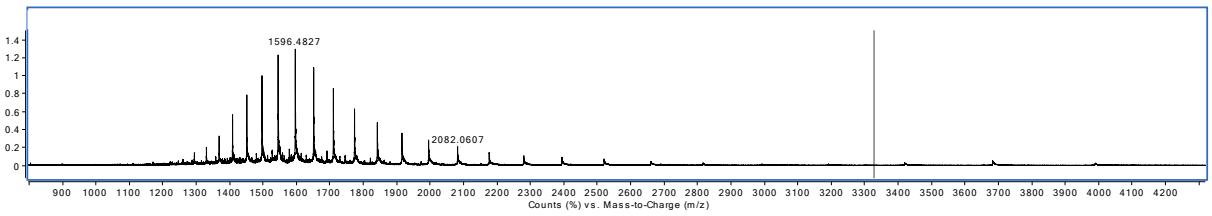
Figure S64: LCMS analysis of conjugate **55**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47865 Da.

24 hours

(a)



(b)



(c)

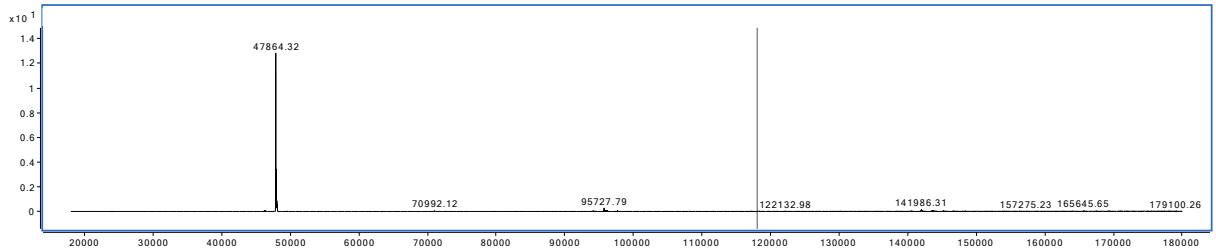
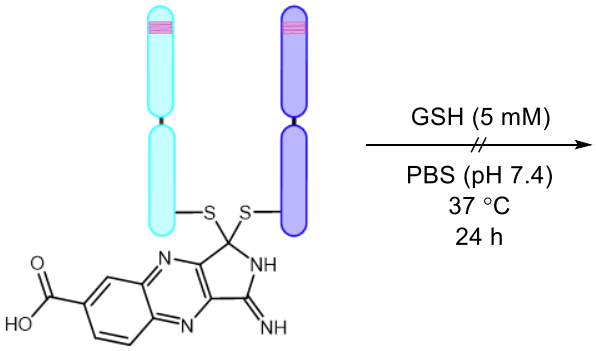


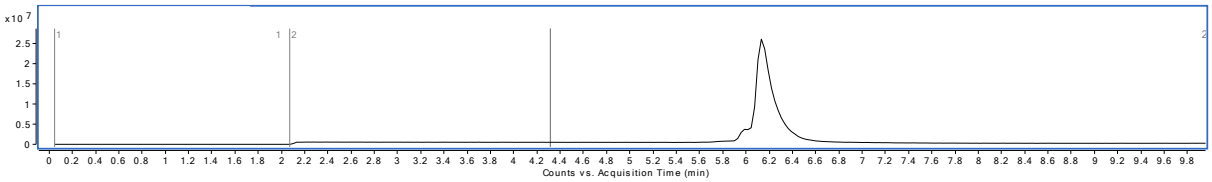
Figure S65: LCMS analysis of conjugate **55**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47864 Da.



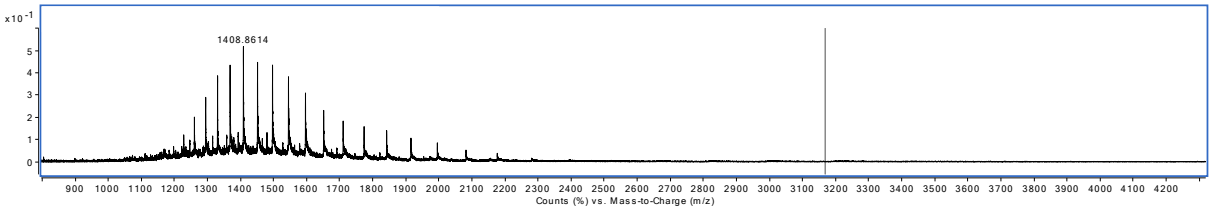
Fab conjugate **55** (50 µL, 7 µM) was buffer exchanged using desalting columns (7000 MWCO, ZebaSpin, Thermo Scientific) into GSH containing buffer (5 mM GSH, in PBS pH 6.5) and maintained at 37 °C for 24 h. Excess reagents were removed using desalting columns (7000 MWCO, ZebaSpin, Thermo Scientific) prior to LCMS analysis.

0 hours

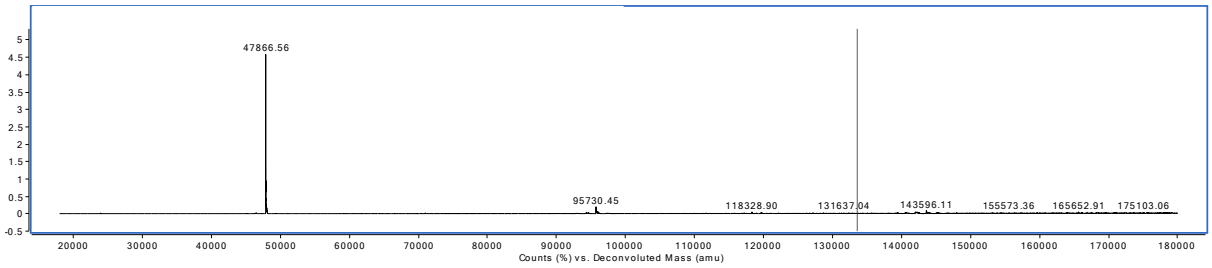
(a)



(b)



(c)



(d)

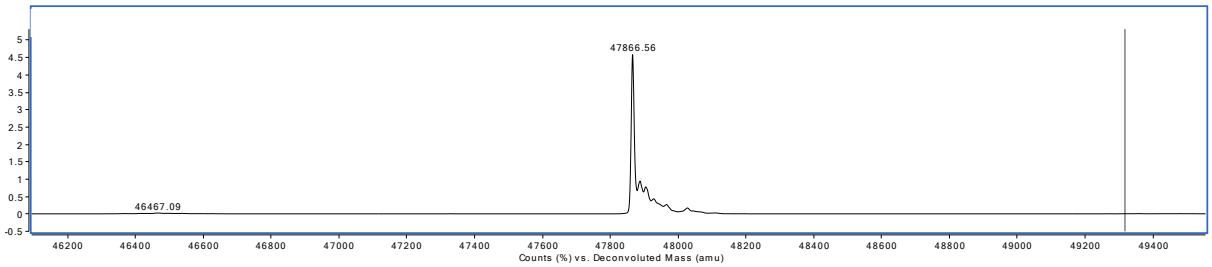
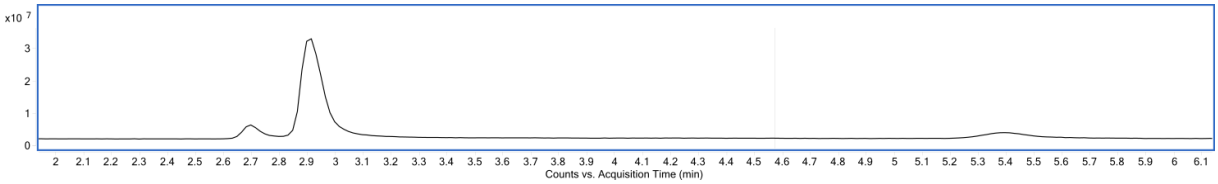


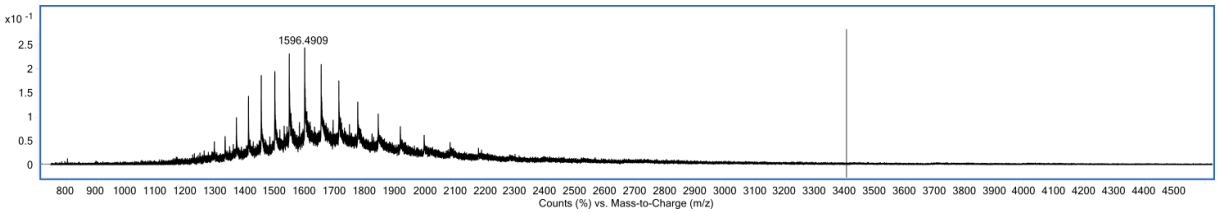
Figure S66: LCMS analysis of conjugate **55**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47866 Da.

6 hours

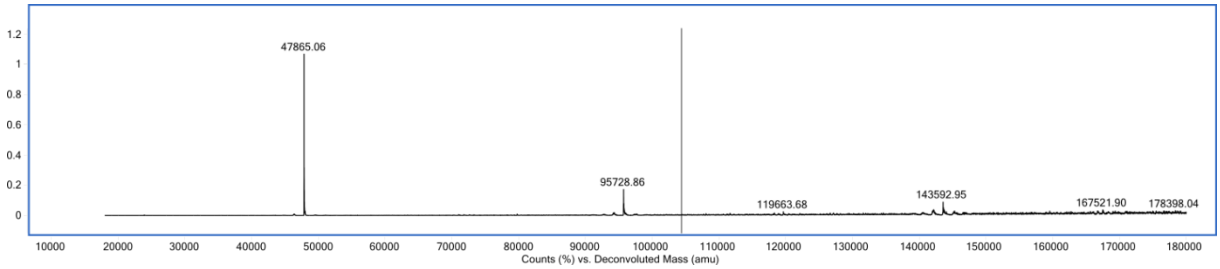
(a)



(b)



(c)



(d)

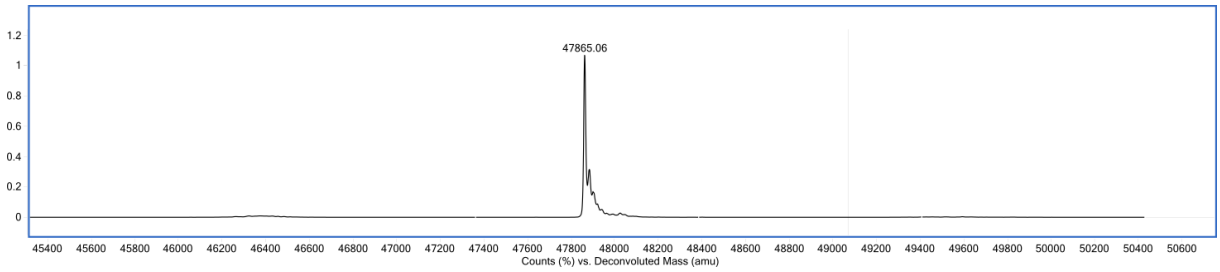
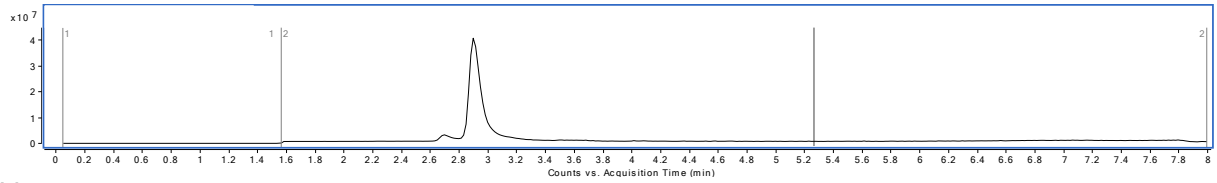


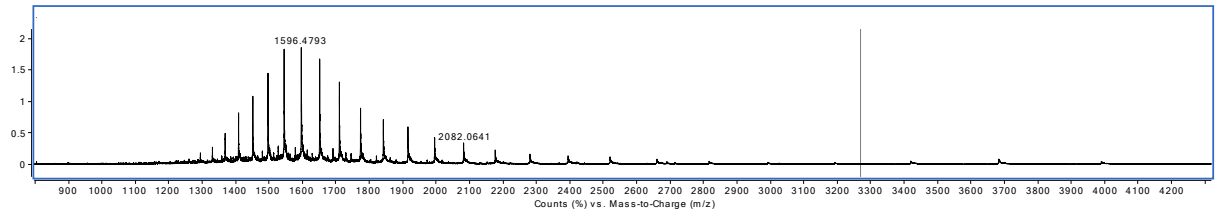
Figure S67: LCMS analysis of conjugate **55**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47865 Da.

24 hours

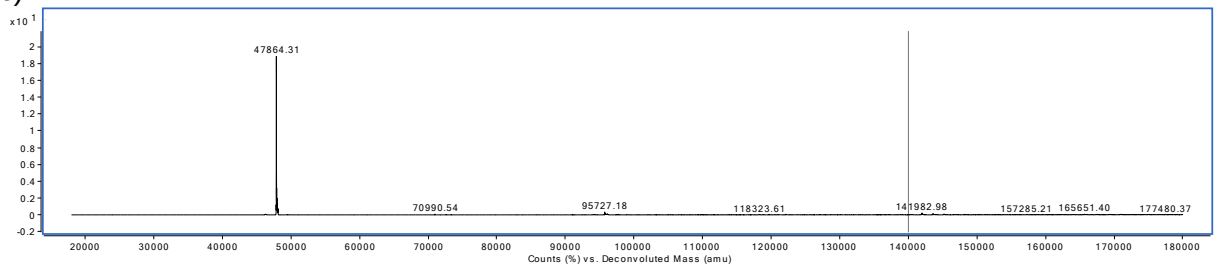
(a)



(b)



(c)



(d)

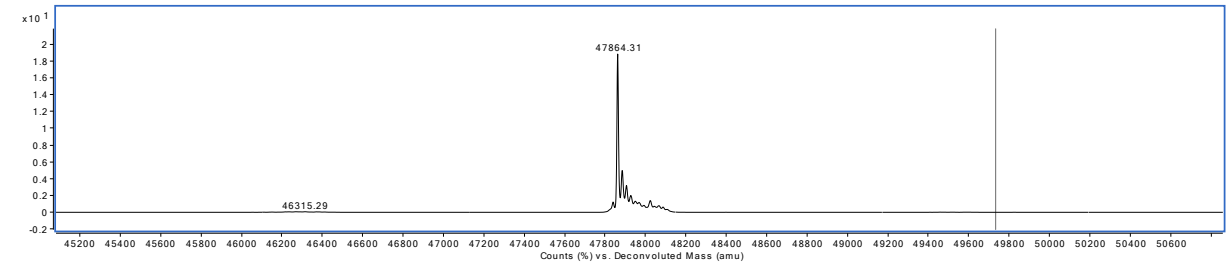
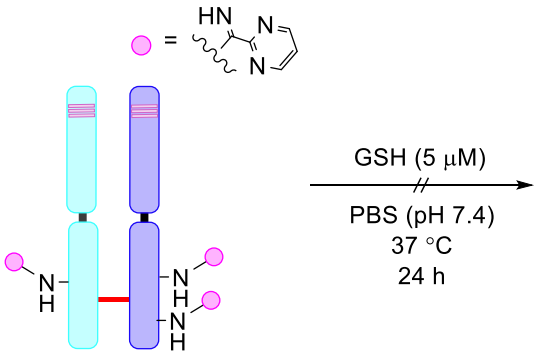


Figure S68: LCMS analysis of conjugate **55**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47864 Da.

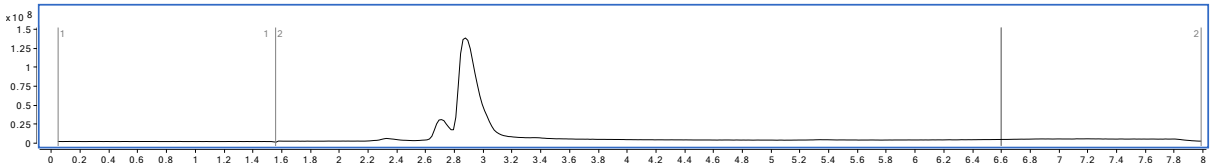
GSH stability studies – CLT



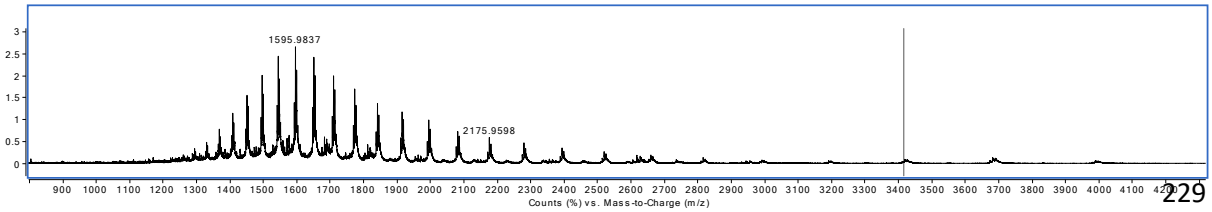
Fab conjugate **18** (50 µL, 7 µM) was buffer exchanged using desalting columns (7000 MWCO, ZebaSpin, Thermo Scientific) into GSH containing buffer (5 µM GSH, in PBS pH 7.4) and maintained at 37 °C for 24 h. Excess reagents were removed using desalting columns (7000 MWCO, ZebaSpin, Thermo Scientific) prior to LCMS analysis.

0 hours

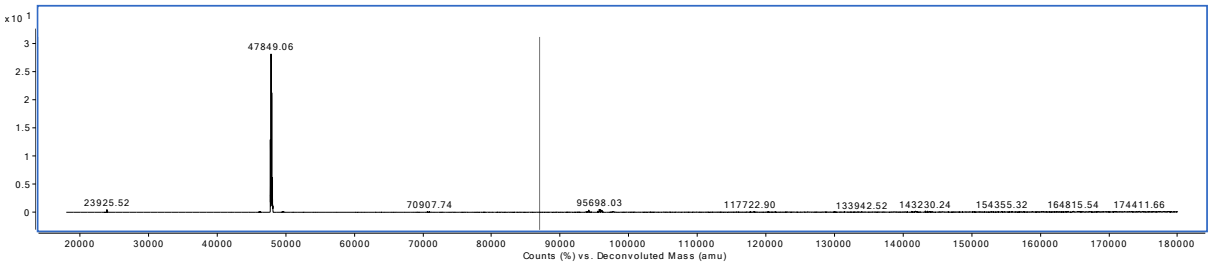
(a)



(b)



(c)



(d)

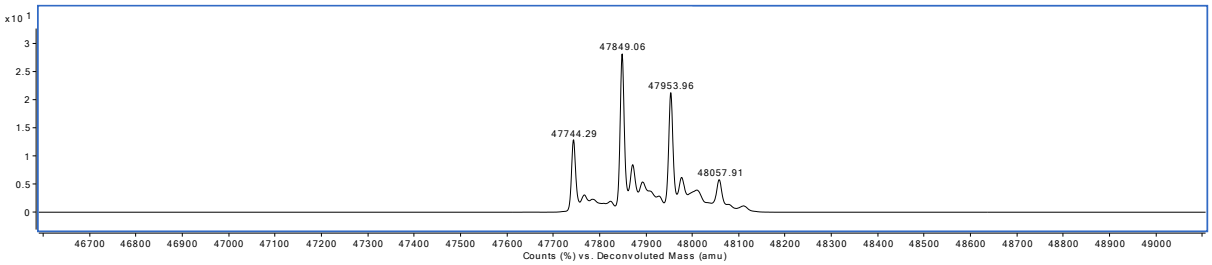
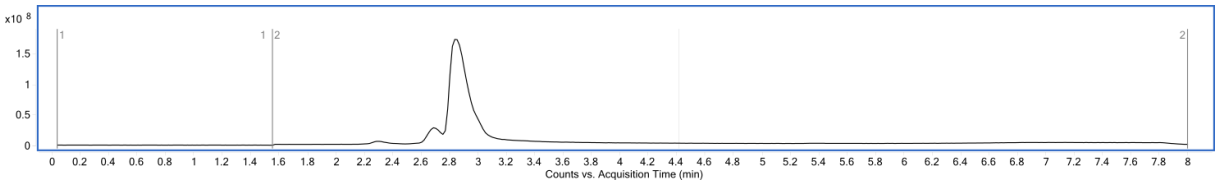


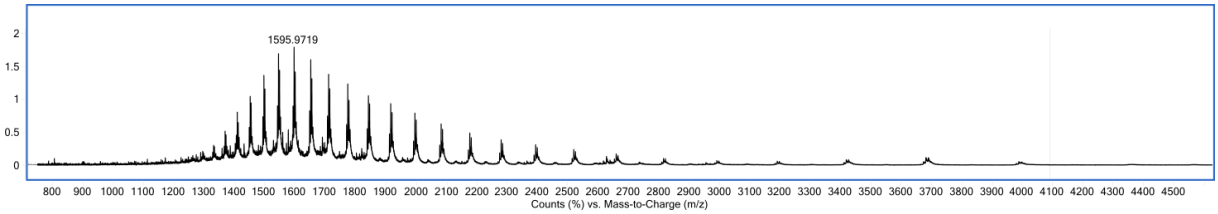
Figure S69: LCMS analysis of conjugate **18**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions), 48058 Da (4 additions). Observed: 47744 Da, 47849 Da, 47954 Da, 48058 Da.

6 hours

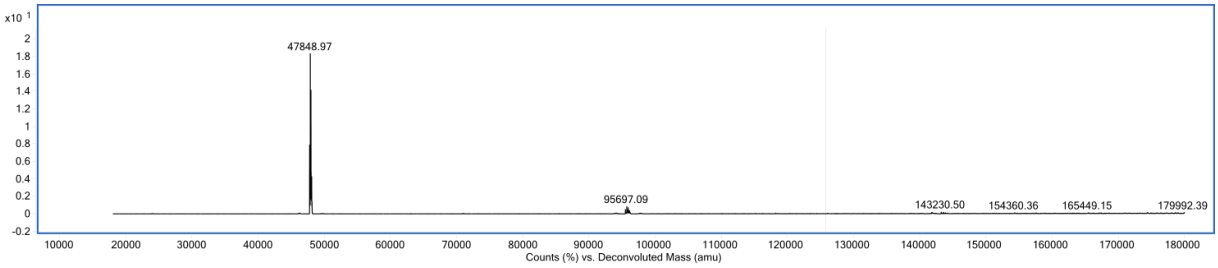
(a)



(b)



(c)



(d)

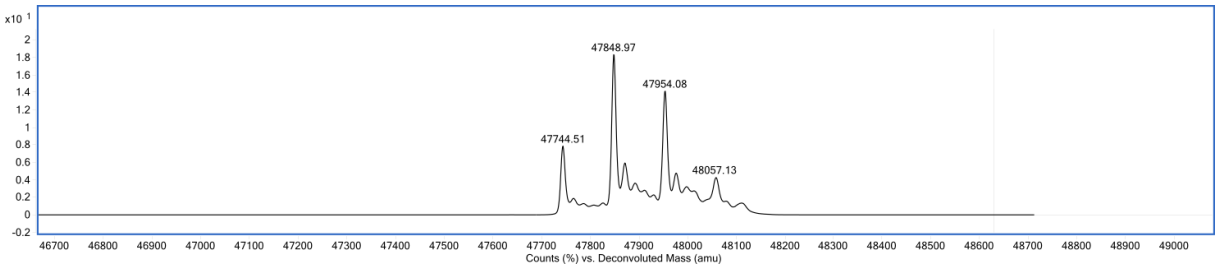
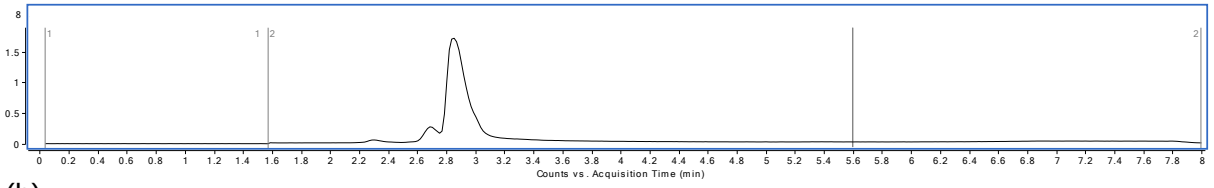


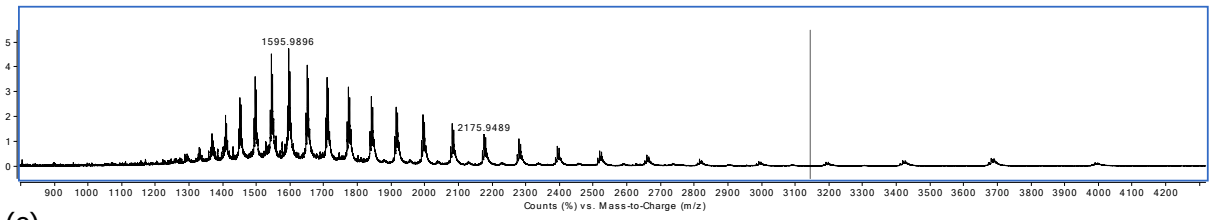
Figure S70: LCMS analysis of conjugate **18**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions), 48058 Da (4 additions). Observed: 47744 Da, 47849 Da, 47954 Da, 48057 Da.

24 hours

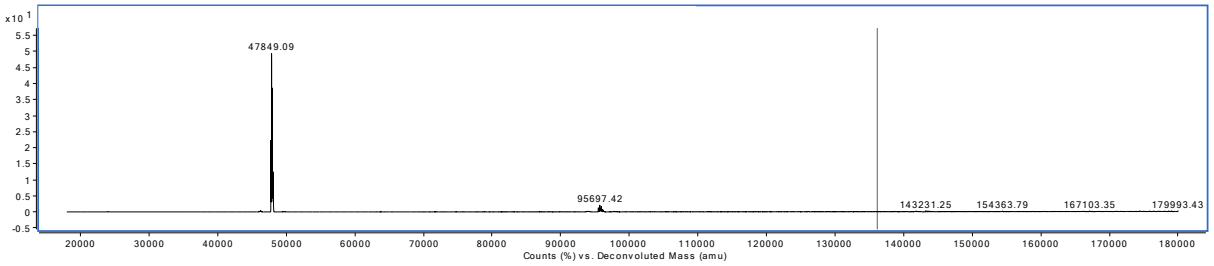
(a)



(b)



(c)



(d)

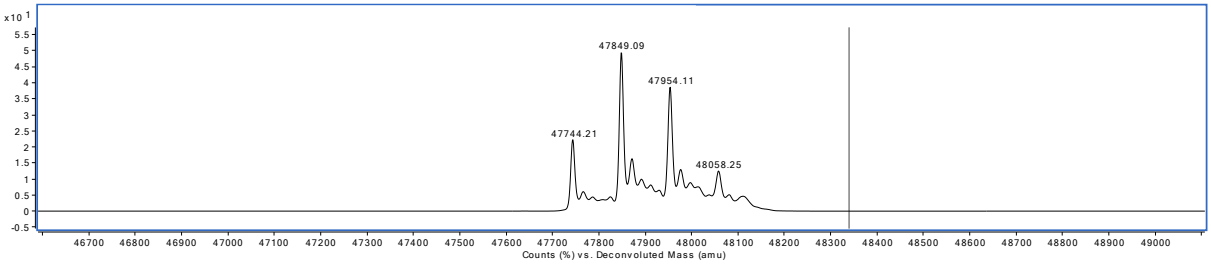
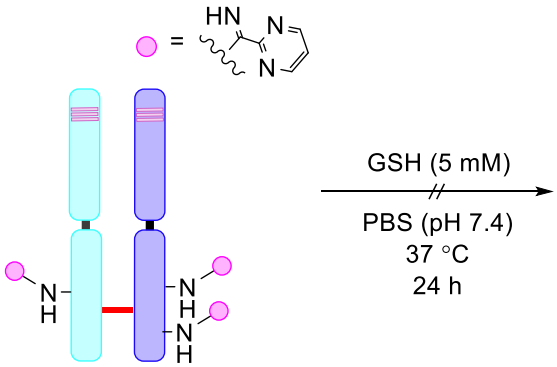


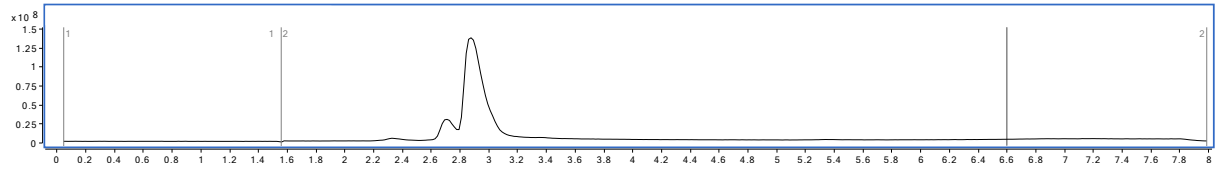
Figure S71: LCMS analysis of conjugate **18**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions), 48058 Da (4 additions). Observed: 47744 Da, 47849 Da, 47954 Da, 48058 Da.



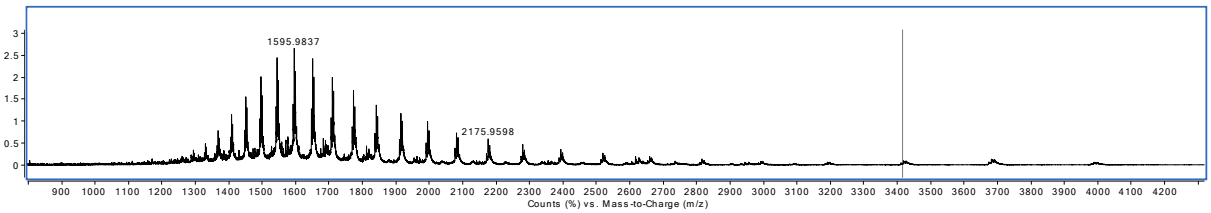
Fab conjugate **18** (50 µL, 7 µM) was buffer exchanged using desalting columns (7000 MWCO, ZebaSpin, Thermo Scientific) into GSH containing buffer (5 mM GSH, in PBS pH 6.5) and maintained at 37 °C for 24 h. Excess reagents were removed using desalting columns (7000 MWCO, ZebaSpin, Thermo Scientific) prior to LCMS analysis.

0 hours

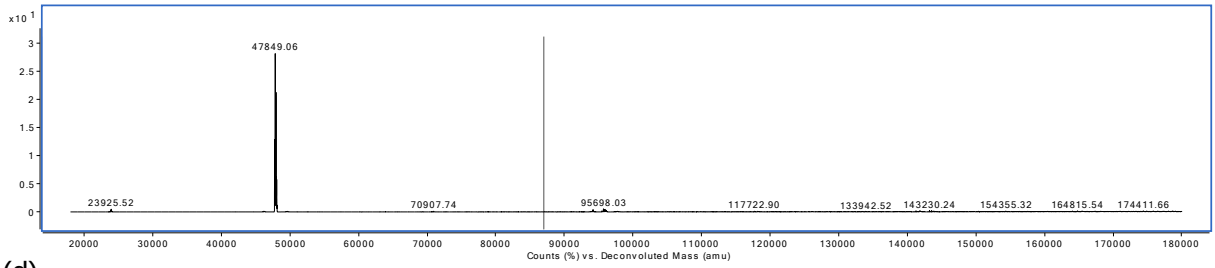
(a)



(b)



(c)



(d)

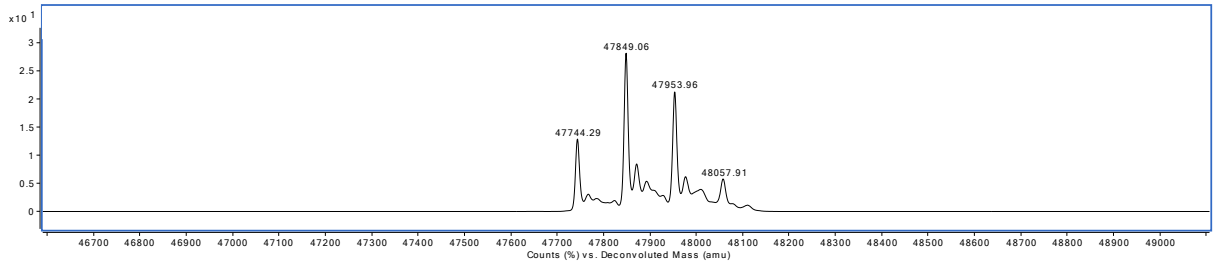
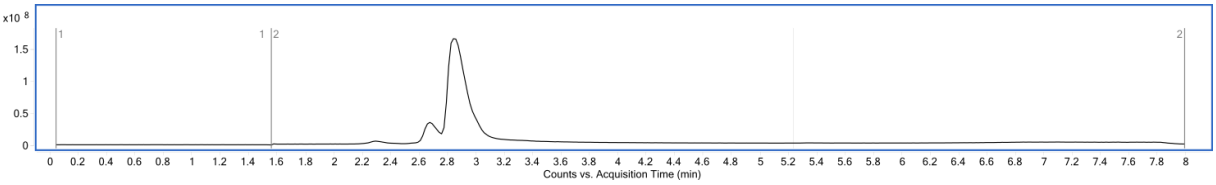


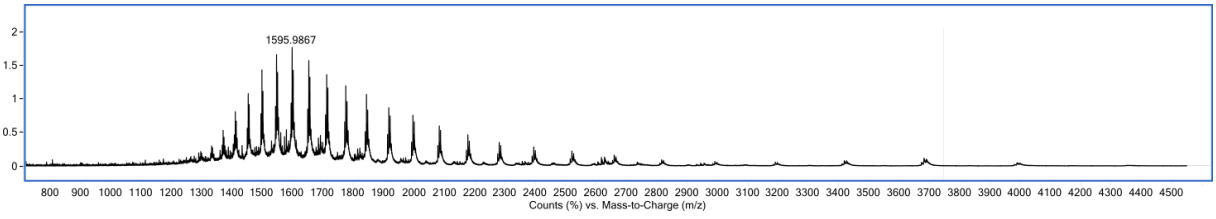
Figure S72: LCMS analysis of conjugate **18**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions), 48058 Da (4 additions). Observed: 47744 Da, 47849 Da, 47954 Da, 48058 Da.

6 hours

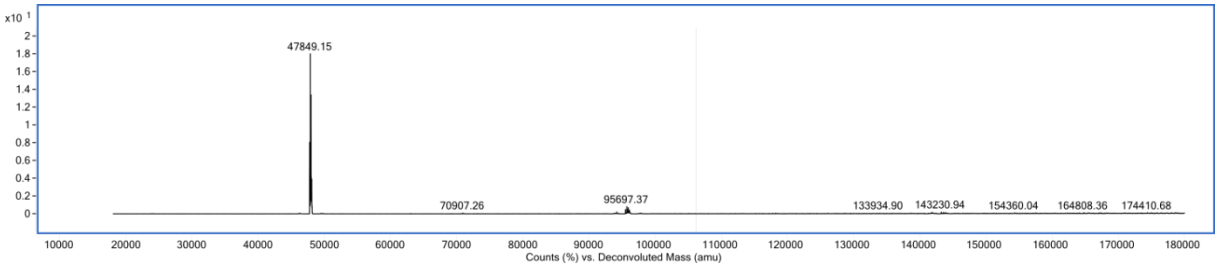
(a)



(b)



(c)



(d)

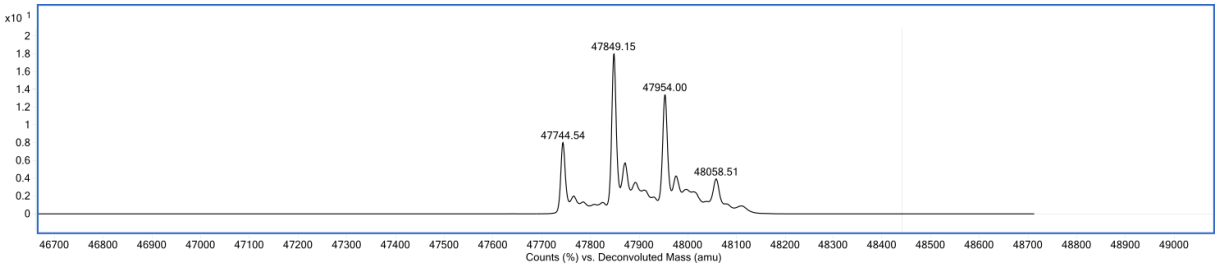
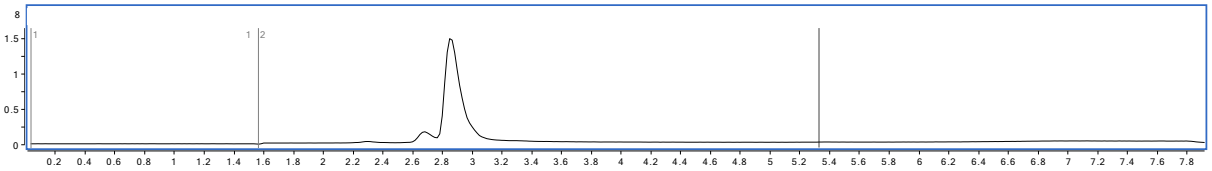


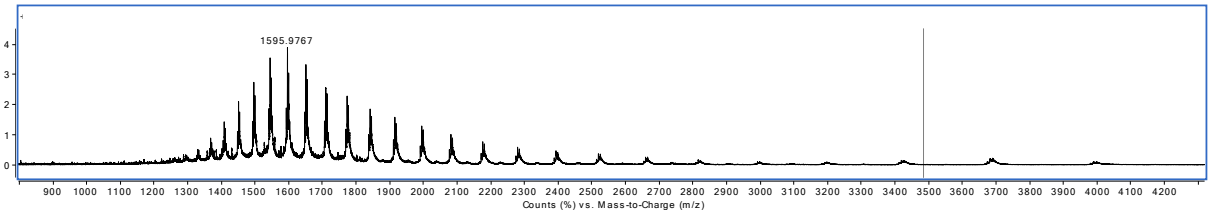
Figure S73: LCMS analysis of conjugate **18**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions), 48058 Da (4 additions). Observed: 47744 Da, 47849 Da, 47954 Da, 48058 Da.

24 hours

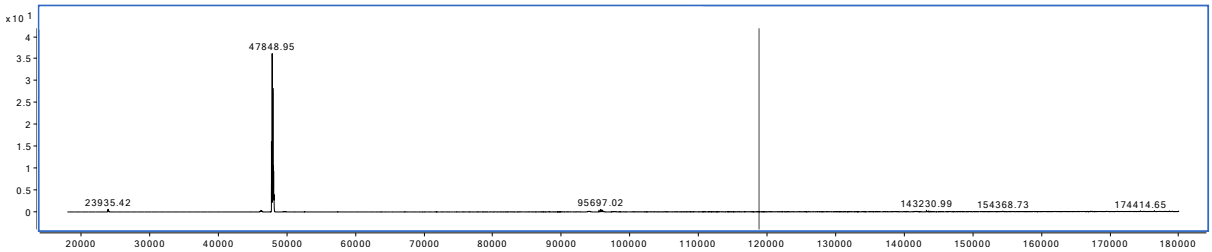
(a)



(b)



(c)



(d)

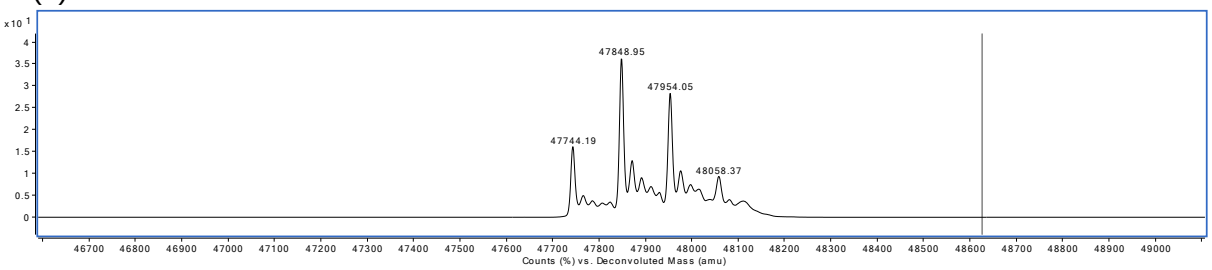
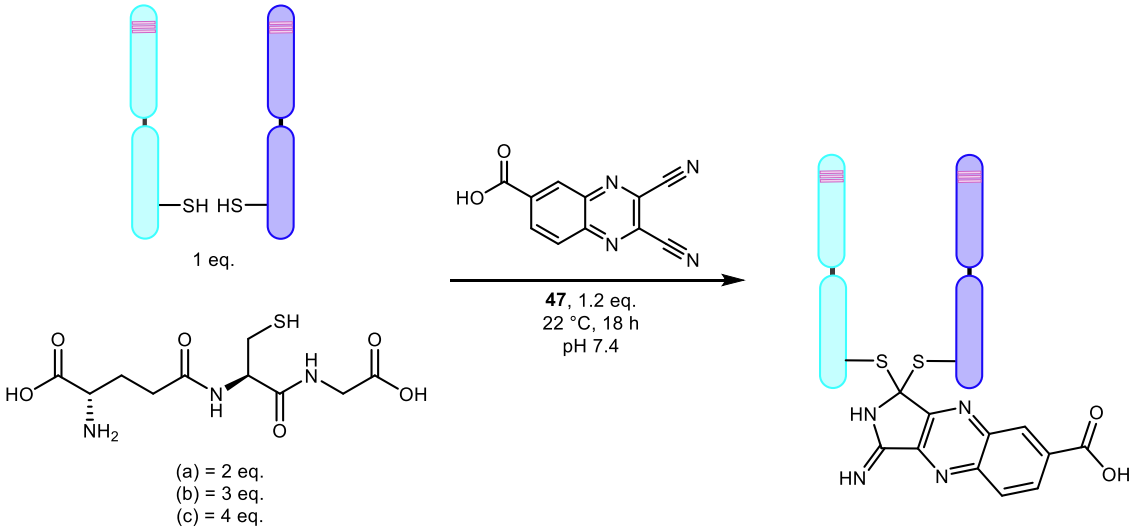


Figure S74: LCMS analysis of conjugate **18**; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of pyrimidine CLT conjugate: 47743 Da (1 addition), 47848 Da (2 additions), 47953 Da (3 additions), 48058 Da (4 additions). Observed: 47744 Da, 47849 Da, 47954 Da, 48058 Da.

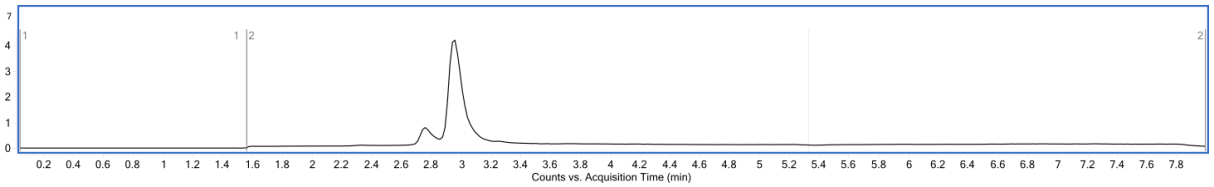
Disulfide competition experiment (GSH Vs Fab)- 2,3-dicyanoquinoxaline-6-carboxylic acid



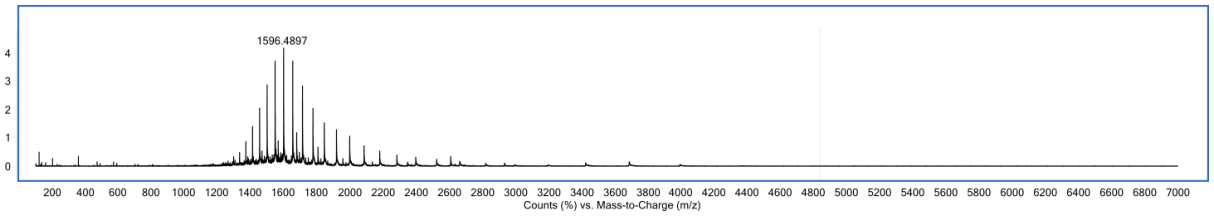
Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with TCEP (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. To this was added glutathione (10 mM in dH_2O) followed closely by 2,3-dicyanoquinoxaline-6-carboxylic acid (0.72 µL, 10 mM in DMF, 1.2 equiv.), the reaction was then incubated at 22 °C for 18 h. Finally, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

2 equiv. GSH

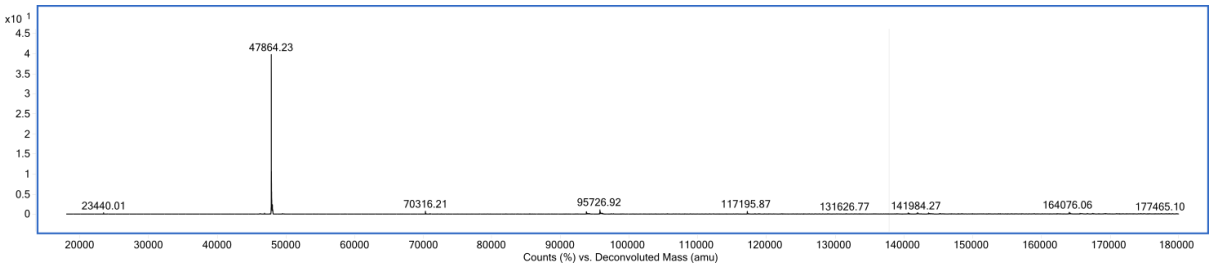
(a)



(b)



(c)



(d)

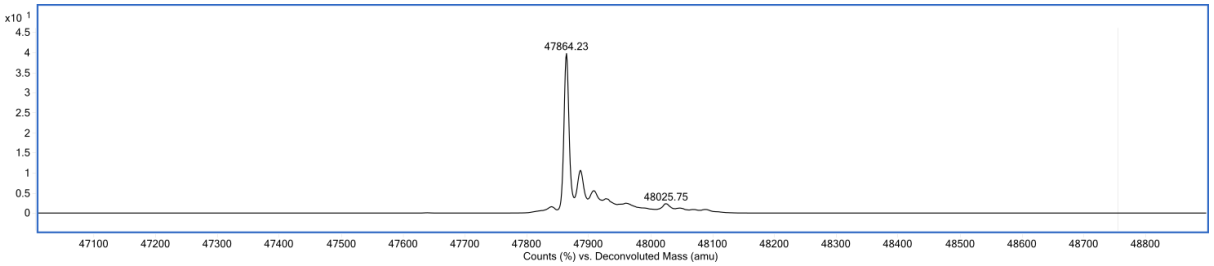
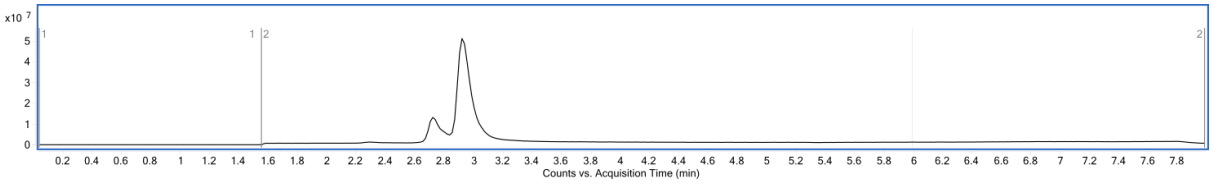


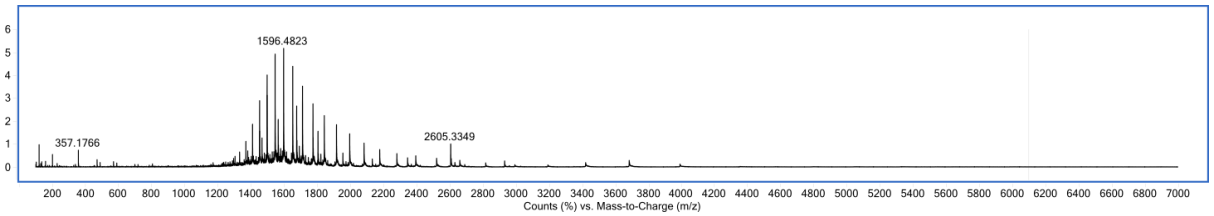
Figure S75: LCMS analysis of disulfide competition experiment; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47864 Da.

3 equiv. GSH

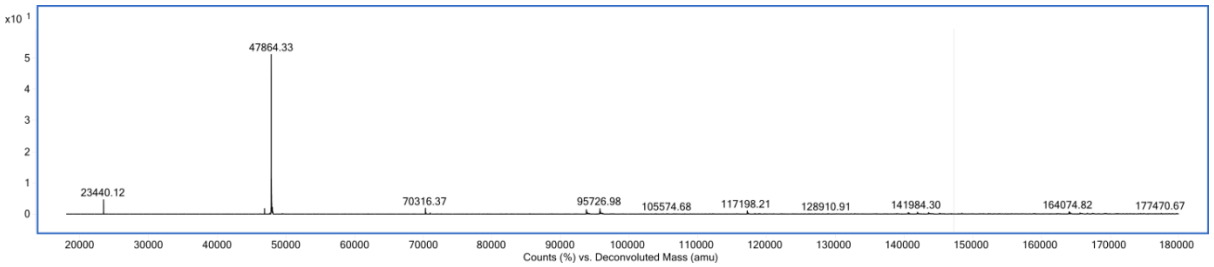
(a)



(b)



(c)



(d)

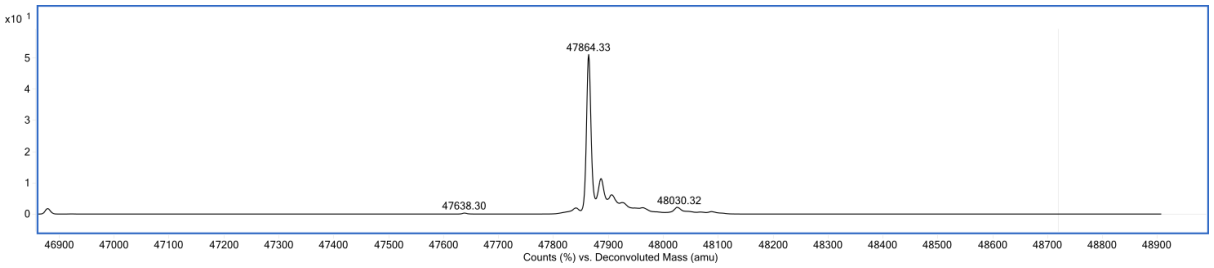
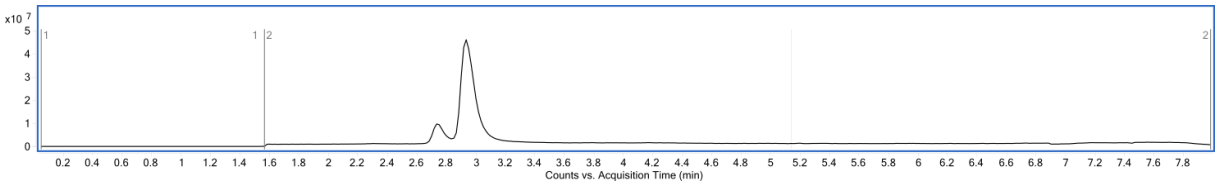


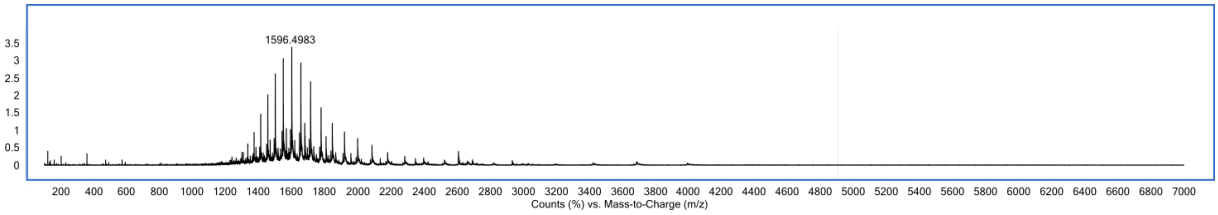
Figure S76: LCMS analysis of disulfide competition experiment; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47864 Da.

4 equiv. GSH

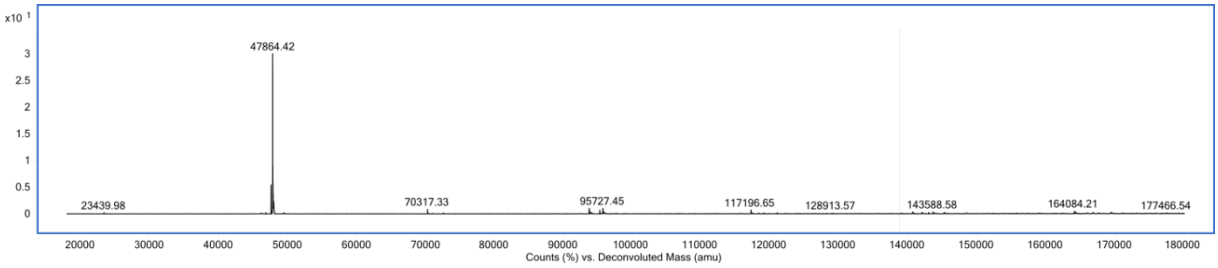
(a)



(b)



(c)



(d)

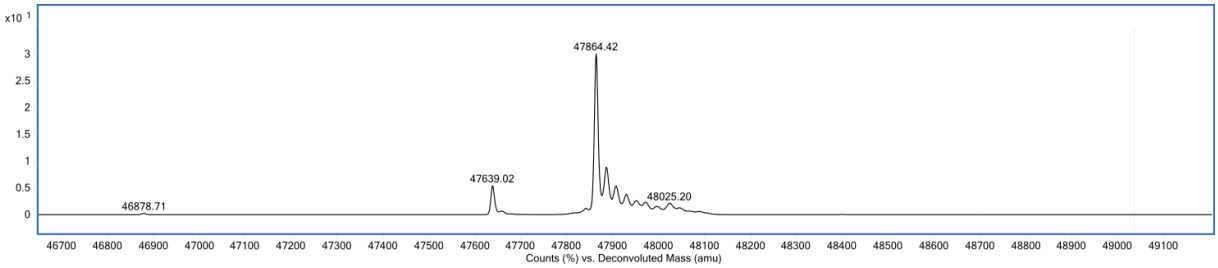
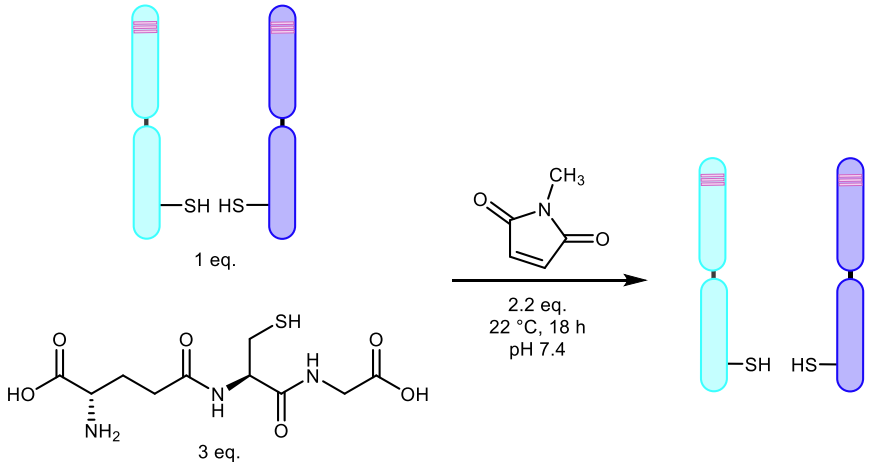


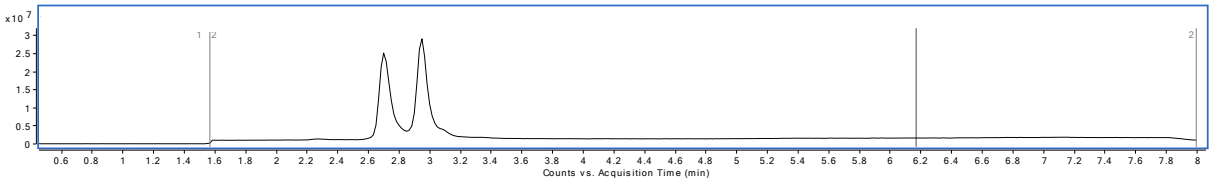
Figure S77: LCMS analysis of disulfide competition experiment; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of native Fab: 47639, expected mass of re-bridged Fab with 2,3-dicyanoquinoxaline-6-carboxylic acid: 47862 Da. Observed: 47639 Da, 47864 Da.

Disulfide competition experiment (GSH Vs Fab)- NMM

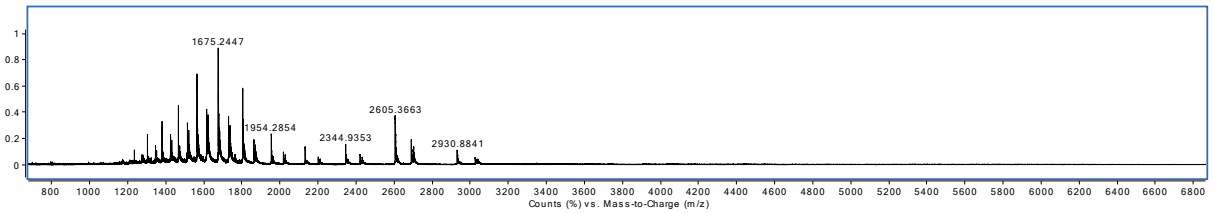


Fab (20 µL, 150 µM, 7.15 mg/mL) in conjugation buffer was reduced with TCEP (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. To this was added glutathione (0.9 µL, 10 mM in dH_2O, 3 equiv.) followed closely by *N*-methyl maleimide (0.72 µL, 10 mM in DMF, 2.4 equiv.), the reaction was then incubated at 22 °C for 18 h. Finally, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 68590 $M^{-1}cm^{-1}$.

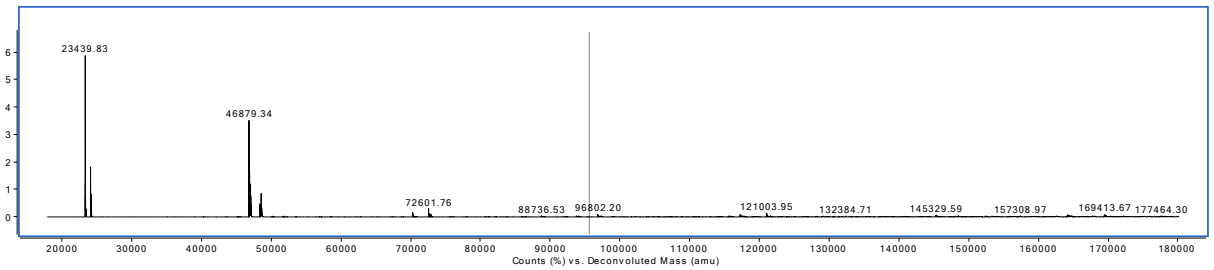
(a)



(b)



(c)



(d)

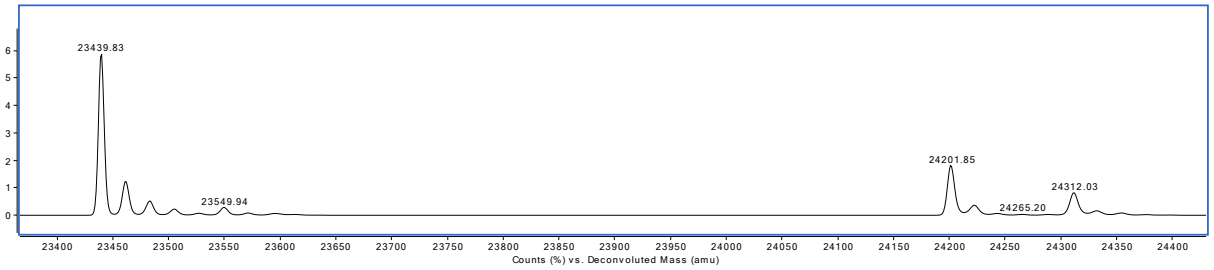
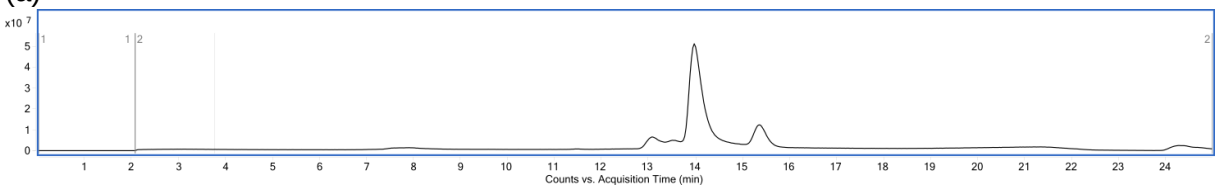


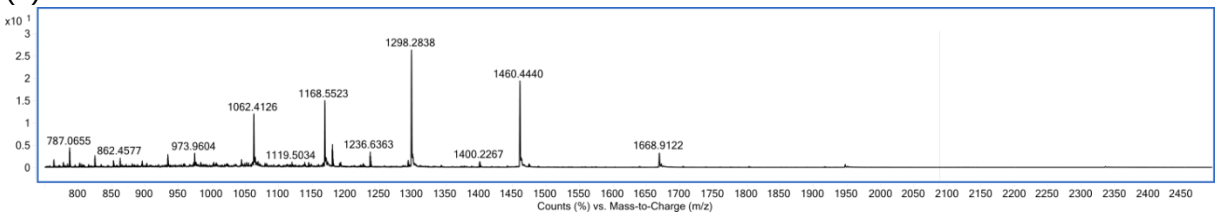
Figure S78: LCMS analysis of N-methyl maleimide competition experiment; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum, d) zoomed in deconvoluted ion series mass spectrum; Expected mass of capped Fab with *N*-methyl maleimide: **LC** 23549 Da, **HC** 24311 Da. Observed: Native **LC** 23439 Da, Modified **LC** 23549 Da, Native **HC** 24201 Da, Modified **HC** 24312 Da. Peaks around 46879 Da (2 x 23440 Da) are proposed to be artifacts formed during ionisation.

Thioredoxin (Trx)

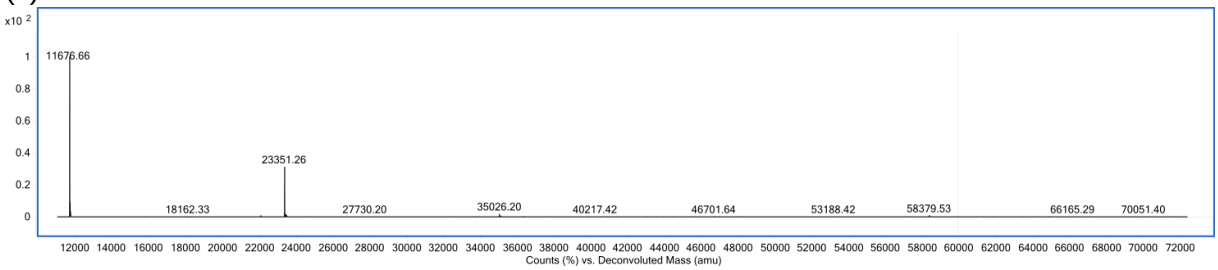
(a)



(b)



(c)



(d)

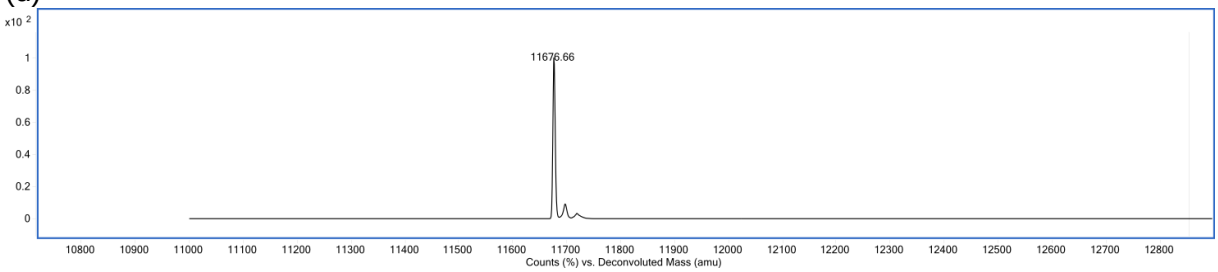
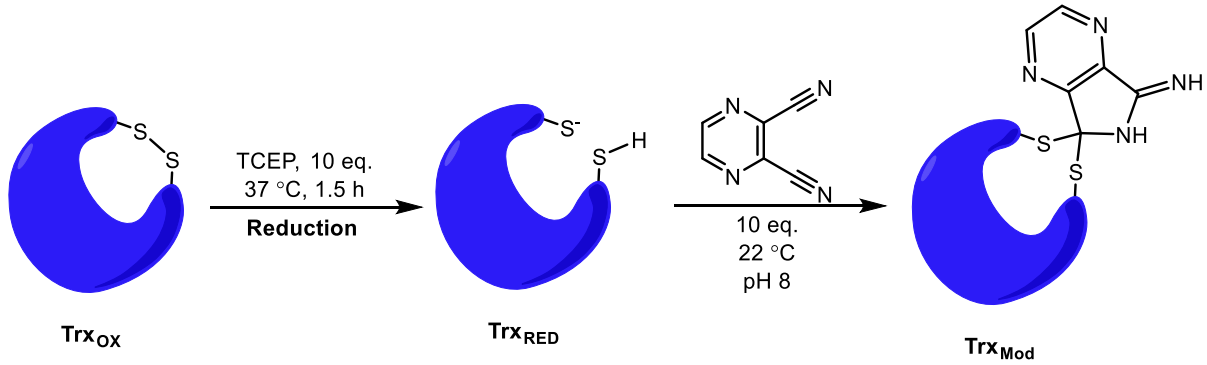


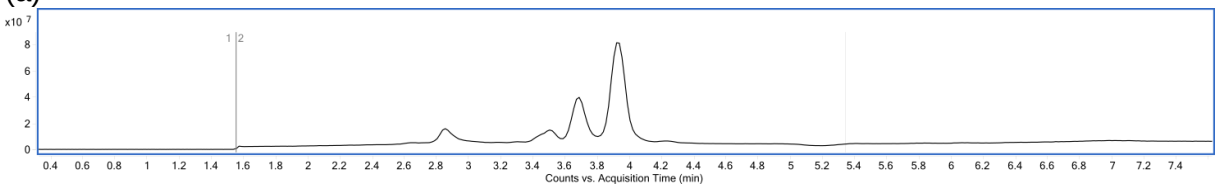
Figure S79: LCMS analysis of conjugate Native Trx; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; native Trx expected 11675 Da. Observed : 11677 Da, Peak at 23351 Da (2 x 11675.5 Da) are proposed to be artifacts formed during ionisation.

Reaction of Trx with pyrazine-2,3-dicarbonitrile

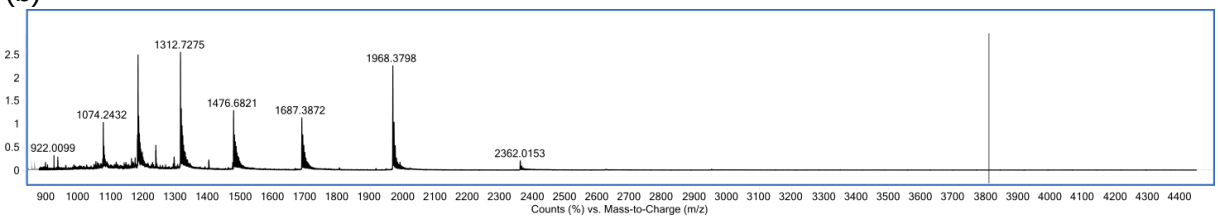


Thioredoxin (30 µL, 100 µM) in BBS (pH 8) was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Pyrazine-2,3-dicarbonitrile (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2.5 h. Lastly, excess reagent was removed, and the sample was desalted (7 kDa MWCO, ZebaSpin) prior to LCMS analysis. Concentration was determined photometrically using ε280= 14180 $M^{-1}cm^{-1}$.

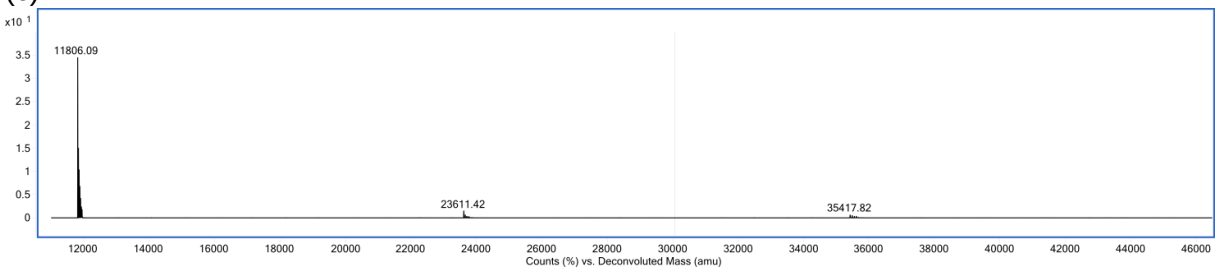
(a)



(b)



(c)



(d)

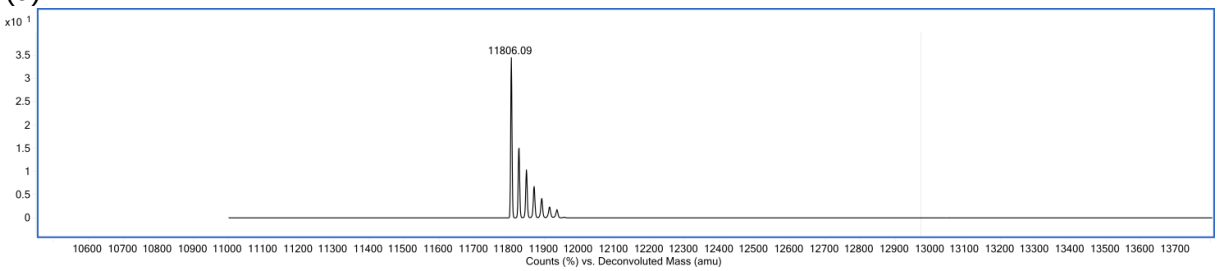
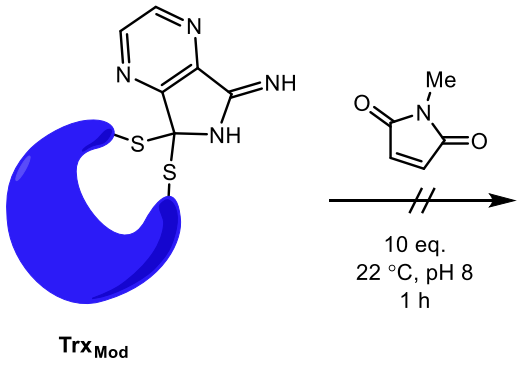


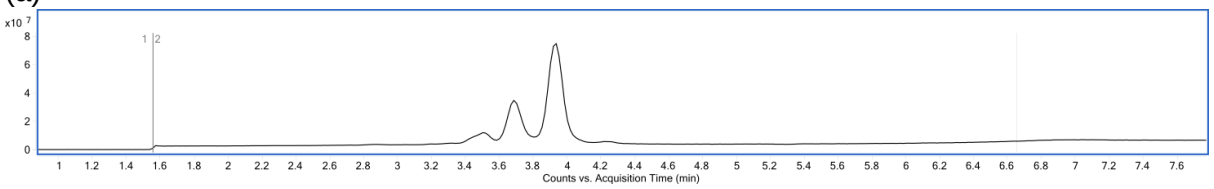
Figure S80: LCMS analysis of modified Trx; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Pyrazine-2,3-dicarbonitrile modified Trx expected 11805 Da. Observed : 11806 Da.

Modified Trx stability studies

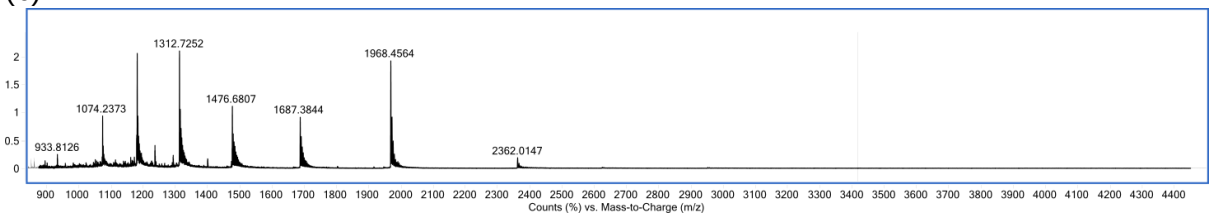


Thioredoxin (30 µL, 100 µM) in BBS (pH 8) was reduced with tris(2-carboxyethyl)phosphine (TCEP) (2.0 µL, 15 mM in dH_2O, 10 equiv.) The mixture was incubated at 37 °C for 1.5 h, 300 rpm. Pyrazine-2,3-dicarbonitrile (2.0 µL, 15 mM in DMF, 10 equiv.) was added and incubated at 22 °C for 2.5 h. Excess reagent was removed, and tris(2-carboxyethyl)phosphine (TCEP) (15 mM in dH_2O, 10 equiv.) was added. The mixture was then incubated at 37 °C for 1.5 h, 300 rpm. NMM (20 mM in DMF, 10 equiv.) was added for 1 h at 22 °C. The reaction was then purified *via* ultrafiltration (10 kDa MWCO) into HPLC grade water prior to LCMS analysis. Concentration was determined photometrically using ε280= 14180 $M^{-1}cm^{-1}$.

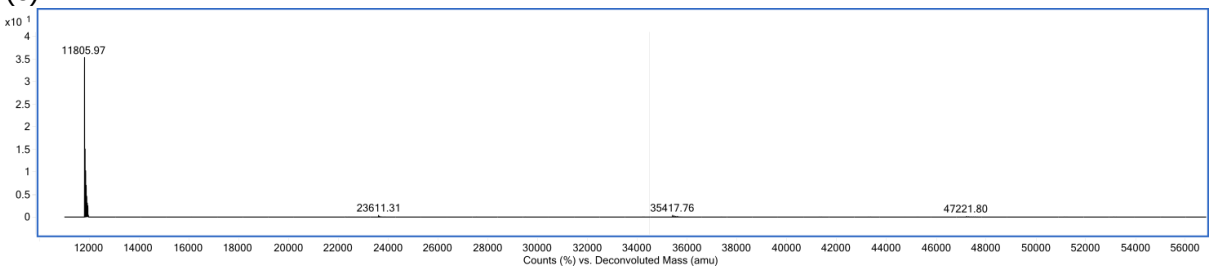
(a)



(b)



(c)



(d)

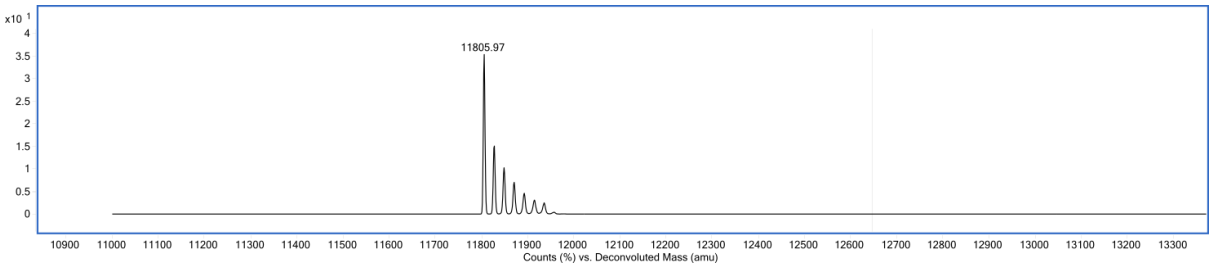
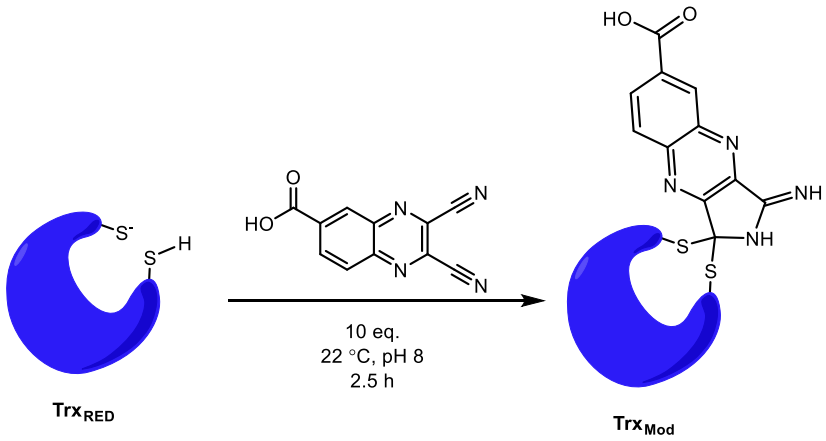


Figure S81: LCMS analysis of modified Trx; a) TIC, b) non-deconvoluted ion-series, c) full range deconvoluted ion series mass spectrum; Pyrazine-2,3-dicarbonitrile modified Trx expected 11805 Da. Observed : 11806 Da.



Bioactivity measurement of Trx and modified Trx (S82)

Trx and modified Trx activity were measured by following the reduction of insulin previously described.[258] The standard Trx (or modified Trx) assay mixture, prepared in 200 µL overall

volume, contained sodium phosphate buffer (50 mM, pH 7.0), 1 mM EDTA, 0.15 mM human insulin (5808 Da) and 1 mM dithiothreitol. The amounts of Trx or modified Trx were varied by concentration (0.8 µM - 4 µM). The increase in turbidity from the reduction of insulin was monitored at 650 nm at RT by a Shimadzu UV-2700i UV-Vis Spectrophotometer with a 1.5x1.5 mm light-path black quartz spectrophotometer cell. The slopes from the kinetic data (ΔmAU/min), were plotted as a function of the concentration of protein.

CLT SDS-page (S83)

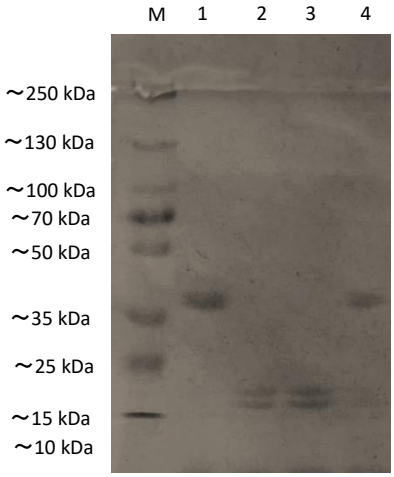


Figure S83: SDS-PAGE of cysteine to lysine transfer (CLT) on trastuzumab Fab; M – molecular marker, 1 – Native Fab, 2 – Reduced Fab, 3 – CLT step 1, transfer, 4 – CLT step 2, reoxidation.

Bridge SDS-page (S84)

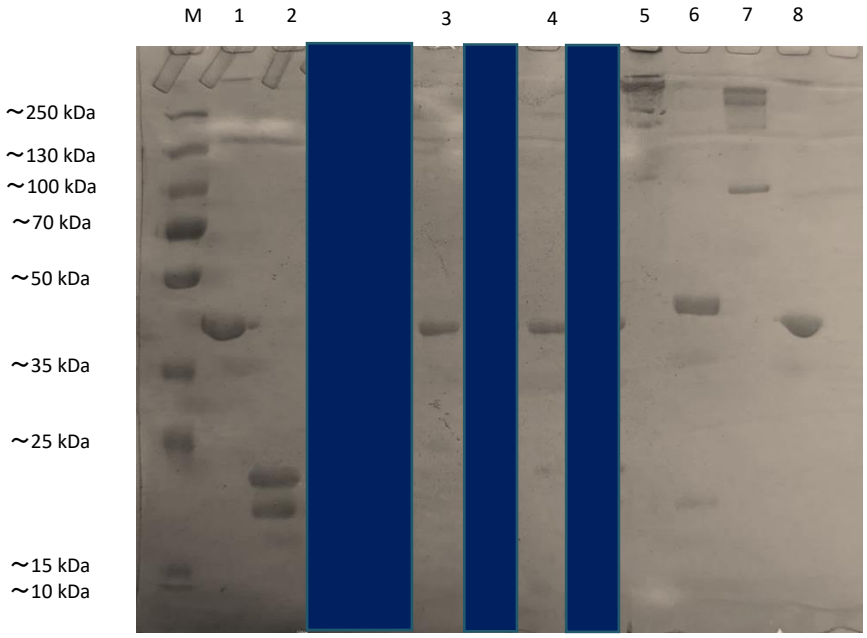


Figure S84: SDS-PAGE of cysteine to lysine transfer (CLT) on trastuzumab Fab; M – molecular marker, 1 – Native Fab, 2 – Reduced Fab, 3 – Re-bridged Fab with **3**, 4 – Re-bridged Fab with **12**, 5 – Full antibody, 6 – Reduced full antibody, 7 – Full antibody re-bridged with **12**, 8 – re-bridged Fab with **11**.

Serum Stability Study (S85)[259]

Fab conjugates **20** and **52** were prepared as 0.2 mg/mL solutions in PBS 140 mM sodium chloride 12 mM sodium phosphates and 2 mM sodium azide at pH 7.4. The conjugates were diluted with 50 % of human blood serum to give a final a concentration of 0.1 mg/mL of 5 or 12 and 1 mM of sodium azide. One aliquot (50 µL) for each conjugate was taken, flash frozen and stored at -20 °C. The remaining solution was incubated at 37 °C under mild shaking (300 rpm) and covered from light. Aliquots (5 µL) were taken at 1, 3, and 5, flash frozen and stored at -20 °C. Samples were diluted with Coomassie Brilliant Blue R-250 (3 µL) and analysed *via* SDS-page (samples were run at 150 V for 80 min). The fluorescent bands were visualised using an AZURE 200® (Epi Blue).

(A) **(B)**

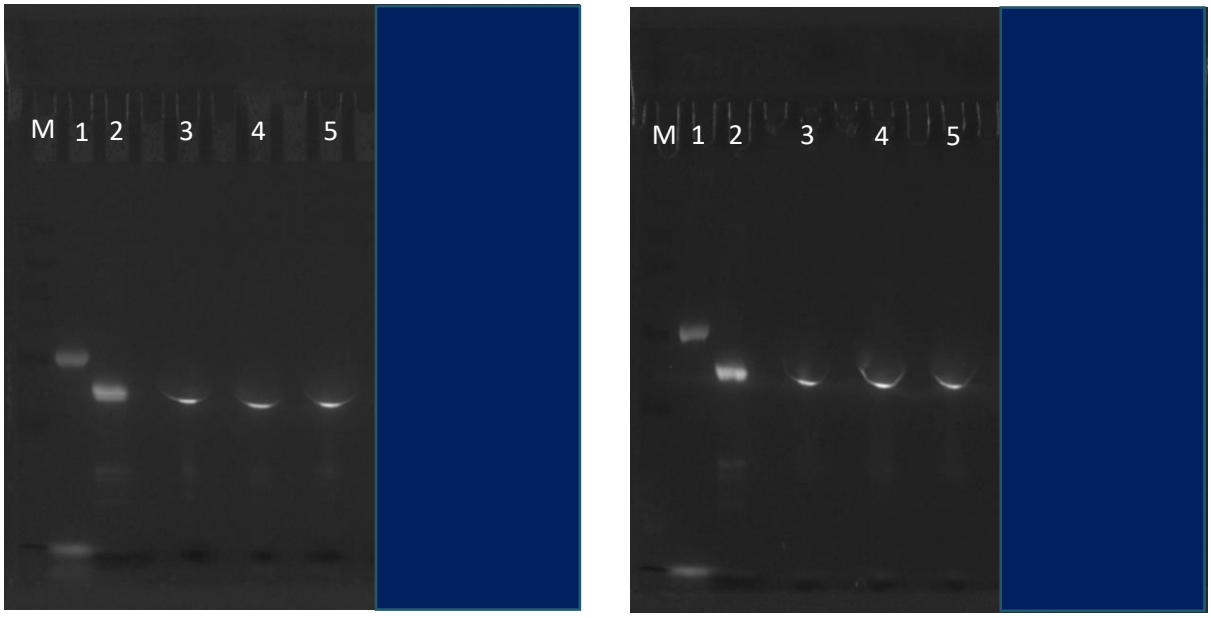


Figure S85: SDS-PAGE of serum stability study; (A) M – molecular marker, 1 – fluorescent HSA reference, 2 – fluorescent CLT Fab reference (conjugate **20**), 3 – Day one, 4 – Day three, 5 – Day five. (B) M – molecular marker, 1 – fluorescent HSA reference, 2 – fluorescent bridged Fab reference (conjugate **52**), 3 – Day one, 4 – Day three, 5 – Day five

Enzyme-linked immunosorbent assay (ELISA) – Trastuzumab against HER2 (S86)

A 96-well plate was coated for 16 h at 4 °C with HER2 (Sino Biological, 100 µL/well, 0.25 µg/mL solution in PBS), all wells except row D and H. After washing (3 × 0.1% Tween® 20 in PBS, followed by 3 × PBS), wells were blocked for 1 h at room temperature with 5% Marvel milk powder (Premier foods) in PBS (200 µL/well). The wells were then washed (3 × 0.1% Tween® 20 in PBS, followed by 3 × PBS), and the following dilutions of Fab conjugate were applied: 810 nM, 270 nM, 90 nM, 30 nM, 10 nM, 3.33 nM, 1.11 nM, 0.37 nM, 0.123 nM, 0.0412 nM, 0.0137 nM, prepared in 1% Marvel solution in PBS (100 µL/well), these were added to rows A to C – 1-11 and E to G – 1-11. The assay was then incubated at room temperature for 1 h. After 1 h, the plate was washed (3 × 0.1% Tween® 20 in PBS, followed by 3 × PBS), and the detection antibody (Anti-Human IgG, Fab specific-horseradish peroxidase (HRP) antibody, Sigma Aldrich, 1:5000 in 1% Marvel solution in 0.1% Tween® 20 in PBS) was added to the whole plate (100 µL/well), and incubated for 1 h at room temperature. After that, the plate was washed (3 × 0.1% Tween® 20 in PBS, followed by 3 × PBS), and o-phenylenediamine hydrochloride (Sigma-Aldrich, 100 µL/well, 0.5 mg/mL in a phosphate-citrate buffer with sodium perborate) was added to the whole plate, left for 15-30 min in the dark (monitor the colour development), room temperature. Once a yellow-orange colour was observed, the reaction was stopped by addition of HCl to the whole plate (4 M, 50 µL/well). Absorbance was immediately measured at 450 nm and was corrected by subtracting the average of negative controls (i.e. PBS had been added to some of the wells instead of HER2 or instead of the 66 samples). Each sample was tested in triplicate and errors are shown as the standard deviation of the average. ELISA data was analysed with Graphpad Prism 7.03 (using equation Sigmoidal, 4PL, X is log(concentration)) and the values have been normalised.

7. References

1 J. Harrow, A. Nagy, A. Reymond, T. Alioto, L. Patthy, S. E. Antonarakis and R. Guigó, *Genome Biol.*, 2009, **10**, 201.

2 S. L. Salzberg, *BMC Biol.*, 2018, **16**, 10–12.

3 E. A. Ponomarenko, E. V. Poverennaya, E. V. Ilgisonis, M. A. Pyatnitskiy, A. T. Kopylov, V. G. Zgoda, A. V. Lisitsa and A. I. Archakov, *Int. J. Anal. Chem.*, , DOI:10.1155/2016/7436849.

4 R. Aebersold, J. N. Agar, I. J. Amster, M. S. Baker, C. R. Bertozzi, E. S. Boja, C. E. Costello, B. F. Cravatt, C. Fenselau, B. A. Garcia, Y. Ge, J. Gunawardena, R. C. Hendrickson, P. J. Hergenrother, C. G. Huber, A. R. Ivanov, O. N. Jensen, M. C. Jewett, N. L. Kelleher, L. L. Kiessling, N. J. Krogan, M. R. Larsen, J. A. Loo, R. R. Ogorzalek Loo, E. Lundberg, M. J. Maccoss, P. Mallick, V. K. Mootha, M. Mrksich, T. W. Muir, S. M. Patrie, J. J. Pesavento, S. J. Pitteri, H. Rodriguez, A. Saghatelian, W. Sandoval, H. Schlüter, S. Sechi, S. A. Slavoff, L. M. Smith, M. P. Snyder, P. M. Thomas, M. Uhlén, J. E. Van Eyk, M. Vidal, D. R. Walt, F. M. White, E. R. Williams, T.

Wohlschlager, V. H. Wysocki, N. A. Yates, N. L. Young and B. Zhang, *Nat. Chem. Biol.*, 2018, **14**, 206–214.

5 C. T. Walsh, S. Garneau-Tsodikova and G. J. Gatto, *Angew. Chemie - Int. Ed.*, 2005, **44**, 7342–7372.

6 L. Wan, K. Xu, Z. Chen, B. Tang and H. Jiang, *Front. Cell. Neurosci.*, , DOI:10.3389/fncel.2018.00290.

7 D. G. Knorre, N. V Kudryashova and T. S. Godovikova, *Acta Naturae*, 2009, **1**, 29–51.

8 S. Ramazi and J. Zahiri, *Database*, 2021, **2021**, 1–20.

9 H. Xu, Y. Wang, S. Lin, W. Deng, D. Peng, Q. Cui and Y. Xue, *Genomics, Proteomics Bioinforma.*, 2018, **16**, 244–251.

10 A. F. L. Schneider and C. P. R. Hackenberger, *COBIOT*, 2017, 48, 61–68.

11 F. Li and R. I. Mahato, *Mol. Pharm.*, 2017, **14**, 1321–1324.

12 W. Hoffman, F. G. Lakkis and G. Chalasani, *Clin. J. Am. Soc. Nephrol.*, 2016, **11**, 137–154.

13 D. N. Forthal, *Antibodies Infect. Dis.*, 2015, 25–48.

14 J. A. Trapani, *Genome Biol.*, 2001, **2**, 1–7.

15 D. R. Anderson, A. Grillo-Löpez, C. Varns, K. S. Chambers and N. Hanna, *Biochem. Soc. Trans.*, 1997, **25**, 705–708.

16 M. X. Sliwkowski, J. A. Lofgren, G. D. Lewis, T. E. Hotaling, B. M. Fendly and J. A. Fox, *Semin. Oncol.*, 1999, **26**, 60–70.

17 M. Z. Tay, K. Wiehe and J. Pollara, *Front. Immunol.*, 2019, **10**, 1–18.

18 G. Vidarsson, G. Dekkers and T. Rispens, *Front. Immunol.*, 2014, **5**, 1–17.

19 M. L. Chiu, D. R. Goulet, A. Teplyakov and G. L. Gilliland, *Antibodies*, 2019, **8**, 55.

20 V. Bayer, *Semin. Oncol. Nurs.*, 2019, **35**, 150927.

21 M. J. Glennie and J. G. J. Van De Winkel, *Drug Discov. Today*, 2003, **8**, 503–510.

22 J. C. Almagro, T. R. Daniels-Wells, S. M. Perez-Tapia and M. L. Penichet, *Front. Immunol.*, , DOI:10.3389/fimmu.2017.01751.

23 Y. Safdari, S. Farajnia, M. Asgharzadeh and M. Khalili, *Biotechnol. Genet. Eng. Rev.*, 2013, **29**, 175–186.

24 F. A. Harding, M. M. Stickler, J. Razo and R. B. DuBridge, *MAbs*, 2010, **2**, 256–265.

25 H. Li, W. Yuan, S. Bin, G. Wu, P. Li, M. Liu, J. Yang, X. Li, K. Yang and H. Gu, *Am. J. Cancer Res.*, 2020, **10**, 688–703.

26 P. R. Pohlmann, I. A. Mayer and R. Mernaugh, *Clin. Cancer Res.*, 2009, **15**, 7479–7491.

27 L. Urquhart, *Nat. Rev. Drug Discov.*, 2021, **20**, 253.

28 H. Kaplon, A. Chenoweth, S. Crescioli and J. M. Reichert, *MAbs*, , DOI:10.1080/19420862.2021.2014296.

29 A. Shahidian, M. Ghassemi, J. Mohammadi and M. Hashemi, eds. A. Shahidian, M. Ghassemi, J. Mohammadi and M. B. T.-B.-E. A. to C. D. and T. Hashemi, Academic

Press, 2020, pp. 69–114.

30 A. L. Nelson, *MAbs*, 2010, **2**, 77–83.

31 K. T. Xenaki, S. Oliveira and P. M. P. van Bergen en Henegouwen, *Front. Immunol.*, , DOI:10.3389/fimmu.2017.01287.

32 M. M. Schmidt and K. D. Wittrup, *Mol. Cancer Ther.*, 2009, **8**, 2861–2871.

33 P. Holliger and P. J. Hudson, *Nat. Biotechnol.*, 2005, **23**, 1126–1136.

34 R. V. Kholodenko, D. V. Kalinovsky, I. I. Doronin, E. D. Ponomarev and I. V. Kholodenko, *Curr. Med. Chem.*, 2019, **26**, 396–426.

35 R. Boushaba, P. Kumpalume and N. K. H. Slater, *Biotechnol. Prog.*, 2003, **19**, 1176–1182.

36 H. W. Schroeder and L. Cavacini, *J. Allergy Clin. Immunol.*, 2010, **125**, S41–S52.

37 A. Frenzel, M. Hust and T. Schirrmann, 2013, **4**, 1–20.

38 R. E. Bird, K. D. Hardman, J. W. Jacobson, S. Y. D. Johnson, B. M. Kaufman, S. Lee, T. Lee, S. H. Pope, G. S. Riordan and M. Whitlow, .

39 C. Duranti and A. Arcangeli, .

40 A. P. Chapman, *Adv. Drug Deliv. Rev.*, 2002, **54**, 531–545.

41 A. P. Herrington-Symes, M. Farys, H. Khalili and S. Brocchini, *Adv. Biosci. Biotechnol.*, 2013, **04**, 689–698.

42 R. Adams, L. Griffin, J. E. Compson, M. Jairaj, T. Baker, T. Ceska, S. West, O. Zaccheo, E. Davé, A. D. G. Lawson, D. P. Humphreys and S. Heywood, *MAbs*, 2016, **8**, 1336–1346.

43 R. Lu, Y. Hwang, I. Liu, C. Lee, H. Tsai, H. Li and H. Wu, 2020, 1–30.

44 *N. Engl. J. Med.*, 1994, **330**, 956–961.

45 Y. Y. Syed, *Am. J. Cardiovasc. Drugs*, 2016, **16**, 297–304.

46 G. Zugmaier, M. Klinger, M. Schmidt and M. Subklewe, *Mol. Immunol.*, 2015, **67**, 58–66.

47 N. A. Yannuzzi and K. B. Freund, *Clin. Ophthalmol.*, 2019, **13**, 1323–1329.

48 L. Ducry and B. Stump, *Bioconjugate Chem.*, 2010, **21**, 5–13.

49 A. M. Wu and P. D. Senter, *Nat. Biotechnol.*, 2005, **23**, 1137–1146.

50 C. D. Spicer and B. G. Davis, *Nat. Commun.*, , DOI:10.1038/ncomms5740.

51 G. Casi and D. Neri, *J. Med. Chem.*, 2015, **58**, 8751–8761.

52 J. R. Adair, P. W. Howard, J. A. Hartley, D. G. Williams and K. A. Chester, *Expert Opin. Biol. Ther.*, 2012, **12**, 1191–1206.

53 K. Strebhardt and A. Ullrich, *Nat. Rev. Cancer*, 2008, **8**, 473–480.

54 N. Joubert, A. Beck, C. Dumontet and C. Denevault-sabourin, 2020, 1–31.

55 P. A. Trail, D. Willner, S. J. Lasch, A. J. Henderson, S. Hofstead, A. M. Casazza, R. A. Firestone, I. Hellström and K. E. Hellström, *Science (80-.).*, 1993, **261**, 212–215.

56 Z. Fu, S. Li, S. Han, C. Shi and Y. Zhang, *Signal Transduct. Target. Ther.*, ,

DOI:10.1038/s41392-022-00947-7.

57 K. Tsuchikama and Z. An, *Protein Cell*, 2018, **9**, 33–46.

58 E. Khera, C. Cilliers, M. D. Smith, M. L. Ganno, K. C. Lai, T. A. Keating, A. Kopp, I. Nessler, A. O. Abu-Yousif and G. M. Thurber, *Neoplasia*, 2021, **23**, 210–221.

59 F. Giugliano, C. Corti, P. Tarantino, F. Michelini and G. Curigliano, *Curr. Oncol. Rep.*, 2022, **24**, 809–817.

60 Cheng KT, *Mol. Imaging Contrast Agent Database*, 2008, 1–7.

61 N. Diamantis and U. Banerji, 2016, 362–367.

62 T. Tedeschini, B. Campara, A. Grigoletto, M. Bellini, M. Salvalaio, Y. Matsuno, A. Suzuki, H. Yoshioka and G. Pasut, *J. Control. Release*, 2021, **337**, 431–447.

63 A. Pharm and O. Chemistry, 2007, **57**, 133–150.

64 W. Leimgruber, V. Stefanović, F. Schenker, A. Karr and J. Berger, *J. Am. Chem. Soc.*, 1965, **87**, 5791–5793.

65

66 G. S. Khan, A. Shah, Zia-ur-Rehman and D. Barker, *J. Photochem. Photobiol. B Biol.*, 2012, **115**, 105–118.

67 G. Wells, C. R. H. Martin, P. W. Howard, Z. A. Sands, C. A. Laughton, A. Tiberghien, C. K. Woo, L. A. Masterson, M. J. Stephenson, J. A. Hartley, T. C. Jenkins, S. D. Shnyder, P. M. Loadman, M. J. Waring and D. E. Thurston, *J. Med. Chem.*, 2006, **49**, 5442–5461.

68 K. M. Rahman, A. S. Thompson, C. H. James, M. Narayanaswamy and D. E. Thurston, *J. Am. Chem. Soc.*, 2009, **131**, 13756–13766.

69 S. J. Gregson, P. W. Howard, J. A. Hartley, N. A. Brooks, L. J. Adams, T. C. Jenkins, L. R. Kelland and D. E. Thurston, *J. Med. Chem.*, 2001, **44**, 737–748.

70 G. Lenglet and M. H. David-Cordonnier, *J. Nucleic Acids*, , DOI:10.4061/2010/290935.

71 A. C. Tiberghien, J. N. Levy, L. A. Masterson, N. V. Patel, L. R. Adams, S. Corbett, D. G. Williams, J. A. Hartley and P. W. Howard, *ACS Med. Chem. Lett.*, 2016, **7**, 983–987.

72 H. Chen, Z. Lin, K. E. Arnst, D. D. Miller and W. Li, *Molecules*, , DOI:10.3390/molecules22081281.

73 C. M. Mckertish and V. Kayser, *Biomedicines*, , DOI:10.3390/biomedicines9080872.

74 M. H. Park, B. I. Lee, J. J. Byeon, S. H. Shin, J. Choi, Y. Park and Y. G. Shin, *Molecules*, 2019, **24**, 1–14.

75 M. Lopus, *Cancer Lett.*, 2011, **307**, 113–118.

76 E. Oroudjev, M. Lopus, L. Wilson, C. Audette, C. Provenzano, H. Erickson, Y. Kovtun, R. Chari and M. A. Jordan, *Mol. Cancer Ther.*, 2010, **9**, 2700–2713.

77 S. Shao, M. H. Tsai, J. Lu, T. Yu, J. Jin, D. Xiao, H. Jiang, M. Han, M. Wang and J. Wang, *Bioorganic Med. Chem. Lett.*, 2018, **28**, 1363–1370.

78 Y. Pommier, 2006, **6**, 789–802.

79 T. Dasgupta and S. Ferdous, 2020, 12–14.

80 H. Okamoto, M. Oitate, K. Hagihara, H. Shiozawa, Y. Furuta, Y. Ogitani and H. Kuga, *Xenobiotica*, 2020, **50**, 1242–1250.

81 M. R. Gordon, M. Canakci, L. Li, J. Zhuang, B. Osborne and S. Thayumanavan, *Bioconjugate Chem.*, 2015, **26**, 2198–2215.

82 J. D. Bargh, A. Isidro-Llobet, J. S. Parker and D. R. Spring, *Chem. Soc. Rev.*, 2019, **48**, 4361–4374.

83 A. W. Tolcher, S. Sugarman, K. A. Gelmon, R. Cohen, M. Saleh, C. Isaacs, L. Young, D. Healey, N. Onetto and W. Slichenmyer, *J. Clin. Oncol.*, 1999, **17**, 478.

84 J. Zeng, O. S. Shirihai and M. W. Grinstaff, *J. life Sci. (Westlake Village, Calif.)*, 2020, **2**, 25–37.

85 J. Lu, F. Jiang, A. Lu and G. Zhang, *Int. J. Mol. Sci.*, 2016, **17**, 561.

86 S. O. Doronina, B. E. Toki, M. Y. Torgov, B. A. Mendelsohn, C. G. Cerveny, D. F. Chace, R. L. DeBlanc, R. P. Gearing, T. D. Bovee, C. B. Siegall, J. A. Francisco, A. F. Wahl, D. L. Meyer and P. D. Senter, *Nat. Biotechnol.*, 2003, **21**, 778–784.

87 R. Sheyi, B. G. de la Torre and F. Albericio, *Pharmaceutics*, , DOI:10.3390/pharmaceutics14020396.

88 Z. Su, D. Xiao, F. Xie, L. Liu, Y. Wang, S. Fan, X. Zhou and S. Li, *Acta Pharm. Sin. B*, 2021, **11**, 3889–3907.

89 H. F. Stanyon and J. H. Viles, *J. Biol. Chem.*, 2012, **287**, 28163–28168.

90 S. Sugio, A. Kashima, S. Mochizuki, M. Noda and K. Kobayashi, *Protein Eng.*, 1999, **12**, 439–446.

91 J. L. Beck, S. Ambahera, S. R. Yong, M. M. Sheil, J. De Jersey and S. F. Ralph, *Anal. Biochem.*, 2004, **325**, 326–336.

92 H. J. Forman, H. Zhang and A. Rinna, *Mol. Aspects Med.*, 2009, **30**, 1–12.

93 L. Kennedy, J. K. Sandhu, M. E. Harper and M. Cuperlovic-culf, *Biomolecules*, 2020, **10**, 1–27.

94 E. Lee and D. H. Lee, *BMB Rep.*, 2017, **50**, 401–410.

95 G. Casi and D. Neri, *Mol. Pharm.*, 2015, **12**, 1880–1884.

96 A. Beck, L. Goetsch, C. Dumontet and N. Corvaïa, *Nat. Rev. Drug Discov.*, 2017, **16**, 315–337.

97 B. A. Kellogg, L. Garrett, Y. Kovtun, K. C. Lai, B. Leece, M. Miller, G. Payne, R. Steeves, K. R. Whiteman, W. Widdison, H. Xie, R. Singh, R. V. J. Chari, J. M. Lambert and R. J. Lutz, *Bioconjugate Chem.*, 2011, **22**, 717–727.

98 J. T. W. Tong, P. W. R. Harris, M. A. Brimble and I. Kavianinia, *Molecules*, 2021, **26**, 5847.

99 G. M. Dubowchik and R. A. Firestone, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 3341–3346.

100 M. Dorywalska, R. Dushin, L. Moine, S. E. Farias, D. Zhou, T. Navaratnam, V. Lui, A. Hasa-Moreno, M. G. Casas, T. T. Tran, K. Delaria, S. H. Liu, D. Foletti, C. J. O'Donnell, J. Pons, D. L. Shelton, A. Rajpal and P. Strop, *Mol. Cancer Ther.*, 2016, **15**, 958–970.

101 J. D. Bargh, S. J. Walsh, A. Isidro-Llobet, S. Omarjee, J. S. Carroll and D. R. Spring, *Chem. Sci.*, 2020, **11**, 2375–2380.

102 J. R. McCombs and S. C. Owen, *AAPS J.*, 2015, **17**, 339–351.

103 N. Ashman, J. D. Bargh and D. R. Spring, *Chem. Soc. Rev.*, 2022, **13**, 9182–9202.

104 Y. V. Kovtun, C. A. Audette, Y. Ye, H. Xie, M. F. Ruberti, S. J. Phinney, B. A. Leece, T. Chittenden, W. A. Blättler and V. S. Goldmacher, *Cancer Res.*, 2006, **66**, 3214–3221.

105 S. Lonial, H. C. Lee, A. Badros, S. Trudel, A. K. Nooka, A. Chari, A. O. Abdallah, N. Callander, N. Lendvai, D. Sborov, A. Suvannasankha, K. Weisel, L. Karlin, E. Libby, B. Arnulf, T. Facon, C. Hulin, K. M. Kortüm, P. Rodríguez-Otero, S. Z. Usmani, P. Hari, R. Baz, H. Quach, P. Moreau, P. M. Voorhees, I. Gupta, A. Hoos, E. Zhi, J. Baron, T. Piontek, E. Lewis, R. C. Jewell, E. J. Dettman, R. Popat, S. D. Esposti, J. Opalinska, P. Richardson and A. D. Cohen, *Lancet Oncol.*, 2020, **21**, 207–221.

106 J. R. Junutula, H. Raab, S. Clark, S. Bhakta, D. D. Leipold, S. Weir, Y. Chen, M. Simpson, S. P. Tsai, M. S. Dennis, Y. Lu, Y. G. Meng, C. Ng, J. Yang, C. C. Lee, E. Duenas, J. Gorrell, V. Katta, A. Kim, K. McDorman, K. Flagella, R. Venook, S. Ross, S. D. Spencer, W. Lee Wong, H. B. Lowman, R. Vandlen, M. X. Sliwkowski, R. H. Scheller, P. Polakis and W. Mallet, *Nat. Biotechnol.*, 2008, **26**, 925–932.

107 I. E. Krop, M. Beeram, S. Modi, S. F. Jones, S. N. Holden, W. Yu, S. Girish, J. Tibbitts, J. H. Yi, M. X. Sliwkowski, F. Jacobson, S. G. Lutzker and H. A. Burris, *J. Clin. Oncol.*, 2010, **28**, 2698–2704.

108 K. J. Hamblett, P. D. Senter, D. F. Chace, M. M. C. Sun, J. Lenox, C. G. Cerveny, K. M. Kissler, S. X. Bernhardt, A. K. Kopcha, R. F. Zabinski, D. L. Meyer and J. A. Francisco, *Clin. Cancer Res.*, 2004, **10**, 7063–7070.

109 N. Forte, V. Chudasama and J. R. Baker, *Drug Discov. Today Technol.*, 2018, **30**, 11–20.

110 X. Sun, J. F. Ponte, N. C. Yoder, R. Laleau, J. Coccia, L. Lanieri, Q. Qiu, R. Wu, E. Hong, M. Bogalhas, L. Wang, L. Dong, Y. Setiady, E. K. Maloney, O. Ab, X. Zhang, J. Pinkas, T. A. Keating, R. Chari, H. K. Erickson and J. M. Lambert, *Bioconjugate Chem.*, 2017, **28**, 1371–1381.

111 S. M. Gheibi Hayat and A. H. Sahebkar, *Antibody drug conjugates for cancer therapy*, 2017, vol. 19.

112 L. Conilh, G. Fournet, E. Fourmaux, A. Murcia, E. L. Matera, B. Joseph, C. Dumontet and W. Viricel, *Pharmaceuticals*, 2021, **14**, 1–17.

113 J. R. Deng, S. F. Chung, A. S. L. Leung, W. M. Yip, B. Yang, M. C. Choi, J. F. Cui, K. K. Y. Kung, Z. Zhang, K. W. Lo, Y. C. Leung and M. K. Wong, *Commun. Chem.*, 2019, **2**, 1–10.

114 S. B. Gunnoo and A. Madder, *ChemBioChem*, 2016, **17**, 529–553.

115 S. Wang, J. Cao, W. Jia, W. Guo, S. Yan, Y. Wang, P. Zhang, H. Y. Chen and S. Huang, *Chem. Sci.*, 2020, **11**, 879–887.

116 R. S. Wible and T. R. Sutter, *Chem. Res. Toxicol.*, 2017, **30**, 729–762.

117 J. A. Flygare, T. H. Pillow and P. Aristoff, *Chem. Biol. Drug Des.*, 2013, **81**, 113–121.

118 C. Sornay, V. Vaur, A. Wagner and G. Chaubet, *An overview of chemo- and site-selectivity aspects in the chemical conjugation of proteins*, 2022, vol. 9.

119 B. H. Northrop, S. H. Frayne and U. Choudhary, *Polym. Chem.*, 2015, **6**, 3415–3430.

120 D. P. Nair, M. Podgórski, S. Chatani, T. Gong, W. Xi, C. R. Fenoli and C. N. Bowman, *Chem. Mater.*, 2014, **26**, 724–744.

121 K. Jendryczko, J. Rzeszotko, M. A. Krzyscik, A. Kocyła, J. Szymczyk, J. Otlewski and A. Szlachcic, *Mol. Pharm.*, 2022, **19**, 1422–1433.

122 G. T. Hermanson, ed. G. T. B. T.-B. T. (Third E. Hermanson, Academic Press, Boston, 2013, pp. 229–258.

123 S. C. Alley, D. R. Benjamin, S. C. Jeffrey, N. M. Okeley, D. L. Meyer, R. J. Sanderson and P. D. Senter, *Bioconjugate Chem.*, 2008, **19**, 759–765.

124 M. E. B. Smith, M. B. Caspersen, E. Robinson, M. Morais, A. Maruani, J. P. M. Nunes, K. Nicholls, M. J. Saxton, S. Caddick, J. R. Baker and V. Chudasama, *Org. Biomol. Chem.*, 2015, **13**, 7946–7949.

125 R. P. Lyon, J. R. Setter, T. D. Bovee, S. O. Doronina, J. H. Hunter, M. E. Anderson, C. L. Balasubramanian, S. M. Duniho, C. I. Leiske, F. Li and P. D. Senter, *Nat. Biotechnol.*, 2014, **32**, 1059–1062.

126 L. M. Tedaldi, M. E. B. Smith, R. I. Nathani and J. R. Baker, *Chem. Commun.*, 2009, 6583–6585.

127 M. E. B. Smith, F. F. Schumacher, C. P. Ryan, L. M. Tedaldi, D. Papaioannou, G. Waksman, S. Caddick and J. R. Baker, *J. Am. Chem. Soc.*, 2010, **132**, 1960–1965.

128 M. E. B. Smith, F. F. Schumacher, C. P. Ryan, L. M. Tedaldi, D. Papaioannou, G. Waksman, S. Caddick and J. R. Baker, *J. Am. Chem. Soc.*, 2010, **132**, 1960–1965.

129 C. P. Ryan, M. E. B. Smith, F. F. Schumacher, D. Grohmann, D. Papaioannou, G. Waksman, F. Werner, J. R. Baker and S. Caddick, *Chem. Commun.*, 2011, **47**, 5452–5454.

130 F. J. Chen and J. Gao, *Chem. - A Eur. J.*, , DOI:10.1002/chem.202201843.

131 B. Bernardim, P. M. S. D. Cal, M. J. Matos, B. L. Oliveira, N. Martínez-Saéz, I. S. Albuquerque, E. Perkins, F. Corzana, A. C. B. Burtoloso, G. Jiménez-Osés and G. J. L. Bernardes, *Nat. Commun.*, 2016, **7**, 1–9.

132 B. Bernardim, M. J. Matos, X. Ferhati, I. Compañón, A. Guerreiro, P. Akkapeddi, A. C. B. Burtoloso, G. Jiménez-Osés, F. Corzana and G. J. L. Bernardes, *Nat. Protoc.*, 2019, **14**, 86–99.

133 A. J. J. Lennox, *Angew. Chemie - Int. Ed.*, 2018, **57**, 14686–14688.

134 H. F. Motiwala, Y. H. Kuo, B. L. Stinger, B. A. Palfey and B. R. Martin, *J. Am. Chem. Soc.*, 2020, **142**, 1801–1810.

135 E. V. Vinogradova, C. Zhang, A. M. Spokoyny, B. L. Pentelute and S. L. Buchwald, *Nature*, 2015, **526**, 687–691.

136 W. Wang, O. Rusin, X. Xu, K. K. Kim, J. O. Escobedo, S. O. Fakayode, K. A. Fletcher, M. Lowry, C. M. Schowalter, C. M. Lawrence, F. R. Fronczek, I. M. Warner and R. M. Strongin, *J. Am. Chem. Soc.*, 2005, **127**, 15949–15958.

137 T. Müller and D. Winter, *Mol. Cell. Proteomics*, 2017, **16**, 1173–1187.

138 S. Balan, J. Choi, A. Godwin, I. Teo, C. M. Laborde, S. Heidelberger, M. Zloh, S. Shaunak and S. Brocchini, *Bioconjugate Chem.*, 2007, **18**, 61–76.

139 P. J. Hogg, *Trends Biochem. Sci.*, 2003, **28**, 210–214.

140 S. L. Kuan, T. Wang and T. Weil, *Chem. – A Eur. J.*, 2016, **22**, 17112–17129.

141 J. A. Burns, J. C. Butler, J. Moran and G. M. Whitesides, *J. Org. Chem.*, 1991, **56**, 2648–2650.

142 J. C. 3rd Lukesh, M. J. Palte and R. T. Raines, *J. Am. Chem. Soc.*, 2012, **134**, 4057–4059.

143 J. P. M. Nunes, M. Morais, V. Vassileva, E. Robinson, V. S. Rajkumar, M. E. B. Smith, R. B. Pedley, S. Caddick, J. R. Baker and V. Chudasama, *Chem. Commun.*, 2015, **51**, 10624–10627.

144 M. Morais, J. P. M. Nunes, K. Karu, N. Forte, I. Benni, M. E. B. Smith, S. Caddick, V. Chudasama and J. R. Baker, *Org. Biomol. Chem.*, 2017, **15**, 2947–2952.

145 A. Maruani, M. E. B. Smith, E. Miranda, K. A. Chester, V. Chudasama and S. Caddick, *Nat. Commun.*, 2015, **6**, 2–10.

146 E. Robinson, J. P. M. Nunes, V. Vassileva, A. Maruani, J. C. F. Nogueira, M. E. B. Smith, R. B. Pedley, S. Caddick, J. R. Baker and V. Chudasama, *RSC Adv.*, 2017, **7**, 9073–9077.

147 D. B. Volkin, H. Mach and C. R. Middaugh, *Appl. Biochem. Biotechnol. - Part B Mol. Biotechnol.*, 1997, **8**, 105–122.

148 S. J., S. Omarjee, W. R. J. D. Galloway, T. T. L. Kwan, H. F. Sore, J. S. Parker, M. Hyvönen, J. S. Carroll and D. R. Spring, *Chem. Sci.*, 2019, **10**, 694–700.

149 C. Bahou, D. A. Richards, A. Maruani, E. A. Love, F. Javaid, S. Caddick, J. R. Baker and V. Chudasama, *Org. Biomol. Chem.*, 2018, **16**, 1359–1366.

150 M. Nisavic, G. J. Wørmer, C. S. Nielsen, S. M. Jeppesen and T. B. Poulsen, 2022, 1–14.

151 I. André, S. Linse and F. A. A. Mulder, *J. Am. Chem. Soc.*, 2007, **129**, 15805–15813.

152 V. Gautier, A. J. Boumeester, P. Lössl and A. J. R. Heck, *Proteomics*, 2015, **15**, 2756–2765.

153 P. Dennler, E. Fischer and R. Schibli, *Antibodies*, 2015, **4**, 197–224.

154 S. J. Walsh, J. D. Bargh, F. M. Dannheim, A. R. Hanby, H. Seki, A. J. Counsell, X. Ou, E. Fowler, N. Ashman, Y. Takada, A. Isidro-Llobet, J. S. Parker, J. S. Carroll and D. R. Spring, *Chem. Soc. Rev.*, 2021, **50**, 1305–1353.

155 M. Haque, N. Forte and J. R. Baker, *Chem. Commun.*, 2021, **57**, 10689–10702.

156 T. Tamura and I. Hamachi, *J. Am. Chem. Soc.*, 2019, **141**, 2782–2799.

157 G. W. Anderson, F. M. Callahan and J. E. Zimmerman, *J. Am. Chem. Soc.*, 1967, **89**, 178.

158 O. Koniev and A. Wagner, *Chem. Soc. Rev.*, 2015, **44**, 5495–5551.

159 K. Yamada, N. Shikida, K. Shimbo, Y. Ito, Z. Khedri, Y. Matsuda and B. A. Mendelsohn, *Angew. Chemie - Int. Ed.*, 2019, **58**, 5592–5597.

160 J. Deisenhofer, *Biochemistry*, 1981, **20**, 2361–2370.

161 C. Yu, J. Tang, A. Loredo, Y. Chen, S. Y. Jung, A. Jain, A. Gordon and H. Xiao, *Bioconjugate Chem.*, 2018, **29**, 3522–3526.

162 S. R. Adusumalli, D. G. Rawale, U. Singh, P. Tripathi, R. Paul, N. Kalra, R. K. Mishra, S. Shukla and V. Rai, *J. Am. Chem. Soc.*, 2018, **140**, 15114–15123.

163 S. R. Adusumalli, D. G. Rawale, K. Thakur, L. Purushottam, N. C. Reddy, N. Kalra, S. Shukla and V. Rai, *Angew. Chemie Int. Ed.*, 2020, **59**, 10332–10336.

164 C. Rader, J. M. Turner, A. Heine, D. Shabat, S. C. Sinha, I. A. Wilson, R. A. Lerner and C. F. Barbas, *J. Mol. Biol.*, 2003, **332**, 889–899.

165 B. K. Hansen, C. J. Loveridge, S. Thyssen, G. J. Wørmer, A. D. Nielsen, J. Palmfeldt, M. Johannsen and T. B. Poulsen, *Angew. Chemie - Int. Ed.*, 2019, **58**, 3533–3537.

166 M. S. Kang, T. W. S. Kong, J. Y. X. Khoo and T. P. Loh, *Chem. Sci.*, 2021, **12**, 13613–13647.

167 J. Gavrilyuk, H. Ban, M. Nagano, W. Hakamata and C. F. Barbas, *Bioconjugate Chem.*, 2012, **23**, 2321–2328.

168 D. Gilis, S. Massar, N. J. Cerf and M. Rooman, *Genome Biol.*, 2001, **2**, 1–12.

169 Y. Seki, T. Ishiyama, D. Sasaki, J. Abe, Y. Sohma, K. Oisaki and M. Kanai, *J. Am. Chem. Soc.*, 2016, **138**, 10798–10801.

170 K. J. Malawska, S. Takano, K. Oisaki, H. Yanagisawa, M. Kikkawa, T. Tsukuda and M. Kanai, *ChemRxiv. Cambridge Cambridge Open Engag. 2022; This content is a Prepr. has not been peer-reviewed.*

171 B. Q. Shen, K. Xu, L. Liu, H. Raab, S. Bhakta, M. Kenrick, K. L. Parsons-Reponte, J. Tien, S. F. Yu, E. Mai, D. Li, J. Tibbitts, J. Baudys, O. M. Saad, S. J. Scales, P. J. McDonald, P. E. Hass, C. Eigenbrot, T. Nguyen, W. A. Solis, R. N. Fuji, K. M. Flagella, D. Patel, S. D. Spencer, L. A. Khawli, A. Ebens, W. L. Wong, R. Vandlen, S. Kaur, M. X. Sliwkowski, R. H. Scheller, P. Polakis and J. R. Junutula, *Nat. Biotechnol.*, 2012, **30**, 184–189.

172 H. L. Perez, P. M. Cardarelli, S. Deshpande, S. Gangwar, G. M. Schroeder, G. D. Vite and R. M. Borzilleri, *Drug Discov. Today*, 2014, **19**, 869–881.

173 S. A. Kularatne, V. Deshmukh, J. Ma, V. Tardif, R. K. V Lim, H. M. Pugh, Y. Sun, A. Manibusan, A. J. Sellers, R. S. Barnett, S. Srinagesh, J. S. Forsyth, W. Hassenpflug, F. Tian, T. Javahishvili, B. Felding-Habermann, B. R. Lawson, S. A. Kazane and P. G. Schultz, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 11863–11867.

174 H. M. Burke, L. McSweeney and E. M. Scanlan, *Nat. Commun.*, 2017, **8**, 1–16.

175 V. Y. Torbeev and S. B. H. Kent, *Angew. Chemie - Int. Ed.*, 2007, **46**, 1667–1670.

176 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776–779.

177 S. B. H. Kent, *Chem. Soc. Rev.*, 2009, **38**, 338–351.

178 J. C. M. Monbaliu, G. Dive, C. V. Stevens and A. R. Katritzky, *J. Chem. Theory Comput.*, 2013, **9**, 927–934.

179 S. S. Panda, C. D. Hall, A. A. Oliferenko and A. R. Katritzky, *Acc. Chem. Res.*, 2014, **47**, 1076–1087.

180 S. S. Kulkarni, J. Sayers, B. Premdjee and R. J. Payne, *Nat. Rev. Chem.*, 2020, **2**, 1–17.

181 B. Dang, T. Kubota, K. Mandal, F. Bezanilla and S. B. H. Kent, *J. Am. Chem. Soc.*, 2013, **135**, 11911–11919.

182 E. C. B. Johnson and S. B. H. Kent, *J. Am. Chem. Soc.*, 2006, **128**, 6640–6646.

183 C. P. R. Hackenberger and D. Schwarzer, *Angew. Chemie - Int. Ed.*, 2008, **47**, 10030–10074.

184 M. J. Matos, B. L. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. M. S. D. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. J. Deery, J. M. Chalker, F. Corzana, G. Jiménez-Osés and G. J. L. Bernardes, *J. Am. Chem. Soc.*, 2018, **140**, 4004–4017.

185 L. J. Chan, D. B. Ascher, R. Yadav, J. B. Bulitta, C. C. Williams, C. J. H. Porter, C. B. Landersdorfer and L. M. Kaminskas, *Mol. Pharm.*, 2016, **13**, 1229–1241.

186 N. Forte, I. Benni, K. Karu, V. Chudasama and J. R. Baker, *Chem. Sci.*, 2019, **10**, 10919–10924.

187 W. Yang and D. G. Drueckhammer, *J. Am. Chem. Soc.*, 2001, **123**, 11004–11009.

188 I. Yanagisawa, Y. Hirata and Y. Ishii, *J. Med. Chem.*, 1984, **27**, 849–857.

189 F. F. Fleming, L. Yao, P. C. Ravikumar, L. Funk and B. C. Shook, *J. Med. Chem.*, 2010, **53**, 7902–7917.

190 X. Wang, Y. Wang, X. Li, Z. Yu, C. Song and Y. Du, *RSC Med. Chem.*, 2021, **12**, 1650–1671.

191 Y. Wang, Y. Du and N. Huang, *Future Med. Chem.*, 2018, **10**, 2713–2728.

192 V. L. Schramm, R. H. Furneaux, P. C. Tyler and G. B. Evans, 2010, 9136–9143.

193 V. Bonatto, R. F. Lameiro, F. R. Rocho, J. Lameira, A. Leitão and C. A. Montanari, *RSC Med. Chem.*, 2022, **14**, 201–217.

194 D. M. Blow, *Acc. Chem. Res.*, 1976, **9**, 145–152.

195 Ö. D. Ekici, M. Paetzel and R. E. Dalbey, *Protein Sci.*, 2008, **17**, 2023–2037.

196 M. W. Hunkapiller, S. H. Smallcombe, D. R. Whitaker and J. H. Richards, *J. Biol. Chem.*, 1973, **248**, 8306–8308.

197 R. Wieczorek, K. Adamala, T. Gasperi, F. Polticelli and P. Stano, *Life*, 2017, **7**, 1–24.

198 E. Altmann, R. Aichholz, C. Betschart, T. Buhl, J. Green, O. Irie, N. Teno, R. Lattmann, M. Tintelnot-Blomley and M. Missbach, *J. Med. Chem.*, 2007, **50**, 591–594.

199 C. P. Ramil, P. An, Z. Yu and Q. Lin, *J. Am. Chem. Soc.*, 2016, **138**, 5499–5502.

200 E. H. White, F. McCapra and G. F. Field, *J. Am. Chem. Soc.*, 1963, **85**, 337–343.

201 H. Ren, F. Xiao, K. Zhan, Y. P. Kim, H. Xie, Z. Xia and J. Rao, *Angew. Chemie - Int. Ed.*, 2009, **48**, 9658–9662.

202 S. G. L. Keyser, A. Utz and C. R. Bertozzi, *J. Org. Chem.*, 2018, **83**, 7467–7479.

203 C. S. Foden, S. Islam, C. Fernández-García, L. Maugeri, T. D. Sheppard and M. W. Powner, *Science (80-.).*, 2020, **370**, 865–869.

204 R. K. Chaturvedi, A. E. MacMahon and G. L. Schmir, *J. Am. Chem. Soc.*, 1967, **89**, 6984–6993.

205 V. A. Savëlova, A. P. Prudchenko and L. P. Drizhd, *Theor. Exp. Chem.*, 1993, **28**, 307–312.

206 E. H. Cordes and W. P. Jencks, *J. Am. Chem. Soc.*, 1962, **84**, 832–837.

207 R. C. Schnur, *J. Org. Chem.*, 1979, **44**, 3726–3727.

208 Y. Qing, M. D. Liu, D. Hartmann, L. Zhou, W. J. Ramsay and H. Bayley, *Angew. Chemie - Int. Ed.*, 2020, **59**, 15711–15716.

209 G. S. Pulcu, E. Mikhailova, L. S. Choi and H. Bayley, *Nat. Nanotechnol.*, 2015, **10**, 76–83.

210 A. Chrzastek, I. A. Thanasi, J. A. Irving, V. Chudasama and J. R. Baker, *Chem. Sci.*, 2022, **13**, 11533–11539.

211 C. Hansch, A. Leo and R. W. Taft, *Chem. Rev.*, 1991, **91**, 165–195.

212 P. R. Campodónico, B. Olivares and R. A. Tapia, *Front. Chem.*, 2020, **8**, 1–9.

213 Y. Karuo, K. Yamada and M. Kunishima, *Chem. Pharm. Bull.*, 2018, **66**, 303–308.

214 B. L. Oliveira, Z. Guo and G. J. L. Bernardes, *Chem. Soc. Rev.*, 2017, **46**, 4895–4950.

215 J. Jouha, F. Buttard, M. Lorion, C. Berthonneau, M. Khouili, M.-A. Hiebel, G. Guillaumet, J.-F. Brière and F. Suzenet, *Org. Lett.*, 2017, **19**, 4770–4773.

216 H. Shimojo, K. Moriyama and H. Togo, *Synth.*, 2013, **45**, 2155–2164.

217 H.-L. Vo, J. L. Arthur, M. Capdevila-Cortada, S. H. Lapidus, P. W. Stephens, J. J. Novoa, A. M. Arif, R. K. Nagi, M. H. Bartl and J. S. Miller, *J. Org. Chem.*, 2014, **79**, 8189–8201.

218 C. Zambaldo, E. V. Vinogradova, X. Qi, J. Iaconelli, R. M. Suciu, M. Koh, K. Senkane, S. R. Chadwick, B. B. Sanchez, J. S. Chen, A. K. Chatterjee, P. Liu, P. G. Schultz, B. F. Cravatt and M. J. Bollong, *J. Am. Chem. Soc.*, 2020, **142**, 8972–8979.

219 M. M. Pichon, D. Drelinkiewicz, D. Lozano, R. Moraru, L. J. Hayward, M. Jones, M. A. McCoy, S. Allstrum-Graves, D.-I. Balourdas, A. C. Joerger, R. J. Whitby, S. M. Goldup, N. Wells, G. J. Langley, J. M. Herniman and M. G. J. Baud, *Bioconjugate Chem.*, , DOI:10.1021/acs.bioconjchem.3c00322.

220 S. J. Siegl, R. Dzijak, A. Vázquez, R. Pohl and M. Vrabel, *Chem. Sci.*, 2017, **8**, 3593–3598.

221 N. R. Dalabehera, S. Meher, B. Bhusana Palai and N. K. Sharma, *ACS Omega*, 2020, **5**, 26141–26152.

222 D. Cheshmedzhieva, S. Ilieva, B. Hadjieva, T. Trayanova and B. Galabov, *Mol. Phys.*, 2009, **107**, 1187–1192.

223 K. KAKEMI, H. SEZAKI, M. NAKANO, K. OHSUGA and T. MITSUNAGA, *Chem. Pharm. Bull. (Tokyo).*, 1969, **17**, 901–905.

224 N. J. Agard, J. A. Prescher and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2004, **126**, 15046–15047.

225 G. A. Bagiyan, I. K. Koroleva, N. V. Soroka and A. V. Ufimtsev, *Russ. Chem. Bull.*, 2003, **52**, 1135–1141.

226 J. R. Winther and C. Thorpe, *Biochim. Biophys. Acta - Gen. Subj.*, 2014, **1840**, 838–846.

227 A. M. Merlot, D. S. Kalinowski and D. R. Richardson, *Front. Physiol.*, 2014, **5 AUG**, 1–7.

228 H. A. Khan and B. Mutus, *Front. Chem.*, 2014, **2**, 1–9.

229 S. P. Brown and A. B. I. I. I. Smith, *J. Am. Chem. Soc.*, 2015, **137**, 4034–4037.

230 Y. Liu, Y. Yu, Q. Zhao, C. Tang, H. Zhang, Y. Qin, X. Feng and J. Zhang, *Coord. Chem. Rev.*, 2021, **427**, 213601.

231 S. D. Schnell, L. V. Hoff, A. Panchagnula, M. H. H. Wurzenberger, T. M. Klapötke, S. Sieber, A. Linden and K. Gademann, *Chem. Sci.*, 2020, **11**, 3042–3047.

232 C. Wiedemann, A. Kumar, A. Lang and O. Ohlenschläger, *Front. Chem.*, 2020, **8**, 1–8.

233 S. V. G. Kudrevich Maria G; van Lier, Johan E, *Synthesis (Stuttg).*, 1994, **1994**, 779–781.

234 E. H. Abdelkader, H. Qianzhu, J. George, R. L. Frkic, C. J. Jackson, C. Nitsche, G. Otting and T. Huber, *Angew. Chemie Int. Ed.*, 2022, **61**, e202114154.

235 Z. Zheng, P. Chen, G. Li, Y. Zhu, Z. Shi, Y. Luo, C. Zhao, Z. Fu, X. Cui, C. Ji, F. Wang, G. Huang and G. Liang, *Chem. Sci.*, 2016, **8**, 214–222.

236 Y. I. Belozerova, S. V. Efimova, A. B. Korzhenevskii and O. I. Koifman, *Russ. J. Gen. Chem.*, 2009, **79**, 2678–2684.

237 O. Ovdiichuk, O. Hordiyenko, E. Fotou, C. Gaucher, A. Arrault and M. C. Averlant-Petit, *Struct. Chem.*, 2017, **28**, 813–822.

238 J. E. Hein and V. V. Fokin, *Chem. Soc. Rev.*, 2010, **39**, 1302–1315.

239 N. Naruse, D. Kobayashi, K. Ohkawachi, A. Shigenaga and A. Otaka, *J. Org. Chem.*, 2020, **85**, 1425–1433.

240 S. Martin, E. M. Lenz, R. Smith, D. G. Temesi, A. L. Orton and M. R. Clench, *Rapid Commun. Mass Spectrom.*, 2017, **31**, 145–151.

241 T. A. Berger, B. K. Berger and K. B. T.-R. M. in C. Kogelman Molecular Sciences and Chemical Engineering, Elsevier, 2022.

242 G. Vistoli and A. Pedretti, eds. J. B. Taylor and D. J. B. T.-C. M. C. I. I. Triggle, Elsevier, Oxford, 2007, pp. 577–602.

243 S. Tariq, K. Somakala and M. Amir, *Eur. J. Med. Chem.*, 2018, **143**, 542–557.

244 M. Morais, J. P. M. Nunes, K. Karu, N. Forte, I. Benni, M. E. B. Smith, S. Caddick, V. Chudasama and J. R. Baker, *Org. Biomol. Chem.*, 2017, **15**, 2947–2952.

245 E. P. Aydin and E. T. Uyar, *Thioredoxin as an antioxidant protein as a marker in depression*, Elsevier Inc., 2021.

246 J.-F. Collet and J. Messens, *Antioxid. Redox Signal.*, 2010, **13**, 1205–1216.

247 J. Lundström and A. Holmgren, *J. Biol. Chem.*, 1990, **265**, 9114–9120.

248 X.-P. Zhang, Q. Yuan, Y.-L. Qi, D.-J. Zheng, Q.-X. Liu, B.-Z. Wang, Y.-S. Yang and H.-L. Zhu, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, 2019, **220**, 117134.

249 W. R. Sherman and E. Robins, *Anal. Chem.*, 1968, **40**, 803–805.

250 J. Tu, D. Svatunek, S. Parvez, A. C. Liu, B. J. Levandowski, H. J. Eckvahl, R. T. Peterson, K. N. Houk and R. M. Franzini, *Angew. Chemie Int. Ed.*, 2019, **58**, 9043–9048.

251 E. H. Mørkved and C. Wang, *J. für Prakt. Chemie/Chemiker-Zeitung*, 1997, **339**, 473–476.

252 M. Karlikova, V. Cermakova, J. Demuth, V. Valer, M. Miletin, V. Novakova and P. Zimcik, *Dalt. Trans.*, 2019, **48**, 6162–6173.

253 P. Zimcik, M. Miletin, V. Novakova, K. Kopecky, M. Nejedla, V. Stara and K. Sedlackova, *Aust. J. Chem.*, 2009, **62**, 425–433.

254 V. Novakova, P. Zimcik, M. Miletin, K. Kopecky and J. Ivincová, *Tetrahedron Lett.*, 2010, **51**, 1016–1018.

255 M. T. W. Lee, A. Maruani, J. R. Baker, S. Caddick and V. Chudasama, *Chem. Sci.*, 2016, **7**, 799–802.

256 F. Bryden, A. Maruani, H. Savoie, V. Chudasama, M. E. B. Smith, S. Caddick and R. W. Boyle, *Bioconjugate Chem.*, 2014, **25**, 611–617.

257 A. Miranda-Vizuete, A. E. Damdimopoulos, J.-Å. Gustafsson and G. Spyrou, *J. Biol. Chem.*, 1997, **272**, 30841–30847.

258 N. Martínez-Sáez, S. Sun, D. Oldrini, P. Sormanni, O. Boutureira, F. Carboni, I. Compañón, M. J. Deery, M. Vendruscolo, F. Corzana, R. Adamo and G. J. L. Bernardes, *Angew. Chemie Int. Ed.*, 2017, **56**, 14963–14967.

259 C. Bahou, P. A. Szijj, R. J. Spears, A. Wall, F. Javaid, A. Sattikar, E. A. Love, J. R. Baker and V. Chudasama, *Bioconjugate Chem.*, 2021, **32**, 672–679.