

Screening for selective ADAMTS-5 substrates to monitor proteinase activity



Milan Mrázek Fowkes, MChem MSc

Wolfson College

Trinity 2021

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

Doctor of Philosophy, DPhil

Abstract

Catabolic proteinase activity plays an important role in the aetiology of diseases such as osteoarthritis (OA), an affliction characterised by the progressive degradation of articular cartilage. Detection of this activity in relevant disease assays would be a useful tool for assessing disease diagnosis, progression and evaluation of targeted therapies. A key proteinase in OA is a disintegrin-like and metalloproteinase with thrombospondin type-I motifs-5 (ADAMTS-5), which cleaves aggrecan, one of the major structural constituents of articular cartilage. Fluorescence resonance energy transfer (FRET) peptide substrates represent a promising means to detect and monitor disease-related proteinase activity, owing to their high detection sensitivity, ease-of-use, and potential to target proteinases selectively. The aim of this thesis was to develop a selective ADAMTS-5 FRET peptide substrate for monitoring early proteinase activity in diseases such as osteoarthritis. Towards this goal, peptide substrates were initially modelled into the active site of ADAMTS-5 and synthesised by 9-fluorenylmethoxycarbonyl-based solid-phase peptide synthesis (Fmoc-SPPS). After screening against ADAMTS-5 on-resin, the lead substrate KY(NO₂)TESESRGK(Abz)IYYKKG was used to generate a combinatorial library of ~20,000 unique peptide sequences, which was then screened against ADAMTS-5 to discover hit peptide substrates. The library parent compound KY(NO₂)TESESRGK(Abz)IYYKKG and hit peptide substrates were evaluated for cleavage by ADAMTS-5 in solution along with their selectivity for this proteinase compared to ADAMTS-4, MMP-2 and MMP-9. This led to the identification of KY(NO₂)TESESRGK(Abz)IYYKKG and KY(NO₂)SENE SRGK(Abz)IYYKKG as the best substrates. When compared to the current ADAMTS-5 literature substrate, they showed ~4-fold and ~3-fold higher overall cleavage rates by ADAMTS-5 and ~2-fold and ~1.5-fold higher catalytic efficiencies, respectively.

In addition, they both showed ~16-fold/~13-fold selectivity for ADAMTS-5 over ADAMTS-4, ~10-fold/~8-fold selectivity over MMP-2 and ~2561-fold/~548-fold selectivity over MMP-9. The potential of both substrates to sensitively measure ADAMTS-5 activity was shown by their ability to detect low nanomolar concentrations of this proteinase. These novel FRET peptide substrates thus have potential application in measuring ADAMTS-5 activity in a variety of *in vitro*, *ex vivo* and potentially *in vivo* settings to support OA research; and may also be applied to other diseases characterised by elevated ADAMTS-5 activity.

Co-authorship statement

The work contained herein was performed solely by the author except for the modelling of the Y(NO₂)TESESRGK(Abz)IYYKKG peptide substrate in Chapter 4 and some of the MALDI and tandem mass spectrometry data shown in Chapter 5. This was carried out by Professor Morten Meldal at the University of Copenhagen, Denmark.

The work described in Chapter 3 and Chapter 5 was performed by the author at the Kennedy Institute of Rheumatology, University of Oxford, UK.

The work described in Chapter 4 was performed by the author at the Center for Evolutionary Chemical Biology (CECB), University of Copenhagen, Denmark.

“Good, better, best. Never let it rest. 'Til your good is better and your better is best.”

-St. Jerome

Acknowledgements

It's difficult to write an acknowledgements section, as this DPhil would not have been possible without the support of so many people. I shall, however, endeavour to do my best to make sure that no-one is omitted from this list. If I have forgotten anyone, I do apologise.

First of all, I would like to thank my supervisor Han for all his help and support during the course of my studies, especially for taking his time to go through my presentations and work, along with giving me the opportunity to be his first DPhil student!

I would especially like to thank Linda, for taking time out of her busy schedule to teach me enzyme kinetics, active-site titrations, and K_i determinations. She patiently mentored me while discussing my data and project to the betterment of my scientific knowledge and training. As a result of this, I have unexpectedly become a big fan of enzyme kinetics!

I would also like to thank my co-supervisor Tonia for her insightful suggestions and support throughout my DPhil, especially towards the end. I am particularly grateful to her for giving me a thorough grounding on presenting to a lay audience. I would not be a well-rounded speaker and scientist without her knowledge and guidance.

A big thanks also goes to my co-supervisor Paul, who was the first to tell me that “getting a PhD is not a race,” and that not doing one directly after your undergraduate degree is not an issue—it is perhaps the better option! Most importantly of all, he helped me be more confident and believe in myself. Along with Paul, I'd also like to extend my thanks to the entire Med Chem Lab at the Target Discovery Institute over the years for their suggestions and encouragement while I was doing my DPhil. Fillippo, Tamas, Moses, Raina, and many, many, others.

I would also like to extend my thanks to the Kennedy Trustees for providing me with a timely funded extension to complete my DPhil after the coronavirus crisis struck. This was very generous and meant a lot to me.

I also appreciate Wolfson College for their support in general and especially during the lockdowns of the coronavirus crisis. We were never in the dark about the situation our college was in, and they were always willing to help.

At the University of Copenhagen, I would like to thank Martin for discussions on Combi-chem and click chemistry, Andrej for keeping me good company during our lunchtime discussions on politics and history, and Frederik for rekindling my interest in organic chemistry.

A huge thank you also goes to Professor Morten Meldal at the University of Copenhagen. His incredible patience in teaching me about Combi-chem, making sure that I understood everything and providing help whenever I needed it was remarkable. His knowledge, commitment and dedication to Chemistry and Science in general was inspiring and highly motivating.

I should also not forget about my family, especially my mum, who convinced me to do a DPhil back in 2011, even when I wasn't sure whether I should or not. I ended up starting it 5 years later, but I got there in the end! The support that came from my fiancée's family, especially from my Eomma and aunt Bethel, was equally invaluable.

Last but not least, I would like to thank my biggest fan, best friend, and life-long partner Sunwoo for everything. Her continual support, care, and love was what kept me motivated. It especially mattered in the final year of my DPhil, during a global pandemic. She is truly incredible and very loving, and I am very fortunate to have her in my life.

Table of Contents

Abstract.....	i
Co-authorship statement	iii
Acknowledgements	v
Table of Contents	vii
List of Figures.....	xiv
List of Tables.....	xix
List of Schemes	xx
List of abbreviations and symbols	xxi
Chapter 1. Introduction	1
1.1. Preamble: proteinases and diseases.....	2
1.2. Proteinases and metalloproteinases	2
1.2.1. Classification of proteinases	2
1.2.2. Metalloproteinases.....	4
1.2.2.1. MMPs: structure and function	4
1.2.2.2. ADAMTSs: structure and function	5
1.2.2.3. The activity of ADAMTSs and MMPs is regulated by tissue inhibitors of metalloproteinases.....	8
1.3. Detecting proteinase activity	10
1.3.1. Methods used in chemistry and biochemistry	10
1.3.2. Measuring proteinase activity through molecular imaging.....	12
1.3.2.1. Monitoring proteinase activity with FRET peptide substrates	15
1.4. Osteoarthritis.....	18
1.4.1. Introduction.....	18
1.4.2. The structure and function of articular cartilage.....	20
1.4.2.1. The zonal anatomy of articular cartilage.....	22

1.4.3.	Metalloproteinase activity in osteoarthritis	23
1.4.3.1.	Catabolic MMPs: <i>in vitro</i> studies with bovine cartilage explants	24
1.4.3.2.	Catabolic MMPs: early <i>in vitro</i> studies with human osteoarthritic cartilage and synovial fluid	25
1.4.3.3.	Catabolic MMPs: later <i>in vitro</i> studies with human cartilage explants revealed MMP-13 as a major collagenase.....	27
1.4.3.4.	The MMP-13 knockout mouse model of osteoarthritis partially protects cartilage from erosion	28
1.4.3.5.	Catabolic ADAMTSs: aggrecanases and the identification of cleavage sites in aggrecan	29
1.4.3.6.	ADAMTSs: <i>in vitro</i> studies with human osteoarthritic cartilage and synovial fluid.....	31
1.4.3.7.	ADAMTS-4 and ADAMTS-5: which is the major aggrecanase in osteoarthritis?	32
1.4.3.8.	Targeting metalloproteinases as a means to image catabolic proteinase activity in osteoarthritis	35
1.4.3.9.	Commercial MMP FRET substrates lack selectivity	38
1.4.3.10.	MMP-13 FRET substrates can diagnose osteoarthritis in rodents but cannot detect and monitor early osteoarthritic processes.....	39
1.5.	Combinatorial chemistry.....	41
1.5.1.	Introduction.....	41
1.5.2.	Screening substrate libraries requires specialised resin beads	44
1.5.3.	Identification of peptide hits through optical encoding.....	46
1.5.4.	Identification of peptide hits through MALDI and tandem mass spectrometry	48
1.6.	Hypothesis	50
1.7.	Aims	50
Chapter 2.	Materials and methods	51
2.1.	Molecular biology and biochemistry	52
2.1.1.	Materials.....	52

2.1.1.1.	Tissue culture	52
2.1.1.2.	Protein purification	52
2.1.1.3.	SDS-PAGE.....	53
2.1.1.4.	Coomassie staining.....	53
2.1.1.5.	Silver staining	53
2.1.1.6.	Western blotting	53
2.1.1.7.	Gelatin zymography	54
2.1.1.8.	Enzyme kinetics	54
2.1.2.	Methods	54
2.1.2.1.	Tissue culture	54
2.1.2.2.	Purification of TS4-4 and TS5-5 by FLAG affinity chromatography ...	56
2.1.2.3.	Purification of proMMP-2ΔC by gelatin-sepharose affinity chromatography	57
2.1.2.4.	Removal of DMSO from collected fractions by native sephacryl affinity chromatography	58
2.1.2.5.	SDS-PAGE.....	59
2.1.2.6.	Coomassie staining.....	59
2.1.2.7.	Silver staining	60
2.1.2.8.	Western blotting	60
2.1.2.9.	Gelatin zymography	61
2.1.2.10.	Concentration and long-term storage of metalloproteinases	61
2.1.2.11.	Fluorophore calibration curves.....	62
2.1.2.12.	General information on kinetic studies with ADAMTSs and MMPs .	63
2.1.2.13.	Linear time frame of substrate hydrolysis by TS4-4 and TS5-5.....	63
2.1.2.14.	Activity studies of ADAMTSs and MMPs.....	64
2.1.2.15.	Surfactant activity studies of TS4-4 and TS5-5.....	65
2.1.2.16.	Active site titrations of ADAMTSs and MMPs with N-TIMP-3 and CT-1746	66

2.1.2.17.	Determination of CT-1746 inhibition constants for TS4-4/TS5-5	68
2.1.2.18.	Selectivity studies with ADAMTSs and MMPs	69
2.1.2.19.	Determination of kinetic parameters for substrate cleavage by TS5-5: k_{cat} and K_m	70
2.1.2.20.	Correcting kinetic parameters for the inner-filter effect	71
2.1.2.21.	Evaluating the rate of cleavage of library parent compound 5 and hit peptide 39 at different substrate concentrations.....	72
2.1.2.22.	Determination of the lowest concentration of TS5-5 that can be detected	72
2.2.	Computational Chemistry	73
2.2.1.	Modelling of peptide substrates into the active site of ADAMTS-5	73
2.3.	Combinatorial chemistry and chemical synthesis.....	74
2.3.1.	General information	74
2.3.2.	Synthesis of Fmoc-Lys(<i>retro</i> -Abz-Boc)-OH (22)	75
2.3.3.	Linear synthesis of peptide substrates on PEGA ₁₉₀₀ resin	76
2.3.4.	Kaiser Test	77
2.3.5.	Cleavage of peptide substrates from PEGA ₁₉₀₀ resin.....	78
2.3.6.	MALDI-TOF-MS of peptide substrates	78
2.3.7.	On-resin screening of modelled peptide substrate against TS5-5.....	78
2.3.8.	Synthesis of a combinatorial library on PEGA ₁₉₀₀ resin	79
2.3.9.	Incubation of the combinatorial library with TS5-5 and isolation of hits .	80
2.3.10.	Cleavage of hit peptide substrates from PEGA ₁₉₀₀ resin	80
2.3.11.	MALDI and tandem mass spectrometry of hit peptide substrates	81
2.3.12.	Analysis of mass spectra with computational software.....	81
Chapter 3.	The purification and characterisation of ADAMTS-4 and ADAMTS-5 proteinases	83
3.1.	Introduction	84
3.2.	Results.....	87

3.2.1.	Expression and selection of ADAMTS-4 and ADAMTS-5 for purification..	87
3.2.2.	Purification of ADAMTS-4 and ADAMTS-5 by FLAG affinity chromatography	88
3.2.3.	Validation of ADAMTS-4 and ADAMTS-5 activities using FRET substrates ...	91
3.2.4.	Determination of ADAMTS-4 and ADAMTS-5 concentrations by active-site titration with N-TIMP-3	92
3.2.5.	Determination of CT-1746 inhibition constants for ADAMTS-4 and ADAMTS-5	94
3.2.6.	Validation of ADAMTS-4 and ADAMTS-5 concentrations by active-site titration with CT-1746	98
3.2.7.	Summary of purified amounts of ADAMTS-4 and ADAMTS-5 proteinases ..	98
3.3.	Discussion	99
Chapter 4. The design, synthesis and screening of an ADAMTS-5 combinatorial library		107
4.1.	Introduction	108
4.2.	Results.....	110
4.2.1	Determination of the optimal surfactant concentration for FRET substrate cleavage by ADAMTS-4 and ADAMTS-5 proteinases.....	110
4.2.2	Determination of the selectivity of FRET substrates for ADAMTS-4 and ADAMTS-5 over a panel of metalloproteinases.....	112
4.2.3	General design and docking of ADAMTS-5 FRET peptide substrates	114
4.2.4	Design and computational docking of a P4' fluorophore series of ADAMTS-5 FRET peptide substrates	117
4.2.5	Design and computational docking of a P7' fluorophore series of ADAMTS-5 FRET peptide substrates	122
4.2.6	Design and computational docking of a P9' series of ADAMTS-5 FRET peptide substrates.....	126
4.2.7	Fmoc-based solid-phase peptide synthesis of modelled ADAMTS-5 FRET peptide substrates.....	130
4.2.8	Evaluation of modelled FRET peptide substrates against the ADAMTS-5 proteinase.....	133

4.3. Discussion	142
Chapter 5. The identification and characterisation of peptide hits from ADAMTS-5 combinatorial library screening	149
5.1. Introduction	150
5.2. Results.....	152
5.2.1. Identification of peptide hits from MALDI and tandem mass spectra using computational software	152
5.2.2. Synthesis of peptide hits by Fmoc-SPPS	156
5.2.3. Evaluation of purified peptide hits for cleavage by ADAMTS-5.....	157
5.2.4. Determination of kinetic parameters for cleavage of peptide hits by ADAMTS-5	159
5.2.5. Expression and selection of proMMP-2 Δ C for purification	164
5.2.6. Purification of proMMP-2 Δ C by gelatin-sepharose and native sephacryl affinity chromatography.....	164
5.2.7. Validation of MMP-2 Δ C activity and determination of concentration through active-site titration with N-TIMP-3	167
5.2.8. Determination of selectivity of peptide hits for ADAMTS-4, MMP-2 and MMP-9.....	168
5.2.9. Determination of peptide 5 and peptide 39 sensitivity for ADAMTS-5..	170
5.3. Discussion	174
Chapter 6. Conclusions and future work.....	181
6.1. Summary of this work.....	182
6.2. The importance of selectivity in the design of FRET substrates	183
6.2.1. The difference between selectivity and specificity.....	183
6.2.2. The importance of selectivity for <i>in vivo</i> FRET peptide substrates.....	184
6.3. The importance of kinetic parameters in the design of FRET substrates	185
6.3.1. Modulating the Michaelis constant and the catalytic rate constant <i>in vitro</i>	185
6.3.2. Considering pharmacokinetic parameters for <i>in vivo</i> FRET peptide substrates	186

6.3.3.	Considering the target protease concentration <i>in vivo</i>	186
6.3.4.	Summary.....	188
6.4.	Potential applications for selective ADAMTS-5 FRET substrates	188
6.4.1.	General <i>in vitro</i> applications.....	188
6.4.2.	Application to tissue culture models of osteoarthritis	189
6.4.3.	Characterisation of pharmacological inhibitors.....	190
6.4.4.	High-throughput screening for drugs	190
6.4.5.	ADAMTS-5 activity as a biomarker of human osteoarthritis	191
6.4.6.	Monitoring ADAMTS-5 activity in animal models of osteoarthritis.....	191
6.4.7.	Protease-activated prodrugs	192
6.5.	Future outlook	193
6.5.1.	Developing selective substrates for other ADAMTS proteinases	193
6.5.2.	Concluding remarks	194
Chapter 7.	References	195
Chapter 8.	Appendices	207
8.1.	Chapter 3 Appendix	208
8.1.1.	Figures	208
8.2.	Chapter 4 Appendix	212
8.2.1.	Figures and tables.....	212
8.3.	Chapter 5 Appendix	224
8.3.1.	MALDI and tandem mass spectra of peptide hits.....	224
8.3.2.	HPLC characterisation data of peptide hits from GL Biochem.....	242
8.3.3.	Figures	253

List of Figures

Figure 1.1. The molecular interaction between a proteinase and a substrate.	3
Figure 1.2. Domain architecture of MMPs described in this thesis.....	4
Figure 1.3. Domain architecture of the ADAMTS family.	6
Figure 1.4. Mechanism of inhibition of MMPs by endogenous and synthetic inhibitors.	10
Figure 1.5. FRET substrates are able to monitor proteinase activity over time.	16
Figure 1.6. Commercially available FRET substrates for monitoring proteinase activity <i>in vivo</i>	17
Figure 1.7. Osteoarthritis is a disease characterised by multiple structural changes. ...	19
Figure 1.8. Zonal arrangement of articular cartilage.....	23
Figure 1.9. Aggrecan possesses ADAMTS and MMP cleavage sites with unique amino acid preferences.	29
Figure 1.10. A comparison between normal and osteoarthritic articular cartilage.	37
Figure 1.11. Split-mix synthesis facilitates the synthesis of large compound libraries. .	42
Figure 1.12. The structure of PEGA resins is tailored to proteinase screening assays. ...	45
Figure 1.13. Peptides show unique fragmentation patterns during tandem mass spectrometry.....	49
Figure 3.1. Representative affinity purification of ADAMTS proteinases using stepwise pH elution shown for ADAMTS-5.....	90
Figure 3.2. Assessment of purity of pooled fractions of ADAMTS proteinases by western blot, Coomassie stain and silver stain.	91
Figure 3.3. Dose-dependent cleavage of FRET substrates by purified ADAMTS-4 and ADAMTS-5 proteinases.....	92
Figure 3.4. Determination of ADAMTS-4 and ADAMTS-5 concentrations through active-site titration with an endogenous TIMP-3 inhibitor.	93
Figure 3.5. FRET substrate hydrolysis by ADAMTS-4 and ADAMTS-5 proteinases over 24 hours.	95
Figure 3.6. Determination of the inhibition constant for a small molecule inhibitor of ADAMTS-4 and ADAMTS-5.	95

Figure 4.1. The effect of different surfactant concentrations on the rate of cleavage of FRET substrates by ADAMTS-4 and ADAMTS-5.	111
Figure 4.2. Selectivity studies of ADAMTS-4 and ADAMTS-5 FRET substrates with ADAMTS and MMP metalloproteinases.	113
Figure 4.3. The most commonly occurring residues by peptide position for known substrates of ADAMTS-5.	116
Figure 4.4. Representative images of FRET peptide substrates docked into the active site of the crystal structure of ADAMTS-5 using computational modelling.	120
Figure 4.5. Representative image of a FRET peptide substrate with a P7' fluorophore docked into the active site of the crystal structure of ADAMTS-5 using computational modelling.	124
Figure 4.6. Representative image of a FRET peptide substrate with a P9' fluorophore docked into the active site of the crystal structure of ADAMTS-5 using computational modelling.	128
Figure 4.7. Change in fluorescence intensity at different time points after incubation of on-resin FRET peptide substrates with ADAMTS-5.	135
Figure 4.8. Qualitative and quantitative assessment of the change in fluorescence intensity of a single resin bead containing a FRET peptide substrate before and after incubation with ADAMTS-5.	138
Figure 4.9. Design of a combinatorial library of FRET peptide substrates for screening against ADAMTS-5.	139
Figure 4.10. Qualitative time-course assessment of the change in on-resin fluorescence intensity of a combinatorial library of FRET peptide substrates before and after incubation with ADAMTS-5.	142
Figure 5.1. MALDI and tandem mass spectra of a peptide hit from ADAMTS-5 library screening.	153
Figure 5.2. Cleavage of the current substrate, parent compound and peptide hits by the ADAMTS-5 proteinase.	158
Figure 5.3. Kinetics of substrate cleavage by the ADAMTS-5 proteinase.	160
Figure 5.4. Representative affinity purification of proMMP-2ΔC through stepwise elution.	166
Figure 5.5. Representative purification of MMP-2ΔC fractions using Sephacryl affinity chromatography.	167
Figure 5.6. Determination of MMP-2ΔC concentration through active-site titration with an endogenous TIMP-3 inhibitor.	168

Figure 5.7. Selectivity studies of the current substrate, parent compound and hit peptides against ADAMTS and MMP proteinases.....	169
Figure 5.8. Comparison between the cleavage of library parent compound 5 and hit peptide 39 at different concentrations of substrate.....	171
Figure 5.9. Probing the lowest concentration of substrate that can detect ADAMTS-5 activity.....	172
Figure 5.10. Further investigation into the lowest concentration of substrate that can detect ADAMTS-5 activity.....	173
Figure 8.1. Titration curves for ADAMTS-4 data shown in Table 3.1.....	208
Figure 8.2. Titration curves for ADAMTS-5 data shown in Table 3.1.....	209
Figure 8.3. Determination of ADAMTS-4 and ADAMTS-5 concentrations listed in Table 3.3 through active-site titration with N-TIMP-3 and CT-1746.....	210
Figure 8.4. Determination of ADAMTS-4 and ADAMTS-5 amounts from different purification batches listed in Table 3.4.....	211
Figure 8.5. Calibration curves for the 5-FAM fluorophore at different concentrations of Brij-35 surfactant.....	212
Figure 8.6. Calibration curves for the Abz fluorophore at different concentrations of Brij-35 surfactant.	213
Figure 8.7. Cleavage of FRET substrate 1 by ADAMTS and MMP metalloproteinases.	214
Figure 8.8. Cleavage of FRET substrate 2 by ADAMTS and MMP metalloproteinases.	215
Figure 8.9. Activity of metalloproteinases from ADAMTS-4 and ADAMTS-5 selectivity studies.	216
Figure 8.10. Determination of MMP concentrations from ADAMTS-4 and ADAMTS-5 selectivity studies through active-site titration with N-TIMP-3 or CT-1746.....	217
Figure 8.11. HPLC Chromatograms of the small molecule fluorophore building block Fmoc-Lys(<i>retro</i> -Abz-Boc)-OH (22).	221
Figure 8.12. ¹ H-NMR spectrum of the small molecule fluorophore building block Fmoc-Lys(<i>retro</i> -Abz-Boc)-OH (22).	222
Figure 8.13. Change in fluorescence intensity at different time points after incubation of on-resin FRET peptide substrates in the presence or absence of ADAMTS-5.	223
Figure 8.14. Mass spectrum of the peptide hit KY(NO ₂)SEGESRGK(Abz)JYFKKG (31) obtained from ADAMTS-5 library screening.....	224

Figure 8.15. Tandem mass spectrum of the peptide hit KY(NO ₂)SEGESRGK(Abz)JYFKKG (31) obtained from ADAMTS-5 library screening.	225
Figure 8.16. Mass spectrum of the peptide hit KY(NO ₂)TENESRGK(Abz)IYYKKG (32) obtained from ADAMTS-5 library screening.....	226
Figure 8.17. Tandem mass spectrum of the peptide hit KY(NO ₂)TENESRGK(Abz)IYYKKG (32) obtained from ADAMTS-5 library screening.	227
Figure 8.18. Mass spectrum of the peptide hit KY(NO ₂)QDSESBK(Abz)JHYUKG (33) obtained from ADAMTS-5 library screening.....	228
Figure 8.19. Tandem mass spectrum of the peptide hit KY(NO ₂)QDSESBK(Abz)JHYUKG (33) obtained from ADAMTS-5 library screening.	229
Figure 8.20. Mass spectrum of the peptide hit KY(NO ₂)TEGESRGK(Abz)JZYKKG (34) obtained from ADAMTS-5 library screening.....	230
Figure 8.21. Tandem mass spectrum of the peptide hit KY(NO ₂)TEGESRGK(Abz)JZYKKG (34) obtained from ADAMTS-5 library screening.	231
Figure 8.22. Mass spectrum of the peptide hit KY(NO ₂)TETESOAK(Abz)IYFKKG (35) obtained from ADAMTS-5 library screening.....	232
Figure 8.23. Tandem mass spectrum of the peptide hit KY(NO ₂)TETESOAK(Abz)IYFKKG (35) obtained from ADAMTS-5 library screening.	233
Figure 8.24. Mass spectrum of the peptide hit KY(NO ₂)SESESRGK(Abz)IYYKKG (36) obtained from ADAMTS-5 library screening.....	234
Figure 8.25. Tandem mass spectrum of the peptide hit KY(NO ₂)SESESRGK(Abz)IYYKKG (36) obtained from ADAMTS-5 library screening.	235
Figure 8.26. Mass spectrum of the peptide hit KY(NO ₂)NDTESOAK(Abz)AHFUKG (37) obtained from ADAMTS-5 library screening.....	236
Figure 8.27. Tandem mass spectrum of the peptide hit KY(NO ₂)NDTESOAK(Abz)AHFUKG (37) obtained from ADAMTS-5 library screening.	237
Figure 8.28. Mass spectrum of the peptide hit KY(NO ₂)NESESKAK(Abz)VHYKKG (38) obtained from ADAMTS-5 library screening.....	238
Figure 8.29. Tandem mass spectrum of the peptide hit KY(NO ₂)NESESKAK(Abz)VHYKKG (38) obtained from ADAMTS-5 library screening.	239
Figure 8.30. Mass spectrum of the peptide hit KY(NO ₂)SENESRGK(Abz)IYYKKG (39) obtained from ADAMTS-5 library screening.....	240
Figure 8.31. Tandem mass spectrum of the peptide hit KY(NO ₂)SENESRGK(Abz)IYYKKG (39) obtained from ADAMTS-5 library screening.	241

Figure 8.32. Cleavage of the current substrate, library parent compound and peptide hits by the ADAMTS-5 proteinase.	253
Figure 8.33. Michaelis-Menten saturation curves for cleavage of the current substrate 2 by ADAMTS-5.....	254
Figure 8.34. Michaelis-Menten saturation curves for cleavage of the library parent compound 5 by ADAMTS-5.	255
Figure 8.35. Michaelis-Menten saturation curves for cleavage of hit peptide 32 by ADAMTS-5.	256
Figure 8.36. Michaelis-Menten saturation curves for cleavage of hit peptide 36 by ADAMTS-5.	257
Figure 8.37. Michaelis-Menten saturation curves for cleavage of hit peptide 39 by ADAMTS-5.	258
Figure 8.38. Cleavage of the library parent compound and peptide hits by ADAMTS-4.	259
Figure 8.39. Cleavage of the library parent compound and peptide hits by MMP-2. ...	260
Figure 8.40. Cleavage of the library parent compound and peptide hits by MMP-9. ...	261

List of Tables

Table 3.1. Summary of inhibition constants determined for CT-1746 inhibition of ADAMTS-4 and ADAMTS-5.	97
Table 3.2. Summary of inhibition constants determined for CT-1746 inhibition of ADAMTS-4 outside the linear range of substrate hydrolysis.....	97
Table 3.3. Summary of active ADAMTS-4 and ADAMTS-5 concentrations obtained by active-site titration with endogenous and small molecule inhibitors.	98
Table 3.4. Summary of purified amounts of ADAMTS-4 and ADAMTS-5 proteinases determined through active-site titration with N-TIMP-3.	99
Table 4.1. Summary of FRET peptide substrates containing a P4' fluorophore docked into the active site of the crystal structure of ADAMTS-5.	121
Table 4.2. Summary of FRET peptide substrates containing a P7' fluorophore docked into the active site of the crystal structure of ADAMTS-5.	125
Table 4.3. Summary of FRET peptide substrates containing a P9' fluorophore docked into the active site of the crystal structure of ADAMTS-5.	129
Table 4.4. Summary of characterisation data for ADAMTS-5 FRET peptide substrates synthesised by manual Fmoc-SPPS using PEGA resin.....	133
Table 5.1. Peptide hits identified through MALDI and tandem mass spectrometry	155
Table 5.2. Kinetic parameters for cleavage of the current substrate, library parent compound and selected hit peptides by the ADAMTS-5 proteinase.	162
Table 8.1. Known peptide substrates cleaved by ADAMTS-5 compiled from CutDB, BRENDA and MEROPS online databases.	218

List of Schemes

Scheme 3.1. Selection of cells, expression of ADAMTS proteinases and collection of conditioned medium for affinity purification.	88
Scheme 4.1. Comparison between on-resin screening of combinatorial library designs containing a FRET peptide substrate with an N-terminal or C-terminal fluorophore. .	115
Scheme 4.2. Multi-step design process for the development of an P4' fluorophore series of ADAMTS-5 FRET peptide substrates informed by computational docking.	119
Scheme 4.3. Two-step design process for the development of a P7' fluorophore series of ADAMTS-5 FRET peptide substrates informed by computational docking.	123
Scheme 4.4. Two-step design process for the development of a P9' fluorophore series of ADAMTS-5 FRET peptide substrates informed by computational docking.	127
Scheme 4.5. Synthesis of the small molecule fluorophore building block Fmoc-Lys(<i>retro</i> -Abz-Boc)-OH used in Fmoc-SPPS of ADAMTS-5 FRET peptide substrates.	130
Scheme 4.6. Synthetic route to ADAMTS-5 FRET peptide substrates through manual Fmoc-SPPS using PEGA resin.	131
Scheme 4.7. Determination of the extent of cleavage of synthesised FRET peptide substrates on-resin after incubation with ADAMTS-5.	134
Scheme 4.8. Incubation of a combinatorial library of FRET peptide substrates with ADAMTS-5, deactivation of the proteinase and isolation of hits.	141

List of abbreviations and symbols

λ_{em}	emission wavelength
λ_{ex}	excitation wavelength
ΔC	catalytic domain
$^1\text{H-NMR}$	proton-nuclear magnetic resonance spectroscopy
1-Melm	1-Methylimidazole
5-FAM	5-Carboxyfluorescein
A_{280}	absorbance at 280 nm
Abz	ortho-aminobenzoic acid/ortho-aminobenzoyl
AcOH	acetic acid
ADAMTS	a disintegrin-like and metalloproteinase with thrombospondin type-I motifs
Boc	<i>tert</i> -butyloxycarbonyl
CaCl_2	calcium chloride
CBB	Coomassie brilliant blue
CCD	charge-coupled device
Cm-Tf	carboxymethylated transferrin
COMP	cartilage oligomeric matrix protein
COMU	(1-Cyano-2-ethoxy-2-oxoethylidenaminoxy)dimethylamino-morpholino-carbenium hexafluorophosphate
CS	chondroitin sulphate
CT	X-ray computed tomography
CV	column volume
Cy5.5	cyanine 5.5

DCM	dichloromethane
DIPEA	<i>N,N</i> -diisopropylethylamine
DMEM	Dulbecco's modified eagle's medium
DMF	dimethylformamide
DMM	destabilisation of the medial meniscus
DMMB	dimethylmethylene blue assay
DMSO	dimethyl sulphoxide
Dnp	2,4-Dinitrophenyl
Dpa	<i>N</i> -3-[2,4-dinitrophenyl]-L-2,3-diaminopropionyl
ECM	extracellular matrix
ELISA	enzyme-linked immunosorbent assay
EtOAc	ethyl acetate
FA	formic acid
FBS	fetal bovine serum
Fmoc-SPPS	9-Fluorenylmethoxycarbonyl-based solid-phase peptide synthesis
FPLC	fast protein liquid chromatography
FRET	Förster resonance energy transfer
G1/2/3	globular domain 1/2/3
GAG	glycosaminoglycan
H ₂ SO ₄	sulphuric acid
HEK293	human embryonic kidney 293
HMBA	4-Hydroxymethylbenzoic acid
HPLC	high-performance liquid chromatography
IGD	interglobular domain

IL-1/IL-6	interleukin-1/-6
JSN	joint space narrowing
k_{cat}	catalytic rate constant
KCl	potassium chloride
K_i	inhibition constant
$K_{i(\text{app})}$	apparent inhibition constant
K_m	Michaelis constant
KO	knockout
KS	keratan sulphate
LEH	lactalbumin enzymatic hydrolysate
LRMS	low resolution mass spectrometry
MALDI	matrix-assisted laser desorption ionisation
Mca	7-Methoxycoumarin-4-acetic acid/7-Methoxycoumarin-4-yl-acetyl
MeCN	acetonitrile
MeOH	methanol
MMP	matrix metalloproteinase
MOE	Molecular Operating Environment
MPEG	methoxy-poly(ethylene)-glycol
MPM	microparticle matrix
MRI	magnetic resonance imaging
mRNA	messenger RNA
MSNT	1-(Mesitylene-2-sulfonyl)-3-nitro-1H-1,2,4-triazole
M_w	molecular weight
NaCl	sodium chloride

NaHCO ₃	sodium hydrogen carbonate
NaN ₃	sodium azide
NEM	4-Ethylmorpholine
NIRF	near-infrared fluorophore
N-TIMP-3	N-terminal domain of tissue inhibitor of metalloproteinases-3
OA	osteoarthritis
OBOC	one bead one compound
PAP	protease-activated prodrug
PEGA	poly(ethylene glycol)-polydimethyl acrylamide
PET	positron emission tomography
PL	poly-L-lysine
PS	penicillin-streptomycin
PVDF	polyvinylidene difluoride
Q	quencher
RFU	relative fluorescence units
RT	room temperature
SD	standard deviation
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	standard error of the mean
siRNA	small interfering RNA
SPD-NGS	substrate phage display–next generation sequencing
TAMRA	<i>N,N,N',N'</i> -tetramethyl-6-carboxyrhodamine
TBST	19 mM tris base, 137 mM NaCl, 2.7 mM KCl and 0.05% (v/v) Tween 20

TBTU	<i>N,N,N',N'</i> -tetramethyl-O-(benzotriazol-1-yl)uronium tetrafluoroborate
TEA	triethylamine
TEMED	<i>N,N,N',N'</i> -tetramethylethane-1,2-diamine
TFA	trifluoroacetic acid
TIMPs	tissue inhibitors of metalloproteinases
TNC	50 mM tris base, pH 7.5, 150 mM NaCl, 10 mM CaCl ₂ and 0.02% (w/v) NaN ₃
TNF- α	tumour necrosis factor- α
Tof	Time-Of-Flight
TS4-4	ADAMTS-4 catalytic and disintegrin domains
TS5-5	ADAMTS-5 catalytic and disintegrin domains
UHPLC	ultra-high performance liquid chromatography
Y(NO ₂)	3-Nitro-L-tyrosine
ZnCl ₂	zinc chloride

The abbreviations for the common amino acids are in accordance with the recommendations of the IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN).

Chapter 1. Introduction

1.1. Preamble: proteinases and diseases

Proteinases are enzymes that play a diverse array of metabolic roles in the body, ranging from the activation of various proteins, hormones and receptors to the degradation of food proteins that release essential amino acids (1). They are also involved in a number of diseases, either through affecting intracellular processes within the cell, or by contributing to the degradation of the extracellular matrix (ECM) of tissues (2, 3). Proteinases are thus important in a plethora of diseases, including those resulting from inborn errors of metabolism (*e.g.* albinism) to those of degenerative diseases of aging (3, 4). One highly prevalent disease of aging is osteoarthritis (OA), where increased activity of catabolic proteinases belonging to the a disintegrin-like and metalloproteinase with thrombospondin type-I motifs (ADAMTS) and matrix metalloproteinase (MMP) families are responsible for the degradation of cartilage ECM lining the surfaces of joints, leading to pain and disability (3, 5). This thesis will discuss how knowledge of these proteinases was used to develop a tool to monitor their activity in a variety of settings. These settings could include animal models of diseases such as osteoarthritis, with potential application in the diagnosis, prognosis and evaluation of therapeutic responses in patients with OA.

1.2. Proteinases and metalloproteinases

1.2.1. Classification of proteinases

Proteinases may be classified according to their domains using the highly detailed MEROPS classification system (6). In addition, they can also be classified by their mechanism of action or the specificity of their cleavage (1).

When an amide (peptide) bond is cleaved in a protein, it typically occurs through the hydrolytic action of a peptidase (1, 7). Depending on the location of the bond being broken, peptidases may be subcategorised into “exopeptidases” (*i.e.* cleavage of the amide bond

occurs at the free N- or C-terminus of the peptide sequence), or “endopeptidases” (*i.e.* cleavage of the amide bond occurs within the peptide sequence) (1, 7). Endopeptidases are more commonly referred to as “proteinases” (1, 7) and cleavage of a peptide or protein substrate occurs in the active site, where the different binding pockets (subsites, “S”) become occupied by the unique side chains of the substrate (positions, “P”), as illustrated in Figure 1.1 (8).

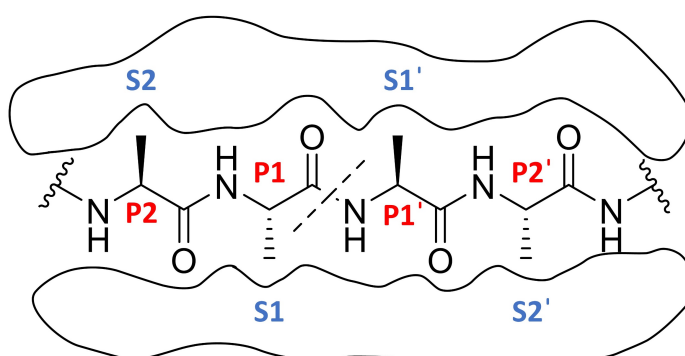


Figure 1.1. The molecular interaction between a proteinase and a substrate. Schematic representation of the molecular interaction between a substrate and a proteinase in the active site. The cleavage site of the substrate is indicated by a dotted line, with both peptide positions (P) and proteinase subsites (S) numbered with or without a prime to the right or left of this site, respectively (8). *Please note that only part of the substrate is shown for clarity.*

Another major classification of proteinases is based on their catalytic type, which determines the mechanism by which they hydrolyse substrates (1). For example, cathepsins are proteinases which utilise a catalytic serine, cysteine, or aspartate residue to process hormones and neuropeptides; along with degrading collagens in cartilage and bone (9). Other examples include the metalloproteinases, which utilise a catalytic zinc ion (*i.e.* Zn^{2+}) to promote substrate hydrolysis (1). Among this group are the MMPs and ADAMTSs mentioned earlier. The MMPs play a crucial role in reproduction, embryogenesis, morphogenesis, and the repair and remodelling of tissues (10, 11); whereas on top of embryogenesis and morphogenesis, the ADAMTSs are also involved in vascular biology and inflammation (12,

13). The structure and function of these two families of metalloproteinases will be discussed in more detail in the next section.

1.2.2. Metalloproteinases

1.2.2.1. MMPs: structure and function

MMPs can be grouped based on their domain organisation. Figure 1.2 shows the general domain organisation of the MMPs described in this thesis.

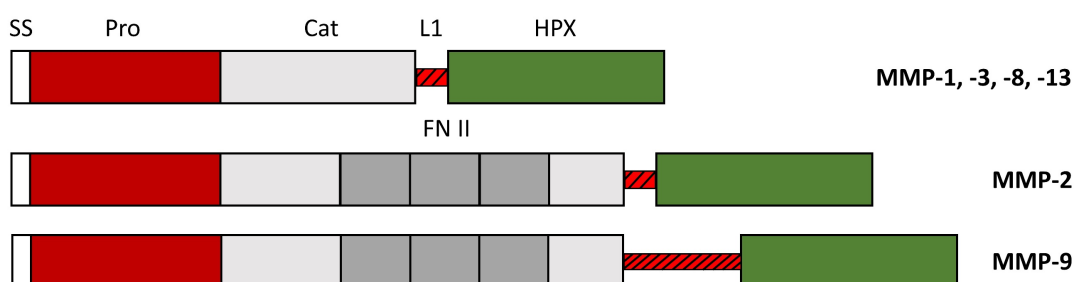


Figure 1.2. Domain architecture of MMPs described in this thesis. Representative figure of the different domains, which are labelled in the following manner: SS = signal sequence; Pro = pro-domain; Cat = catalytic domain; L1 = linker region 1; HPX = hemopexin-like region; FN II = fibronectin type II motif. This figure is drawn to scale and was derived from Nagase *et al* (10, 11).

MMPs are composed of a signal peptide (SS), which is deleted during the course of translation to generate the inactive proMMP proteinase or zymogen. The pro-domain (Pro) of this zymogen is comprised of ~80 residues. The adjacent catalytic domain (Cat) has ~170 amino acids and contains a conserved Met residue eight amino acids downstream from the zinc binding region HEXXHXXGXXH (Zn^{2+}), where X is any amino acid. This forms a “Met-turn” that supports the area near the catalytic Zn^{2+} ion. This Cat domain is followed by a linker region (L1) whose length (when present) differs depending on the MMP. A hemopexin-like domain (HPX) completes the domain organisation and consists of ~200 residues. Of the MMPs discussed in this thesis, the exceptions to this typical domain architecture are MMP-2 and MMP-9, both of which contain an additional fibronectin type II sequence (FN II) within the Cat domain (10, 11).

MMPs can also be grouped based on the preferred substrates that they cleave. These groups include collagenases (*e.g.* MMP-1, -8, -13), gelatinases (*e.g.* MMP-2, -9) and stromelysins (*e.g.* MMP-3). In terms of their proteolytic functions, collagenases degrade type I, II and III collagens into smaller fragments of different lengths. On the other hand, gelatinases bind to gelatin through their fibronectin type II motifs and digest it. They also show collagenase activity against type IV, V and XI collagens, with MMP-2 also digesting type I, II and III collagens. Stromelysins such as MMP-3 (stromelysin-1) degrade several molecules within the ECM including fibronectin and laminin (10, 11).

1.2.2.2. ADAMTSs: structure and function

Like MMPs, ADAMTSs may also be grouped by their domain organisation. Figure 1.3 shows the domain organisation of the complete ADAMTS family.

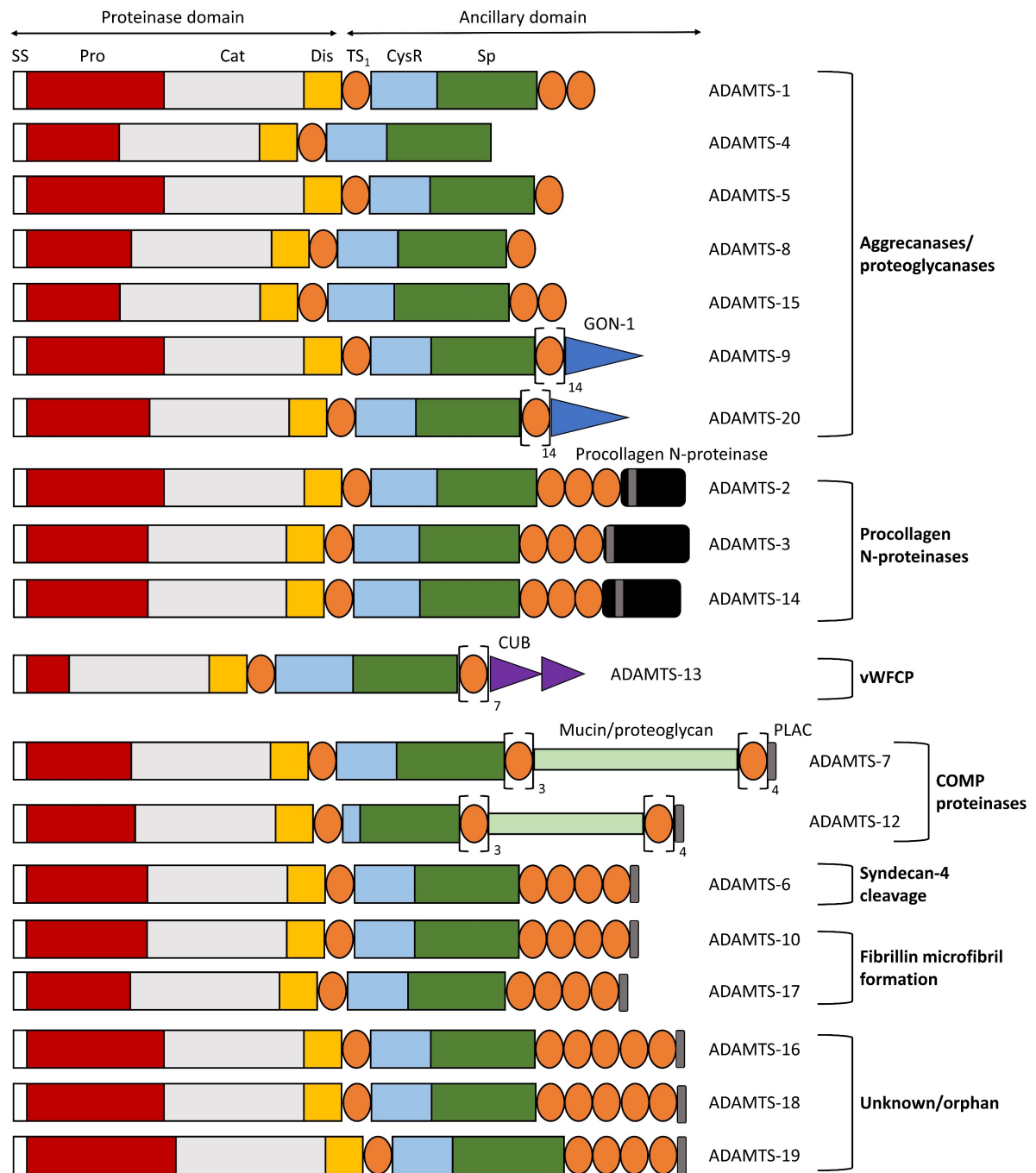


Figure 1.3. Domain architecture of the ADAMTS family. Representative figure showing the different domains, which are abbreviated as follows: SS = signal peptide; Pro = pro-domain; Cat = catalytic domain; Dis = disintegrin-like domain; TS₁ = thrombospondin type-I motif; CysR = cysteine-rich domain and Sp = spacer domain; GON-1 = from *C. elegans gon-1* gene; CUB = complement C1r/C1s, Uegf, Bmp1; PLAC = protease and lacunin. Key to proteinase groups: vWFCP = von Willebrand factor cleaving proteinase; COMP = cartilage oligomeric matrix protein. This figure is drawn to scale and was derived from Kelwick *et al* (13).

ADAMTSs are comprised of a signal sequence (SS), which targets the proteinase to the secretory pathway and is removed during translation. The next domain is the pro-domain (Pro), which differs in length between the members of the family (180–260 residues) and is involved in managing correct folding of the proteinase (13, 14). Removal of this domain occurs through cleavage *via* pro-protein convertases such as furin. For example, in the case of proADAMTS-1/-4, this activation occurs in the Golgi apparatus, whereas for proADAMTS-5 this takes place in the ECM. In the case of proADAMTS-9, this occurs on the cell surface (13). The catalytic domain (Cat) includes a consensus HEXXHXBG(/N/S)BXHD sequence, where the Zn^{2+} cation is coordinated by three histidine residues and B denotes a large non-polar amino acid. Much like the MMPs, a Met amino acid lies 12-24 residues downstream from the final histidine in this sequence and forms a Met-turn (a hydrophobic area underneath the His residues); a signature common to all 19 members of the ADAMTS family (13, 15). The disintegrin-like domain appears to be a misnomer, as it does not interact with integrins but consists of an area rich in cysteine residues that forms the side of the active-site.

The subsequent domains are considered “ancillary” as they regulate ADAMTS activity by controlling association with the ECM and defining substrate binding, examples of which include ADAMTS-4 and ADAMTS-5 (13, 16, 17). The first of these modules—namely the thrombospondin type-I motif (TS_1)—comprises ~50 residues and displays similarity to the thrombospondin 1 and 2 type-I repeats (13). This is followed by a cysteine-rich region (CysR) with ~100 residues (of which 10 consist of conserved cysteines), and a spacer domain (Sp) consisting of 103–160 amino acids where cysteines are absent. This general structure is only found in ADAMTS-4, as other ADAMTS proteinases differ in the number and type of ancillary domains following the Sp domain. Additional domain types include: a PLAC module with six

conserved cysteine residues (ADAMTS-2, -3, -6, -7, -10, -12, -14, -16, -17, -18, -19); a procollagen N-proteinase domain (ADAMTS-2, -3, -14); a mucin/proteoglycan domain (ADAMTS-7, -12); CUB domains (ADAMTS-13); and a GON-1 domain with 10 conserved cysteine residues (ADAMTS-9, -20) (13).

ADAMTSs can additionally be grouped by their preferential cleavage of certain substrates. The largest group is the aggrecanase or proteoglycanase group (ADAMTS-1, -4, -5, -8, -9, -15, -20) which degrades aggrecan and other proteoglycans such as brevican, neurocan and versican (13). Another group is that of the procollagen N-proteinases (ADAMTS-2, -3, -14), which generate type I, II, III and V collagen fibrils through removal of the N-terminal pro-peptide (13, 18). Additionally, ADAMTS-7 and ADAMTS-12 cleave the cartilage oligomeric matrix protein (COMP), which is a structural constituent of cartilage ECM (13, 18); and ADAMTS-13 plays a key role in blood coagulation by cleaving the pre-proteins of the von-Willebrand factor to generate active proteins that are involved in this process (13, 19). ADAMTS-6 appears to cleave syndecan-4 and may regulate transforming growth factor beta signalling, while ADAMTS-10 and ADAMTS-17 play a role in the formation of fibrillin microfibrils—but only ADAMTS-10 is currently known to proteolyse fibrillin-1/-2 (20). The remainder of the ADAMTS family (ADAMTS-16, -18, -19) have unknown substrates (19, 20).

1.2.2.3. The activity of ADAMTSs and MMPs is regulated by tissue inhibitors of metalloproteinases

The activity of ADAMTSs and MMPs is controlled by endogenous inhibitors known as tissue inhibitors of metalloproteinases (TIMPs), of which four have been described in humans: TIMP-1, TIMP-2, TIMP-3 and TIMP-4 (21). TIMPs are 188–194 residues in length and are comprised of an N-terminal- and C-terminal domain, each of which has three conserved disulphide bonds (11). The N-terminal domain can fold and function as a separate inhibitor

of MMPs (*e.g.* N-TIMP-3) (11, 21). TIMPs inhibit all MMPs—although to varying degrees—but only TIMP-3 inhibits ADAMTSs (11, 21, 22). The mechanism of MMP inhibition is shown in Figure 1.4A and is typically based on two processes occurring in the active site of the proteinase: i) coordination of the Zn^{2+} ion by both the N-terminal amino group and carbonyl group of the first Cys residue of the TIMP; and ii) an interaction between the hydroxyl group of the second Ser/Thr residue with the nucleophilic glutamate residue of the MMP, which expels the water nucleophile responsible for amide bond hydrolysis from the active site (22). Interestingly, N-TIMP-3 inhibition of ADAMTS-4 and ADAMTS-5 does not appear to rely on the coordination of Zn^{2+} by the N-terminal cysteine residue mentioned above (23). This is because the addition of up to two N-terminal alanine residues maintains the inhibition constant at < 100 nM for ADAMTSs compared to MMPs, where it is typically > 1 μM (23). In contrast, small molecule hydroxamic acid inhibitors such as CT-1746 coordinate to Zn^{2+} *via* the syn (*i.e.* Z) rotamer of the hydroxamate anion, which results in displacement of the water molecule from the active site of the proteinase (24) (Figure 1.4B).

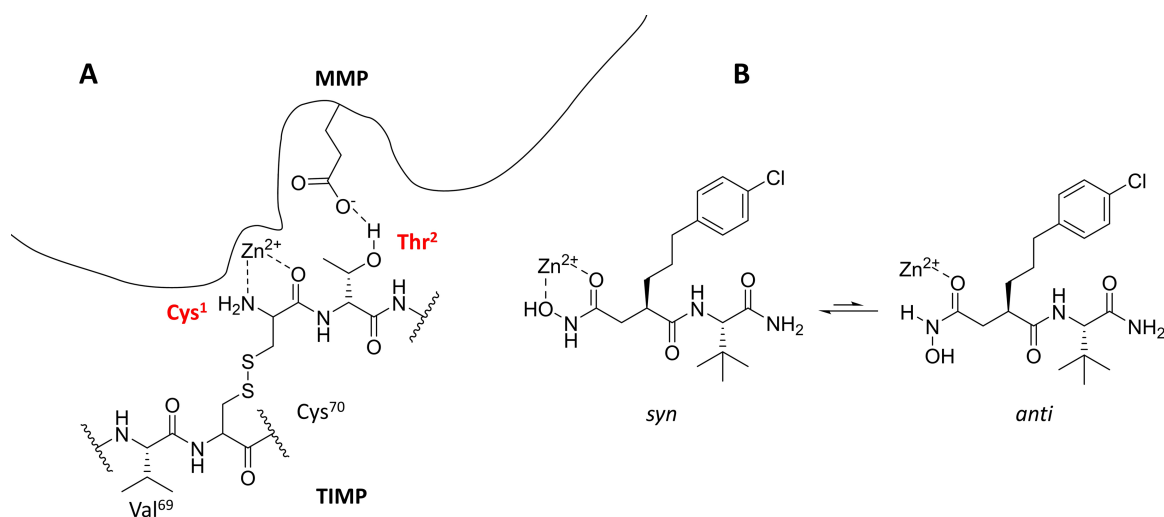


Figure 1.4. Mechanism of inhibition of MMPs by endogenous and synthetic inhibitors. MMPs are inhibited by endogenous TIMP inhibitors *via* a mechanism involving the coordination of the Zn^{2+} ion and an interaction with the glutamate amino acid of the MMP (A). Small molecule inhibitors such as CT-1746 inhibit MMPs through coordination of the Zn^{2+} ion *via* the syn rotamer (B). Part A of this figure was derived from Murphy *et al* (22).

1.3. Detecting proteinase activity

1.3.1. Methods used in chemistry and biochemistry

To develop a tool for monitoring the activity of an ADAMTS or MMP proteinase it is important to consider the available methods. Proteinase activity can be detected using a number of analytical techniques. These can be sub-divided into those more frequently associated with chemistry and those more closely tied to biochemistry, and each of these will be considered in turn.

One way to observe proteinase activity would be to employ high-performance liquid chromatography (HPLC). In this experiment, a substrate would be incubated with a proteinase in a reaction vessel for a fixed period of time along with a separate control containing substrate alone. To ascertain whether substrate cleavage had taken place, aliquots would be collected at regular timepoints, analysed by HPLC and compared to the

control. Fragment peaks resulting from substrate hydrolysis would be detected through increased absorbance on the HPLC chromatogram compared to the absence of such peaks in the control sample. These peaks could be integrated to determine the percentage conversion of substrate to product, which would give an indication of the reaction rate. In addition, to determine the molecular masses of the fragments being generated, HPLC could be coupled with mass spectrometry (HPLC-MS). The success of these techniques would depend on the ability of the substrate and its fragments to elute from the column and be detected by both absorbance and mass. A similar approach would be to use proton NMR (^1H -NMR) spectroscopy as the monitoring tool. However, the spectra of the reaction mixture would likely be difficult to analyse due to the quantity and complexity of couplings present.

In a similar vein, common biochemistry methods such as sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (25) could be employed with Coomassie staining to follow the appearance of low molecular weight (M_w) bands resulting from the cleavage of a synthetic proteinase substrate such as carboxymethylated transferrin (Cm-Tf) (16, 17). Additionally, western blotting could also be used following SDS-PAGE to detect the specific degradation fragments resulting from the proteolysis of substrates over time (26).

Other examples of techniques for determining proteinase activity are focused on detecting fragments resulting from the proteolysis of natural, full-length substrates. For example, aggrecan is a proteoglycan that consists of a core protein bound to negatively charged polysaccharide side chains of keratan sulphate (KS) and chondroitin sulphate (CS)—known as glycosaminoglycans (GAGs) (27, 28). The degradation of aggrecan by ADAMTSs results in the release of GAGs, which can be detected through the dimethylmethylene blue (DMMB) assay (29, 30). Similarly, the hydroxyproline assay detects

hydroxyproline discharged through cleavage of type II collagen by collagenase MMPs (31, 32).

Unfortunately, these techniques suffer from inherent limitations. On the one hand, the DMMB and hydroxyproline assays determine proteinase concentrations by measuring absorbance at specific wavelengths (*i.e.* 525 nm and 560 nm, respectively), which can be influenced by fluctuations in buffer and pH (29-32). In addition, they also possess a poor linear range of detection, namely 0.5–5 µg of GAGs for the DMMB assay (33) and 0.2–1.0 µg of hydroxyproline for the hydroxyproline assay (34). On the other hand, the detection sensitivity of Coomassie staining is limited to protein amounts > 50 ng (35), while the effectiveness of western blotting depends on the nature and sensitivity of the antibodies used.

Most crucially of all, while all of the aforementioned methods used in both chemistry and biochemistry detect the end products of proteinase activity, they do not measure activity directly. The development of a tool which would enable the visualisation and monitoring of proteinase activity *in situ* in real time would enable a better understanding of the mechanism of action of proteinases, their role in biological processes, and how their activity in both healthy and diseased tissues *ex vivo* (or subjects *in vivo*) might change over time. Such a tool would be non-invasive and can be developed through molecular imaging.

1.3.2. Measuring proteinase activity through molecular imaging

The field of molecular imaging emerged in the mid-1990s (36-39) and may be defined as the non-invasive visualisation and characterisation of biological processes at both the molecular and cellular level in living systems (37, 40, 41). It covers a variety of disciplines including chemistry, biology, physics, engineering, and medicine, amongst others. The most common imaging techniques that it utilises may be sub-divided into pre-clinical modalities

such as optical imaging and clinical modalities including X-ray computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET) (41). The greatest strength of molecular imaging lies in its ability to visualise biochemical processes occurring in living organisms *in situ* without disturbing the system under examination (41). This makes it well suited to measuring proteinase activity. Nonetheless, it is important to consider the most appropriate imaging modality to use for this task in order that activity is measured most effectively.

Among the clinical imaging modalities, CT is the most commonly used as it is inexpensive, widely available and has an excellent spatial resolution (0.5–1 mm clinical, 2.3 μm pre-clinical) (41, 42). In this technique, the contrast in each image is generated through the differential absorption of X-rays by tissues in the body (41). As soft tissues have poor X-ray absorption, a number of radiocontrast agents have been developed to improve their visualisation through the use of radiopaque elements such as iodine (43). Unfortunately, these contrast agents have a poor detection sensitivity (2 mg iodine/ml),¹ and thus detecting low levels of proteinase activity would require a large amount of contrast agent, thereby increasing the likelihood of renal toxicity (41).

On the other hand, MRI generates images based on the alignment of protons to an applied magnetic field (41). Smart MRI contrast agents have been developed for measuring the activity of proteinases (44-46). These typically consist of a peptide substrate connected to a chelator complexed to a paramagnetic lanthanide ion (*e.g.* thulium, Tm^{3+}). Cleavage of the peptide substrate results in a change in the rate of exchange of protons near the chelator with those of surrounding water molecules, whose proton magnetic resonance signal

¹ Fowkes, M.M.; Das Neves Borges, P. *et al*, unpublished work.

subsequently changes (47). This effect has been utilised to detect the activity of caspases (44, 45) and cathepsins (46). Similar to CT contrast agents however, these probes suffer from poor detection sensitivity (1–10 mM) (44). In addition, MRI is expensive and generally not widely available.

In PET, the signal is detected based on the release of photons from an annihilation reaction (41). This occurs when a positron emitted from a radioisotope collides with an electron from the surrounding tissues. This is perhaps the most suitable imaging modality for monitoring proteinase activity as it possesses excellent detection sensitivity (10^{-11} – 10^{-12} M) (41). As a consequence of this, several different radiotracers have been developed for detecting proteinase activity. For example, a dual contrast agent consisting of a cell-penetrating peptide and a cleavable targeting peptide connected to a chelated [^{68}Ga] radioisotope and fluorophore was used to image cathepsin D activity in a mouse model of Alzheimer's disease (48). In addition, selective inhibitors have been developed for imaging MMP-12 (49, 50). These have included small molecules radiolabelled with [^{18}F] (49) and peptidomimetic compounds radiolabelled with [$^{99\text{m}}\text{Tc}$], which were used to image MMP-12 activity in a murine model of inflammatory disease (50). Unfortunately, the synthesis of PET radiotracers requires efficient radiolabelling methods in place due to rapid decay of the radioisotope, and these can be challenging and time-consuming to establish. Furthermore, access to a cyclotron is also needed to produce the short-lived radioisotope *in situ*.

Optical fluorescence imaging records the emission of light at a particular wavelength resulting from excitation of a fluorophore at a different wavelength (41). It addresses some of the combined limitations of CT, MRI and PET in that it is relatively inexpensive and possesses excellent detection sensitivity (10^{-9} – 10^{-12} M) without the need for laborious radiolabelling procedures (41). The limited spatial resolution of optical fluorescence imaging

compared to CT (*cf.* 2–3 mm vs 2.3 μm) does not pose a problem, as it would not be used to visualise detailed morphology (41, 42). Furthermore, while MRI and PET probes are able to measure proteinase activity much like the analytical methods described earlier in section 1.3.1, they are unable to monitor this activity dynamically over time. Optical fluorescence imaging is able to accomplish this with Förster resonance energy transfer (FRET) substrates (51).

1.3.2.1. Monitoring proteinase activity with FRET peptide substrates

FRET substrates generally consist of a peptide sequence with a near-infrared fluorophore (NIRF) and quencher (Q) flanking the proteolysis site (Figure 1.5) (52, 53). In its inactive state the probe is optically silent, as fluorescence from the NIRF (the donor) is extinguished by the quencher (the acceptor) over distances between 10–100 Å (54). For *in vivo* imaging, NIRFs are preferred to conventional fluorophores as the higher wavelength of NIR light results in decreased tissue auto-fluorescence and absorption by water, lipids and haemoglobins (55). This leads to a depth of penetration of < 1 cm, which is sufficient for use in pre-clinical small animal models of disease (52).

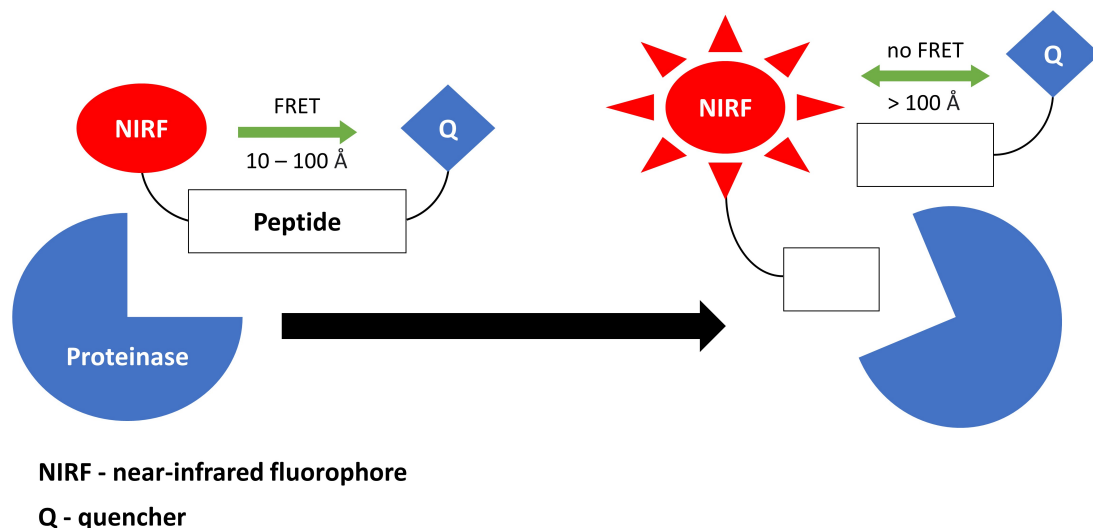


Figure 1.5. FRET substrates are able to monitor proteinase activity over time. A FRET substrate generally consists of a peptide sequence bordered by a near-infrared fluorophore (NIRF) and quencher (Q). The activity of a target proteinase can be measured following cleavage of this substrate, as this results in separation of the NIRF and Q and thus an increase in fluorescence intensity. Continuous proteolysis of the substrate leads to an increase in signal intensity, which can be used to monitor proteinase activity *via* an optical imaging system.

When a proteinase identifies and cleaves the peptide substrate, a rise in fluorescence occurs by virtue of an increase in space between the NIRF and Q ($> 100 \text{ \AA}$) and thus disruption of the FRET process (52). By measuring the relative fluorescence produced at a specific wavelength, the real-time activity of the proteinase can be monitored with each successive peptide that is cleaved, thereby resulting in signal amplification and enhanced detection *in vivo* (52). Signal amplification is especially important for proteinases with a low abundance, as it allows them to be detected. To date, a considerable number of FRET substrates have been developed for monitoring proteinase activity *in vivo*. These include commercially available probes such as ProSense 680 and Cat K 680 for cathepsins, and MMPSense 680 and MMPSense 750 for MMPs (Figure 1.6) (56-61). However, no such substrates are currently available for ADAMTSs for *in vivo* applications and existing *in vitro* probes such as ortho-aminobenzoyl (Abz)-TESE~SRGAIY-*N*-3-[2,4-dinitrophenyl]-L-2,3-diaminopropionyl (Dpa)-KK-NH₂ (62), 5-carboxyfluorescein (5-FAM)-AE~LQGRPISIAK-*N,N,N',N'*-tetramethyl-6-

carboxyrhodamine (TAMRA) (63), K(6-FAM)-DVQE~FRGVTAVIRC(QSY-9)KGK (64) and Abz-TEGE~ARGSVI-Dpa-KK-NH₂ (65) lack sufficient selectivity.

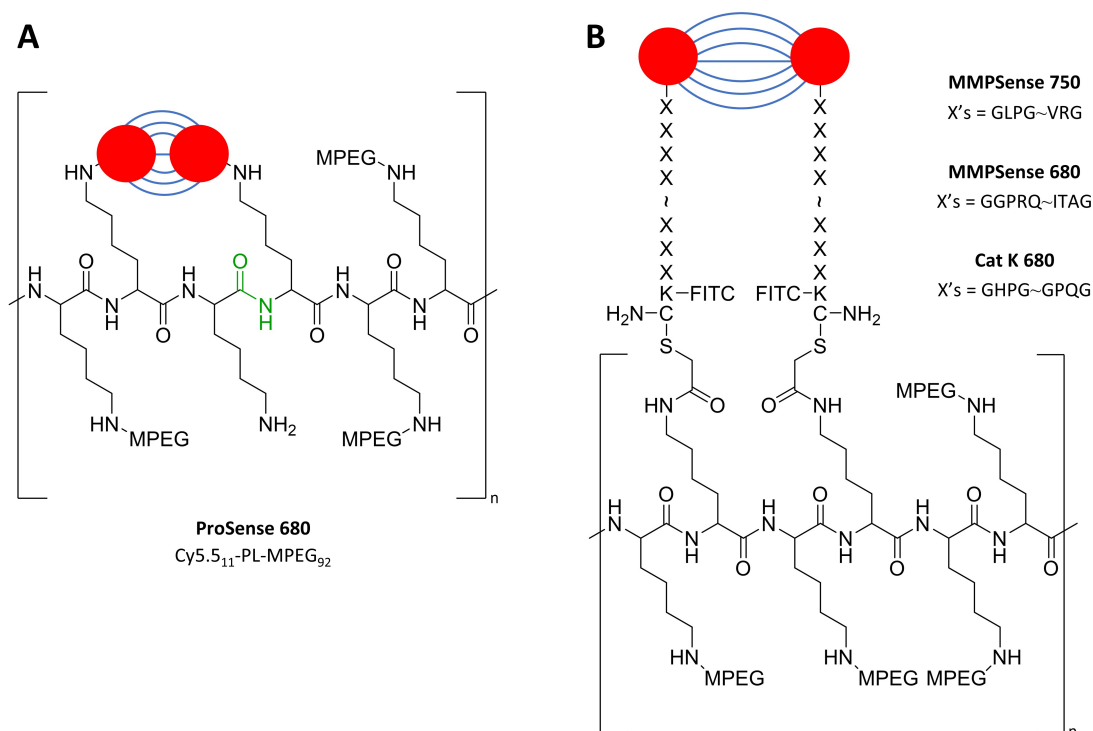


Figure 1.6. Commercially available FRET substrates for monitoring proteinase activity *in vivo*.

Examples of commercially available probes for *in vivo* imaging of proteinases include ProSense 680 for cathepsins, MMPsense 750 and MMPsense 680 for MMPs and Cat K 680 for cathepsin K. In each case, the probe consists of a graft co-polymer with a poly-L-lysine (PL) backbone and methoxy-poly(ethylene glycol) (MPEG) side chains with either a pair of Cyanine5.5 (Cy5.5) fluorophores (red circles) attached to free lysine residues in each repeating unit (A); or the absence of free lysine residues and a pair of fluorescein isothiocyanate fluorophores attached to lysine side chains with a cleavable peptide sequence indicated by "X"s (B). The cleavage sites are indicated with a green highlighted amide bond for (A) and with a ~ for (B). This figure was adapted from Weissleder *et al* (57) and Tung *et al* (58).

From a disease perspective, there is also an unmet clinical need for the development of an ADAMTS imaging probe for osteoarthritis; to detect the activity of catabolic ADAMTS proteinases that degrade the ECM of articular cartilage (3). The primary aim of this thesis is

thus to develop a selective ADAMTS FRET substrate as a tool for monitoring proteinase activity. This aim and the direct clinical demand for such a probe in OA will become clearer following an exploration of this disease and the specific ADAMTS and MMP proteinases involved.

1.4. Osteoarthritis

1.4.1. Introduction

Osteoarthritis is the most prevalent type of joint disease, with more than 300 million people affected worldwide (5, 66). In the UK, it is estimated that 33% of individuals over the age of 45 sought treatment for this condition between 2004 and 2010 (67, 68). Osteoarthritis is characterised by structural changes within the joint, including the degradation of articular cartilage, sclerosis of underlying subchondral bone, the establishment of osteophytes, inflammation of the synovium and hypertrophy of the joint capsule (5, 69).² These changes are illustrated for a human knee joint in Figure 1.7.

² "Sclerosis" = thickening/hardening of bone; "subchondral bone" = the bone found underneath articular cartilage; "osteophytes" = bony spurs formed during the development of osteoarthritis; "synovium" = the inner layer of the joint capsule; "hypertrophy" = increased growth.

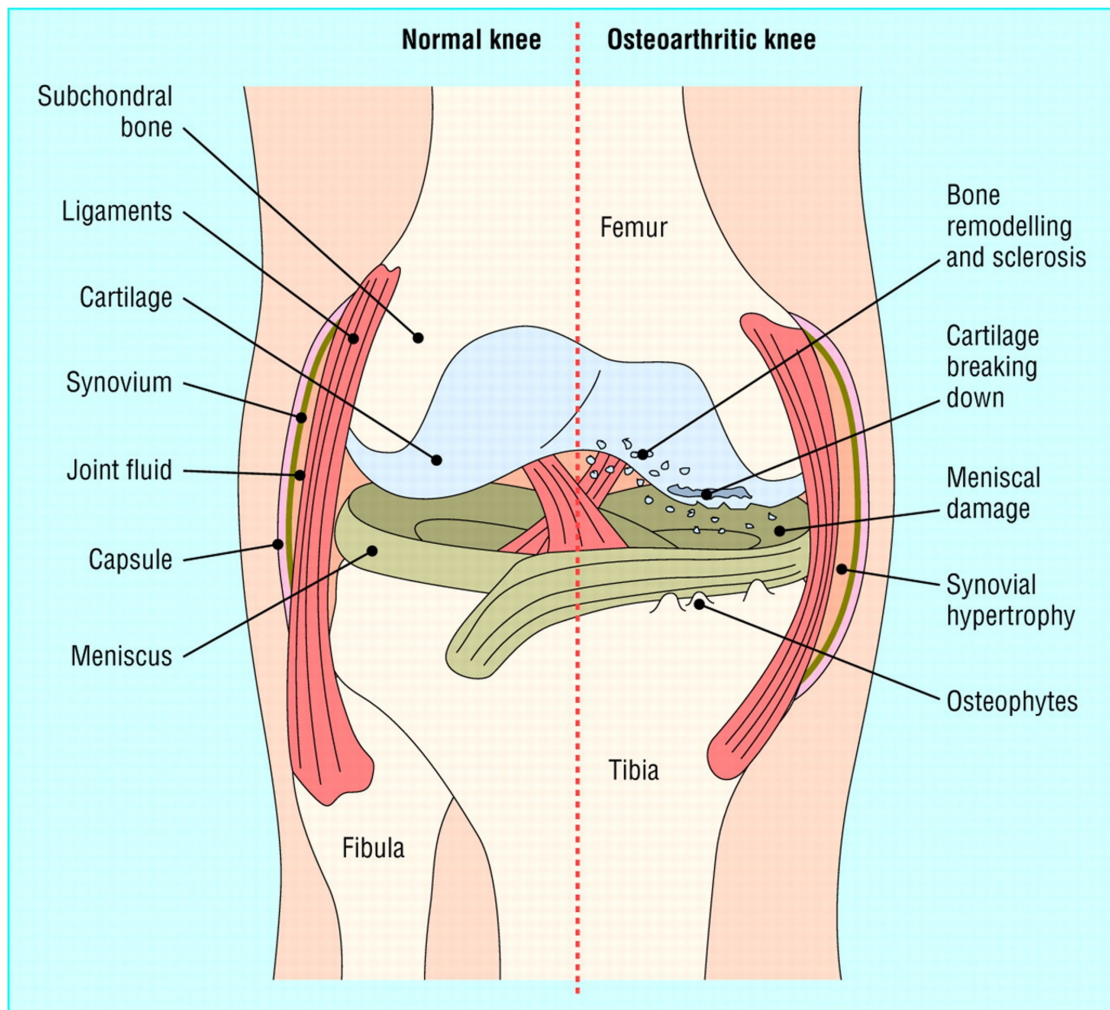


Figure 1.7. Osteoarthritis is a disease characterised by multiple structural changes. The anatomy of a normal human knee joint is shown on the left-hand side, with its osteoarthritic counterpart shown on the right-hand side. Structural changes in an osteoarthritic knee include the degradation of articular cartilage, damage to the medial meniscus, the formation of osteophytes, hypertrophy of the synovial membrane as well as sclerosis and remodelling of subchondral bone. *Reproduced from Osteoarthritis, David J. Hunter and David T. Felson, BMJ 332, 639-642, 2006 with permission from BMJ Publishing Group Ltd.*

The most common risk factors for OA are ageing, obesity, gender, genetic predisposition, previous injury to the joint and mechanical problems including poorly aligned or misshapen joints (69, 70). Epidemiologically, any joint in the body can be affected, although the most frequently affected are those in the knees, hips, hands, neck and lower back, as they bear the greatest mechanical load during daily activities (71, 72). Of these articulations, the knee joint accounts for 87% of the global burden of OA (66, 73). There is

strong evidence in animal models of guinea pigs and dogs that osteoarthritis begins in cartilage (74). The degradation of this tissue is thought to arise from abnormal mechanical stresses on the load-bearing joints mentioned above, suggesting that it is a mechanically induced disease (75). These stresses arise due to a multitude of factors, including joint injury and a reduced capacity of the joint to withstand mechanical loads due to age-related weakening of supporting muscles—such as the quadriceps muscles surrounding the knee joint (75). The structure and function of cartilage will be covered in the next section.

1.4.2. The structure and function of articular cartilage

Articular cartilage is a connective tissue that lines the surfaces of diarthrodial joints in the body (27). It is characterised by a complete absence of blood vessels, lymphatic vessels and neurons. From a biomechanical perspective, its primary role is to lubricate the surfaces of joints—with the assistance of synovial fluid—in order to facilitate movement during perambulation, as well as absorb shocks arising from external stresses. More specifically, the structure of articular cartilage is biphasic: it consists of a fluid phase and a solid phase. The former phase consists of water with ionic and non-ionic particles that make up 65–80% of the total weight; whilst the latter contains chondrocytes surrounded by a complex architecture of proteoglycans, collagens, glycoproteins and lipids known as the extracellular matrix (27, 28). Chondrocytes are the only metabolically active cells that function in the maintenance, growth and regeneration of the ECM (27, 76). Each chondrocyte is encompassed by an extracellular matrix sub-divided into different regions. These regions enable it to react to a mixture of stimuli such as growth factors, mechanical loads and piezoelectric forces, amongst others (27). The three regions of the ECM encompassing each chondrocyte are the pericellular region, the territorial region and the inter-territorial region (27, 28). The pericellular region directly envelops the plasma membrane of the chondrocyte

and contains non-collagenous proteins, glycoproteins and proteoglycans (27). It is actively involved in biochemical and mechanical cell signalling (28). By contrast, adjacent to this area, the territorial region comprises a thick meshwork of collagen fibrils that provide the cell with resistance to mechanical stress (27, 28). Finally, the inter-territorial region is the largest and furthest from the chondrocyte. It is characterised by an abundance of proteoglycans and collagen fibrils that differ in their physical arrangement depending on the depth of the cartilage. This region makes the largest contribution to the biomechanical features of cartilage (27, 28).

The most abundant structural components of the three distinct regions of the ECM are aggrecan and type II collagen (27, 28). Aggrecan was described earlier as a proteoglycan that contains a core protein with a M_w of ~250 kDa bound to unbranched, negatively charged GAGs of keratan sulphate and chondroitin sulphate (27, 28). Aggrecan exists as large aggregates where multiple monomers bind to a hyaluronic acid backbone through link proteins. In addition, the GAG chains of aggrecan contain carboxylate and sulphate groups which give it an overall negative charge, thereby allowing cartilage to remain hydrated and providing it with a resistance to compression (27, 28). Type II collagen accounts for ~90% of the collagen in the ECM and is comprised of fibrils, each of which consists of a triple helix of three identical $\alpha 1(\text{II})$ polypeptide chains with a repeating amino acid motif of Gly-X-Y; where X and Y are frequently proline and hydroxyproline (27, 28, 77). These fibrils bundle together with type IX and XI collagens to form heterofibrils interlinked with COMP and the proteoglycans decorin and fibromodulin (27, 28, 77). Type II collagen provides cartilage with its stability and tensile strength (27, 28).

1.4.2.1. The zonal anatomy of articular cartilage

Articular cartilage is also partitioned into five separate zones, which differ in the arrangement of both type II collagen and chondrocytes (Figure 1.8) (27, 28). The zone closest to the surface of articular cartilage and in contact with the synovial fluid is the most superficial zone, and can be separated into three layers, the first of which contains proteins, proteoglycans and lipids; the second an abundance of small fibrils; and the third type I–III collagen fibrils (28). The second zone is the tangential zone, which comprises a dense mixture of type II and type IX collagen fibrils intertwined with proteoglycans. The chondrocytes are elongated, flat and resemble fibroblasts. The type II collagen fibrils run parallel to the cartilage surface (27, 28). The third zone is the middle zone, which constitutes proteoglycans and collagen fibrils, the latter of which have a random organisation. The chondrocytes are spherical and are not numerous. This zone forms the first line of defence against compressive loads (27, 28).

The fourth zone is the deep zone, where type II collagen fibrils run perpendicular to the cell surface alongside chondrocytes aligned into columns. This area is also characterised by a low water content and high proteoglycan content, and provides the greatest resistance to compressive forces (27, 28). The final zone is the calcified zone, which fastens the cartilage to the subchondral bone through collagen fibrils extending from the deep zone. Here, there are a small number of hypertrophic chondrocytes embedded in a mineralised matrix of type X collagen (27, 28).

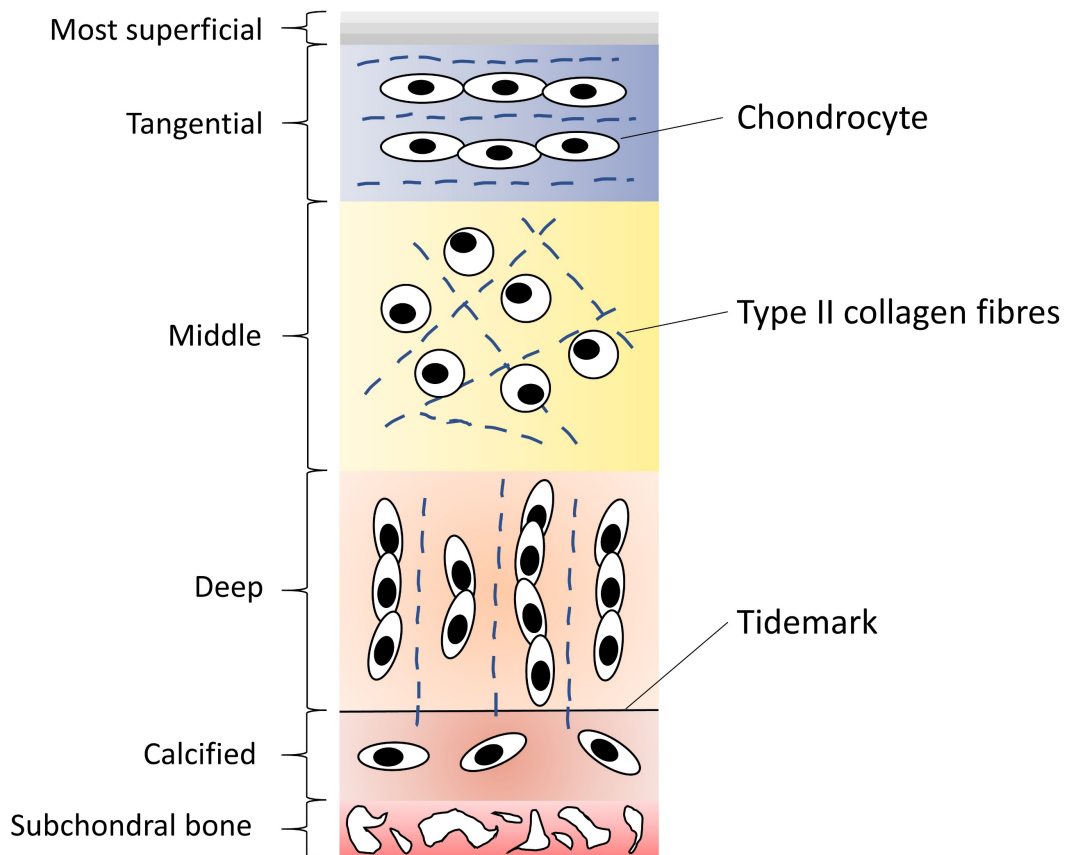


Figure 1.8. Zonal arrangement of articular cartilage. Schematic representation of the different zones found in articular cartilage, of which there are five: most superficial, tangential, middle, deep and calcified. The arrangement of chondrocytes is different within each zone: parallel to the cartilage surface in the tangential zone, randomised in the middle zone, perpendicular in the deep zone and hypertrophic in the calcified zone. Likewise, the type II collagen fibres display similar orientations to the chondrocytes within each zone, except in the middle zone where they form a meshwork. This figure was derived from Bottini *et al* (28).

1.4.3. Metalloproteinase activity in osteoarthritis

As discussed in section 1.4.2, the two major structural constituents of articular cartilage are aggrecan and type II collagen. Osteoarthritis is thought to proceed from the initial degradation of aggrecan followed by type II collagen (78-82), primarily by two classes of catabolic metalloproteinases: ADAMTSs (*i.e.* aggrecanases) and MMPs (*i.e.* collagenases) (3). The next sections will expand on the role of both of these metalloproteinases in the pathogenesis of OA, beginning with the MMPs.

The role of MMPs in osteoarthritis has been heavily researched in both *in vitro* and *in vivo* models of this disease *via* the detection of catabolic collagenase activity. For the sake of clarity, these models have been sub-divided into the most commonly studied species: bovine, human and rodent.

1.4.3.1. Catabolic MMPs: *in vitro* studies with bovine cartilage explants

To determine the time frame of MMP activity in osteoarthritis, research has been carried out using bovine cartilage explants *in vitro*, of which several investigations are highlighted here (78-80, 83). These studies were performed prior to those with mice, as explants are extracted from bovine nasal septa, which are easily obtained from an abattoir. This is in contrast to human explants, which are difficult to acquire due to the strict ethical approval required (84), and critically, very few control samples are available from healthy individuals.

In one study, Mitchell and co-workers treated bovine nasal cartilage with interleukin-1 α (IL-1 α) and monitored the release of hydroxyproline from the subsequent breakdown of type II collagen (83). Hydroxyproline release began after 12 days of tissue culture, with collagenase activity first detected around the same time (day 14) using [^{14}C]labelled type I collagen (83). Both collagenase activity and hydroxyproline release reached a maximum 18 days into tissue culture (83). Similarly, two separate studies showed that bovine cartilage explants stimulated with IL-1 released hydroxyproline 10 days and 12 days later, and this secretion peaked at 15 days and 20 days, which are comparable timeframes to those found by Mitchell *et al* (78, 80). Kozaci *et al* corroborated these findings by showing that a significant quantity of type II collagen degradation was detected by enzyme-linked immunosorbent assay (ELISA) only after incubation of bovine nasal cartilage explants with IL-1 α for 3–4 weeks (79). Elevated MMP activity was only observed in the third week of tissue

culture and was measured using the FRET substrate 7-methoxycoumarin-4-yl-acetyl (Mca)-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-NH₂ (79, 85). The MMPs responsible for type II collagen degradation were revealed as collagenases (*i.e.* MMP-1, -13) and gelatinase A (*i.e.* MMP-2) through western blotting and gelatin zymography, respectively (79). Overall, the experiments with bovine cartilage explants revealed the absence of a clear consensus on the specific MMPs involved in catabolic collagenase activity in *in vitro* models of OA.

1.4.3.2. Catabolic MMPs: early *in vitro* studies with human osteoarthritic cartilage and synovial fluid

In 1989, Dodge and Poole were the first to study the degradation of type II collagen by immunostaining for the resultant α -chain fragments with a polyclonal antibody (86). In their experiments, explants from bovine articular cartilage as well as those from patients with normal and osteoarthritic knee cartilage were analysed (86). In bovine cartilage explants, immunofluorescent staining was found to be most intense around the pericellular region after IL-1 treatment compared to those explants left untreated. In human osteoarthritic cartilage, increased staining for degradation fragments of type II collagen was observed in the pericellular and territorial regions compared to control normal human cartilage. In terms of zonal staining, osteoarthritic cartilage was typically stained to a greater extent in the superficial and middle zones (86).³

In contrast to this, Hollander and co-workers developed a monoclonal antibody specific to degraded type II collagen (87). Once again, the percentage denaturation of type II collagen was more pronounced in the superficial zone as opposed to the deep zone of osteoarthritic cartilage (87). When comparing human osteoarthritic knee cartilage to normal knee cartilage *via* immunoassay, a significantly higher amount of degraded type II collagen

³ Please refer to Figure 1.8 for orientation.

was observed as a percentage of the total type II collagen content (*cf.* 6.0% to 1.1%) (87). In both sample types, initial type II collagen damage was primarily found in the pericellular region surrounding the chondrocytes and in the superficial zone and upper part of the middle zone (81).³ As cartilage became more degenerated, this damage spread to the lower part of the middle zone and deep zone (81).

With a focus on identifying the specific metalloproteinases involved in human cartilage degradation, Wolfe *et al* used northern blotting to analyse the relative messenger RNA (mRNA) levels of MMP-1 and MMP-3 in *ex vivo* human osteoarthritic synovium and articular cartilage (88). Similar mean levels of MMP-1 and MMP-3 mRNA were found in the synovium, but a significantly higher mean level of mRNA was found in articular cartilage for MMP-3 compared to MMP-1. In addition, a comparison between the protein levels of MMP-3 and MMP-1 in synovial fluid by sandwich ELISA uncovered a significantly higher mean level of MMP-3 protein (88). In another study, Lohmander and co-workers also used a sandwich ELISA, but instead looked at the protein levels of MMP-3 and MMP-1 in the synovial fluid of patients with different forms of knee osteoarthritis or knee joint injury (89). The concentrations of both metalloproteinases were found to be higher in all patient disease groups compared to healthy controls (89). The concentration of MMP-3 exhibited the highest increase and differentiated between healthy cohorts and all diseased cohorts with a specificity of 90% and sensitivity of 84% (89). However, the majority of MMP-3 protein present in synovial fluid exists in the proform, thus rendering it proteolytically inactive (90).

1.4.3.3. Catabolic MMPs: later *in vitro* studies with human cartilage explants revealed MMP-13 as a major collagenase

To determine the MMPs responsible for type II collagen degradation in human cartilage explants, studies from the late 1990's and early 2000's by Mitchell *et al* (91), Billinghamurst *et al* (92) and Dahlberg *et al* (93) revealed MMP-13 as a strong candidate.

In the study by Mitchell and co-workers, both MMP-1 and MMP-13 mRNA were expressed in human osteoarthritic knee cartilage and stimulation with IL-1 α led to a considerable increase in the mRNA and protein levels of both metalloproteinases (91). In another, Billinghamurst *et al* developed antibodies to detect the different neo-epitopes resulting from degradation of type II collagen by the collagenases MMP-1, MMP-8, and MMP-13 (92). Treatment with a selective MMP-13 inhibitor resulted in a significant decrease in the release of one of these neo-epitopes during explant culture of human osteoarthritic cartilage, indicating that MMP-13 may play a key role in the degradation of type II collagen (92). Part of the study by Dahlberg and co-workers aimed to determine whether collagenase-mediated degradation of type II collagen was increased in samples of human osteoarthritic cartilage compared to non-arthritic cartilage, and whether this cleavage could be abrogated by a preferential inhibitor for both MMP-8 and MMP-13, but not MMP-1 (93). Digestion of type II collagen was assessed through neo-epitope release, and this was significantly increased in OA cartilage samples compared to control cartilage samples (93). The use of a potent MMP-13 inhibitor (which also partially inhibited MMP-8) was sufficient to reduce type II collagen degradation, indicating that MMP-13 and MMP-8 may play a more important role in this process compared to MMP-1 (93).

Taken together, the evidence from these studies suggests that MMP-13 is a strong candidate for type II collagen degradation in osteoarthritis. This knowledge, coupled with the

fact that MMP-1 is not expressed in adult mice (although orthologs are expressed) (94), led to the development of an MMP-13 knockout (KO) mouse model of OA to see whether cartilage would be protected from degradation *in vivo*.

1.4.3.4. The MMP-13 knockout mouse model of osteoarthritis partially protects cartilage from erosion

The MMP-13 knockout mouse model was first described by Little *et al* in 2009 (95). In this model, the severity of osteoarthritis induced by surgical destabilisation of the medial meniscus (DMM) was compared between knockout and wild-type mice. At 4 weeks post-induction of disease, wild-type and KO mice could not be differentiated in terms of structural cartilage erosion in the femur or the tibia when assessed by histology. However, a significant amount of structural cartilage damage was observed in the tibiae of wild-type mice compared to KO mice at 8 weeks following disease induction. This indicated that MMP-13 may be a major catabolic MMP in cartilage degradation (95). The evidence for potentially higher MMP-13 activity in wild-type osteoarthritic animals is strengthened by a study in a surgical model of OA in rats (96). Here, significant increases in levels of *Mmp13* gene expression were found 4 weeks after surgery in the articular cartilage of diseased rats compared to sham-operated controls. Higher levels of MMP-13 protein expression were also found in osteoarthritic cartilage compared to sham-operated or un-operated controls at this time point (96). Overall, these studies show that MMP-13 is either directly involved in cartilage degradation through type II collagenolysis or indirectly through increased activation of other MMPs (*e.g.* MMP-9) (95).

Catabolic ADAMTSs also play a major role in osteoarthritis through the degradation of aggrecan, the second major structural component of articular cartilage. The role of ADAMTS proteinases—specifically the aggrecanase group—has been most extensively

researched in relation to osteoarthritis in numerous *in vitro* and *in vivo* models since the beginning of the 1990s. The following studies will once again be divided by species into bovine, human and mouse for clarity.

1.4.3.5. Catabolic ADAMTSs: aggrecanases and the identification of cleavage sites in aggrecan

In the early 1990s, the term “aggrecanase” was used to describe the aggrecan-degrading activity of an unknown proteinase (97-102). The structure of human aggrecan is shown in Figure 1.9 along with the location of the major ADAMTS and MMP cleavage sites. The position of the cleavage sites along with the nature of the proteinases responsible will be covered in this section.

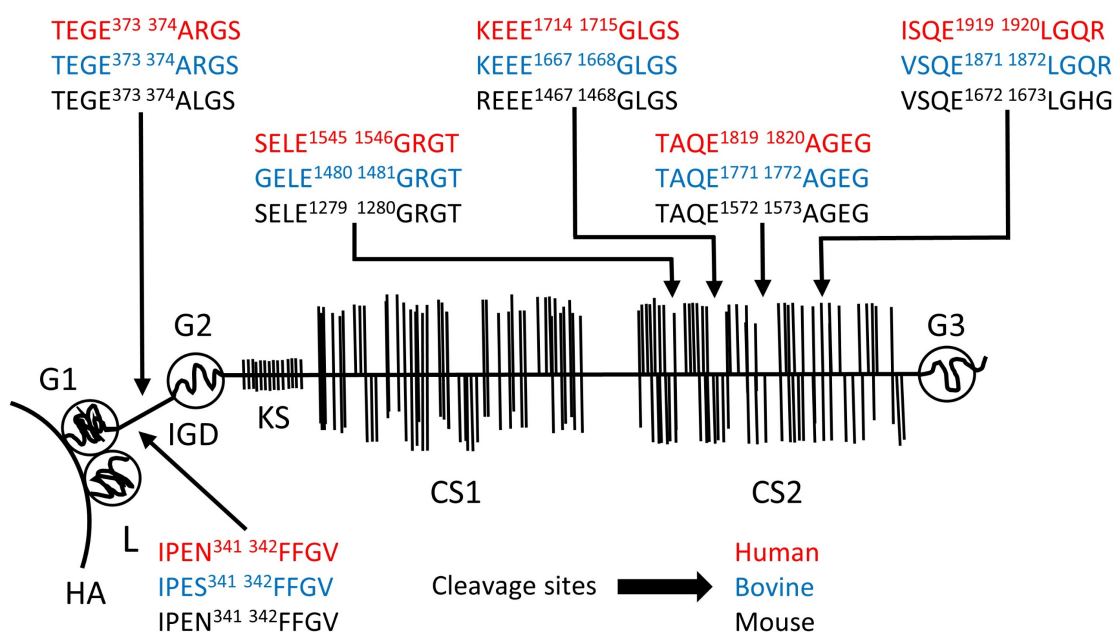


Figure 1.9. Aggrecan possesses ADAMTS and MMP cleavage sites with unique amino acid preferences. Representative structure derived from Wayne *et al* (63). ADAMTS cleavage sites are shown above the structure, with the colour of the text designating the location in different species: red for human, blue for bovine and black for murine. The major MMP cleavage site is shown below the structure with the same colour scheme. Please note that other MMP cleavage sites have been omitted as they are beyond the scope of this thesis. Key to annotations: HA = hyaluronic acid, L = link protein, G1-G3 = globular domains 1-3, IGD = interglobular domain, KS = keratan sulphate region, CS1/CS2 = chondroitin sulphate regions 1/2.

In 1991, Sandy *et al* identified and quantified 15 marker peptides resulting from the digestion of bovine aggrecan by trypsin (97). These peptides were also quantified in aggrecan extracted from untreated or IL-1 α treated bovine articular cartilage explants and their medium. Under both conditions, aggrecan secreted into the medium contained a lower amount of the globular G1 domain, which was more pronounced for IL-1 treated samples. Overall, this indicated that the degradation of aggrecan in articular cartilage comprised cleavage of the core protein primarily within the interglobular domain (IGD) between the G1 and G2 domains (97). Ten days later, a follow up publication by the same research group revealed the specific location of the IGD cleavage site between E³⁷³⁻³⁷⁴A, which resulted in detachment of the G1 domain from the rest of the core protein (98). In bovine chondrocytes stimulated with retinoic acid or IL-1 α , Lark *et al* also showed cleavage of aggrecan at E³⁷³⁻³⁷⁴A by “aggrecanase”, revealing that proteolysis at this site was key to aggrecan catabolism in this system (101). In addition, the role of “aggrecanase” in cartilage degradation was shown by catabolism of bovine nasal and articular cartilage explants treated with IL-1 β and/or tumour necrosis factor- α (TNF- α), where cleavage at E³⁷³⁻³⁷⁴A and increased staining for the ³⁷⁴ARGSVIL neo-epitope by western blotting correlated with greater GAG release (102).

Further cleavage sites within the aggrecan core protein were identified by Ilic and co-workers through analysis of eight N-terminal peptides released into the medium from bovine cartilage explants left untreated or treated with either retinoic acid or calf serum (99). In addition to TEGE~ARGS, the other identified sites were TAQE~AGEG and VSQE~LGQR. The same N-terminal peptide sequences were found in age-matched samples of bovine serum and synovial fluid, thereby suggesting the same cleavage sites (99). In another study, Loulakis *et al* cultured bovine cartilage explants with or without IL-1 α and utilised N-terminal sequencing to determine the amino acid composition of the degradation fragments secreted

into the medium (100). This confirmed the three cleavage sites described by Ilic *et al* along with the KEEE~GLGS site in the CS2 region of aggrecan (100).

The first proteinase member of the aggrecanase group was discovered by Kuno *et al* in 1997 and was designated ADAMTS-1, although its aggrecanase activity was not known at the time (103). It wasn't until 1999 that the first two aggrecanases were described three months apart, namely aggrecanase-1 (*i.e.* ADAMTS-4) (104) and aggrecanase-2 (*i.e.* ADAMTS-5) (105). With these discoveries, another cleavage site in bovine aggrecan was revealed by Tortorella and co-workers as GELE~GRGT (106). This established the total number of major cleavage sites in aggrecan as five: one in the IGD and four in the CS2 region (Figure 1.9).

1.4.3.6. ADAMTSs: *in vitro* studies with human osteoarthritic cartilage and synovial fluid

Prior to the discovery of ADAMTS-4 and ADAMTS-5, early work on aggrecanases focused on human synovial fluid samples from patients with knee injury and those with early or late-stage OA. They predominantly showed aggrecan fragments with an N-terminal ³⁷⁴ARGS resulting from cleavage at E³⁷³⁻³⁷⁴A (107, 108). A later study by Lark *et al* using antisera to stain for the TEGE and IPEN neo-epitopes showed that both ADAMTSs and MMPs were implicated in the catabolism of aggrecan in normal and osteoarthritic human articular cartilage (109). In normal samples, the amount of staining for both fragments rose with age and intense staining for both fragments was observed in osteoarthritic cartilage in areas with considerable degradation (109).

Following the discovery of ADAMTS-4 and ADAMTS-5 and the development of neo-epitope antibodies for the specific aggrecan cleavage sites (106, 110), more detailed analyses were performed on the synovial fluid and articular cartilage of patients with normal, injured and osteoarthritic joints (111). For normal human articular cartilage and synovial fluid, the

primary degradative pathway of aggrecan was due to ADAMTSs, while MMP-like activity (which could not be assigned but was similar to aggrecan digestion by MMP-3) served only to shorten aggrecan at the C-terminus (111). Cartilage and synovial fluid from injured joints were characterised by large amounts of the TEGE neo-epitope and indicated an acceleration of aggrecan breakdown by ADAMTSs (111). However, aggrecan originating from the cartilage of osteoarthritic patients was comparable to normal cartilage, whilst the synovial fluid revealed low molecular mass CS2 fragments that differed from the other sample groups (111). Another study compared synovial fluid and articular cartilage from patients with osteoarthritis (112). The degradation pattern of aggrecan in the synovial fluid was akin to that when aggrecan was cleaved by ADAMTS-4 *in vitro* (112). Nevertheless, in contrast to TEGE fragments, the fluid contained very little IPEN fragments, whereas osteoarthritic cartilage contained considerable quantities of both, thereby corroborating the results from Lark *et al* (109, 112).

1.4.3.7. ADAMTS-4 and ADAMTS-5: which is the major aggrecanase in osteoarthritis?

The major focus of *in vitro* studies has been ADAMTS-4 and ADAMTS-5, as these have been demonstrated to be the most active in the degradation of aggrecan. For example, in bovine articular cartilage stimulated with IL-1 α or TNF, both ADAMTS-4 and ADAMTS-5 were found to be responsible for aggrecan degradation in the IGD at E³⁷³⁻³⁷⁴A (113). For IL-1 treated cartilage, the order of proteolysis at the ADAMTS-4 and ADAMTS-5 cleavage locations correlated with that when purified aggrecan was incubated with either of these aggrecanases (113). In addition, media from IL-1 treated cartilage explants with ADAMTS-4 and ADAMTS-5 (removed *via* a sepharose column containing bound anti-ADAMTS-4/-5 antibody), showed an ~90% decrease in aggrecanolytic activity compared to controls when incubated with aggrecan and tested for cleavage at E³⁷³⁻³⁷⁴A (113). A considerable amount

of work has therefore been carried out to determine whether the degradation of aggrecan in osteoarthritis is primarily due to ADAMTS-4 or ADAMTS-5. This research has focused on both gene expression and knockout mouse models of this disease as well as *in vitro* studies on human chondrocytes, osteoarthritic cartilage and cytokine-stimulated normal cartilage.

Studies on ADAMTS gene expression have been carried out in mouse and rat models of OA. For example, 6 hours following surgical induction of OA in mice, a two-fold increase in mRNA levels of *Adamts5* was observed in the articular cartilage of diseased mice compared to sham-operated control mice (114). Following surgical joint immobilisation, levels of *Adamts5* in OA mice were significantly reduced compared to those of sham-operated control mice, thereby revealing the mechanosensitive nature of *Adamts5* (114). Similarly, in a surgical model of OA in rats, a significant increase in the amount of *Adamts5* gene expression was observed in the articular cartilage of osteoarthritic rats compared to sham-operated controls 4 weeks after surgery (96).

In a surgical DMM model of murine osteoarthritis (115), knockout of the gene that encodes most of the Cat domain of ADAMTS-4 did not reduce the severity of osteoarthritis when assessed by histological scoring of knee joints up to 8 weeks post-induction of disease (116). In addition, the degree of aggrecan breakdown was similar to that found in wild-type mice when immunostaining for the TEGE neo-epitope (116). This indicated that ADAMTS-4 was not the major aggrecanase responsible for aggrecanolysis, as knockout mice were not protected from developing osteoarthritis in this model. In stark contrast to this, deletion of the gene encoding most of the Cat domain of ADAMTS-5 resulted in a significant reduction in the severity of osteoarthritis compared to wild-type mice when joints were scored by histology (117). Furthermore, the TEGE neo-epitope was not detected in murine articular cartilage treated with IL-1 and retinoic acid up to 8 weeks after establishment of disease

(117). These results proved that ADAMTS-5 was the major aggrecanase in murine osteoarthritis and that mice were protected from cartilage and aggrecan degradation (117). This is despite proteolysis of aggrecan being demonstrated at the SELE~GRGT and REEE~GLGS sites in the CS2 region due to either ADAMTS-4 or another aggrecanase (118). The finding that ADAMTS-5 was the major aggrecanase in mice was further substantiated when weekly treatment of DMM mice with a monoclonal antibody specific for ADAMTS-5 significantly decreased cartilage degradation compared to the antibody control when joints were scored by histology (119).

Contrastingly, in human osteoarthritis the general picture is uncertain. This is because of the fact that only *in vitro* models of disease can be employed. In addition, studies with human osteoarthritic cartilage are typically from patients undergoing joint replacement at the end stages of disease, a time point where aggrecanase activity is unlikely to be as relevant. Regardless, the studies do give some indication of which aggrecanase could be playing a more important role. For example, Song *et al* showed that small interfering RNA (siRNA) knockdown of either *Adamts4* or *Adamts5* in normal cartilage explants treated with TNF- α /oncostatin M—or untreated osteoarthritic cartilage explants—reduced immunoblotting for AGE3 and ARGS neo-epitopes. This suggests that both aggrecanases played a role in aggrecan degradation in this system (120). In a similar vein, Ismail and co-workers developed an aggrecan overlay assay, where human chondrocytes were covered with bovine aggrecan and either left unstimulated or were stimulated with IL-1 (121). The presence of aggrecan fragments in the medium was assessed by western blotting for the same neo-epitopes as Song *et al* (121). Here, siRNA knockdown of *Adamts5* in IL-1 treated wells resulted in no staining for the AGE3 neo-epitope and reduced staining for the ARGS neo-epitope when compared to knockdown of *Adamts4*. In the case of unstimulated wells, a

small amount of blotting was observed for ARGS and AGE3 neo-epitopes with *Adamts4* knockdown, yet none for *Adamts5* knockdown. This suggests that human ADAMTS-5 appeared to be more important in bovine aggrecan degradation than ADAMTS-4 under these conditions (121). Larkin and co-workers developed two ADAMTS-4 and ADAMTS-5-selective monoclonal antibodies and found that treating osteoarthritic cartilage with the ADAMTS-5 antibody significantly reduced the production of the ARGS neo-epitope compared to treatment with the ADAMTS-4 antibody (119). Overall, this data bolstered support for ADAMTS-5 as the major aggrecanase in human osteoarthritis and a more important therapeutic target than ADAMTS-4 (119). Taken together, the literature demonstrates that although the major aggrecanase in human osteoarthritis is not clear-cut, the weight of evidence is arguably more in favour of ADAMTS-5 than ADAMTS-4.

1.4.3.8. Targeting metalloproteinases as a means to image catabolic proteinase activity in osteoarthritis

The numerous studies on ADAMTSs and MMPs described in the previous sections have highlighted the key role that both of these proteinases play in the degradation of articular cartilage in osteoarthritis. They are thus potential biomarkers for disease diagnosis, prognosis and the evaluation of therapeutic responses. From a clinical standpoint, there are currently no biomarkers for the diagnosis of osteoarthritis, despite a considerable amount of research to date. Numerous targets have been investigated, ranging from cytokines such as IL-6 to degradation fragments of type II collagen (*e.g.* the urinary C-terminal telopeptide of type II collagen) to building blocks for collagen formation (*e.g.* the N-terminal pro-peptide of collagen IIA) (122, 123). Instead, OA is diagnosed primarily on the basis of symptoms such as pain, which is a complicated and dynamic phenomenon that is poorly understood (5, 71). More specifically, the National Institute for Health Care and Excellence (NICE) provides three criteria for the clinical diagnosis of osteoarthritis in the UK: age at or greater than 45, joint

pain due to activity, and the presence of joint stiffness in the morning, lasting up to a maximum of 30 minutes (71, 124). The use of an imaging modality such as X-ray radiography is not generally considered necessary, although it can be used to evaluate structural changes through the analysis of joint space narrowing (JSN) between the constituent bones (71, 124, 125). Unfortunately, radiographic JSN is a poor method for detecting early disease and monitoring disease progression. This is because cartilage is radiolucent and thus cannot be imaged by X-rays, making analysis of JSN both an indirect and insensitive measure of cartilage loss (125). In addition, the progression of disease is assessed by scoring methods such as the Kellgren-Lawrence scale (126), which are subjective and inadequate. Collectively, this makes X-ray radiography only useful towards the end stages of disease, where a complete loss of joint space width and exposed bone are signs that total joint replacement is necessary (125, 127). There is thus a real unmet clinical need for the development of a diagnostic biomarker for osteoarthritis to detect early osteoarthritic processes and track disease progression.

At the beginning of this thesis, the broad aim was to develop a tool that would monitor the activity of ADAMTSs and MMPs in osteoarthritis, with potential application to human disease. Later on, FRET substrates were chosen as the ideal probe for carrying out this task, due to their excellent detection sensitivity and ability to amplify signal intensity through continuous substrate cleavage. The existence of commercial *in vivo* MMP FRET substrates and the absence of such probes for ADAMTSs indicated that the latter proteinases could be targeted for FRET substrate development. In addition, it is currently thought that the degradation of aggrecan precedes that of type II collagen in the progression of osteoarthritis (78-82). This has been shown by *in vitro* tissue culture experiments with bovine cartilage explants stimulated with IL-1: aggrecan breakdown occurs in the first week of culture followed by the loss of type II collagen (78-80). Furthermore, staining for aggrecan is

lost prior to type II collagen staining in both murine and human osteoarthritic cartilage explants (81, 82). A representative example of the loss in aggrecan staining is shown for human tissues in Figure 1.10.

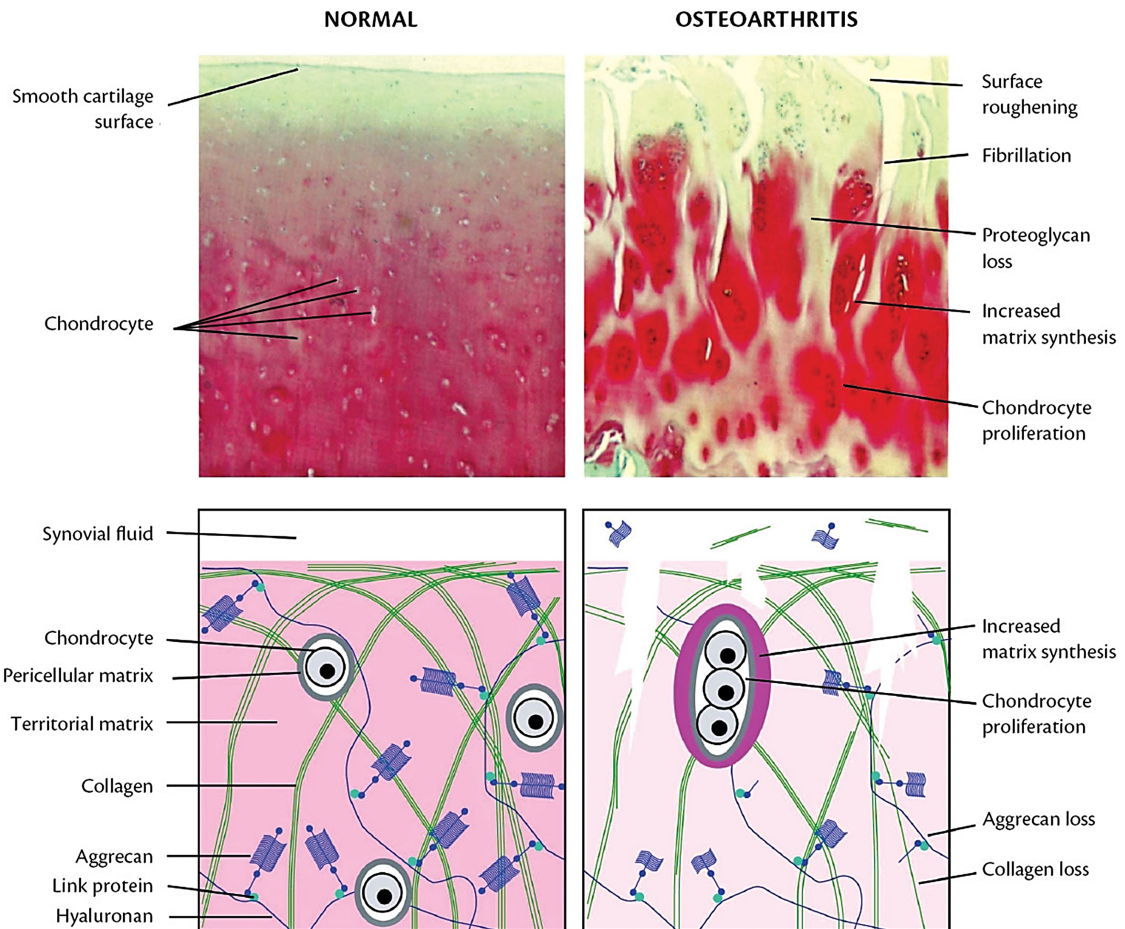


Figure 1.10. A comparison between normal and osteoarthritic articular cartilage. Histological slices (top) and corresponding schematic figures (bottom) are shown for normal human and osteoarthritic cartilage. In each histological section, the proteoglycans are stained with Safranin O dye. In normal cartilage, uniform staining for aggrecan can be observed, whereas in osteoarthritic cartilage, a loss of staining can be seen. The schematic figures illustrate the chondrocytes and key matrix components within each tissue type. In osteoarthritic cartilage, an increase in matrix synthesis is observed for certain chondrocytes and is indicated by greater safranin O staining. From *“Pathogenesis of osteoarthritis”* by Tonia L. Vincent and Linda Troeberg. Reproduced with permission of the Licensor through PLSclear. © Oxford University Press.

Therefore, targeting catabolic ADAMTS activity as opposed to catabolic MMP activity could help resolve the clinical issues associated with early osteoarthritis diagnosis and

monitoring of disease progression. Nevertheless, the initial FRET substrates for imaging proteinase activity in OA targeted MMPs, and these probes will be covered in the next section. In this section, the drawbacks of targeting MMPs will provide further evidence for why developing a FRET substrate to target an ADAMTS proteinase is the primary aim of this thesis.

1.4.3.9. Commercial MMP FRET substrates lack selectivity

A number of commercial FRET substrates have been used to monitor the combined activity of collagenases (*e.g.* MMP-13), gelatinases (*e.g.* MMP-2, -9) and stromelysins (*e.g.* MMP-3) non-selectively in mouse models of osteoarthritis *in vivo* (see Figure 1.6 for structures). In addition, FRET substrates targeting cathepsins will also be covered, as they were often used in conjunction with the MMP probes. This is probably because aggrecan and type II collagen are also degraded by cathepsins B/L and K, respectively (128, 129). Furthermore, knockout of cathepsin K significantly delayed the development of OA in a surgical mouse model of this disease (130), and cathepsin B activity was found to correlate to areas of degenerated human osteoarthritic cartilage (131).

In a DMM mouse model of osteoarthritis (115), the commercial FRET substrates MMPsense 680 and ProSense 680 were used to non-specifically monitor general MMP activity (*e.g.* MMP-2, -3, -9 and -13) and the cleavage of cathepsins B, L and S, respectively (132). Only MMP activity was significantly increased in DMM compared to sham knees 8 weeks into disease, a time point which represents moderate to severe OA by histology (132). This result was confirmed by Leahy *et al* with the same probe and disease model at 7 weeks (133). A similar result was reported with the MMPsense 750 probe (which can monitor the activity of MMP-2, -3, -7, -9, -12 and -13) in a murine mechanical compression model of osteoarthritis for injured knees vs contralateral knees 8 weeks after disease induction (134).

In another study involving three FRET substrates, proteinase activity from cathepsins B, L and S (ProSense 680), cathepsin K (Cat K 680) and MMPs (MMPSense 680) was also evaluated in a murine mechanical compression model of osteoarthritis (135). The activity of all these proteinases was significantly elevated in injured knees compared to contralateral knees throughout the 8-week study, but was highest between 1–7 days after injury (135). Unfortunately, this study lacked histological validation of fluorescent signal intensity, so proteinase activity could not be correlated directly with tissue damage from osteoarthritis (135).

Unfortunately, the FRET substrates used in these studies lack the selectivity necessary to identify the exact proteinase(s) linked to the onset and progression of osteoarthritis. This is particularly problematic for the design of therapeutic inhibitors, which (ideally) require a particular proteinase to target. Selecting a specific proteinase activity to monitor would thus be more helpful. Several of the *in vitro* studies mentioned previously, as well as the MMP-13 knockout mouse model and a rat OA model have pointed towards MMP-13 as a potential target. FRET substrates selectively targeting MMP-13 have thus been developed for monitoring the progression of osteoarthritis in rodent models *in vivo*.

1.4.3.10. MMP-13 FRET substrates can diagnose osteoarthritis in rodents but cannot detect and monitor early osteoarthritic processes

Lee and co-workers developed an MMP-13 FRET substrate through computational modelling and tested this in a surgical DMM model of osteoarthritis in rats (136). An ~3-fold and ~7-fold higher signal intensity was observed in diseased knees compared to contralateral knees⁴ 6-weeks and 8-weeks after disease induction, respectively (136). This signal intensity was significantly reduced in the presence of an MMP-13 inhibitor (136). However, the

⁴ The knee opposite the one that has been operated on, effectively an internal control.

selectivity of this probe was only evaluated against MMP-7 and not against other MMPs present in articular cartilage (*e.g.* MMP-2 and MMP-3). Subsequently, the same group developed a more selective substrate with an ~6-fold and ~3-fold preference for MMP-13 over MMP-9 and MMP-2, respectively (137). This showed an ~4-fold higher signal in osteoarthritic rat knees compared to contralateral knees 8 weeks into disease, with signal intensity once again diminished in the presence of an MMP-13 inhibitor (137). Utilising a library screening approach, Lim and Meinjohanns *et al* described an even more selective FRET substrate with an ~13-fold, ~16-fold, ~20-fold and ~2-fold preference for MMP-13 over MMP-1, MMP-2, MMP-3 and MMP-12, respectively (138, 139). This substrate was deployed in a DMM model of osteoarthritis in mice (139). The MMP-13 probe was only able to distinguish between DMM knees and sham/contralateral knees 8 weeks after disease induction—once significant cartilage degradation had occurred—making it unsuitable for imaging early disease processes in osteoarthritis (139). More recently, a highly selective MMP-13 probe consisting of a poly-L-glutamic acid-FRET substrate conjugate incorporating un-natural amino acids was described (140). With respect to cleavage by MMP-13, this showed very little cleavage by MMP-1, -3, -7, -8 (< 1 %) and a small amount of cleavage by MMP-2, -9, -12 and -14 (~5–17 %). In a DMM mouse model of osteoarthritis, this substrate was able to distinguish between diseased knee joints and non-diseased knee joints at an earlier time point of 6 weeks after DMM surgery (140). Further, the application of an MMP-13 inhibitor after 2 weeks was found to offer significant cartilage protection to DMM mice compared to untreated controls (140).

The research performed with FRET substrates specifically targeting MMP-13 has revealed that differentiating between osteoarthritic joints and healthy joints appears difficult in mice or rats prior to 6 weeks after the onset of disease. Given the evidence for the

degradation of aggrecan occurring prior to that of type II collagen in the pathogenesis of OA, developing a FRET substrate for monitoring ADAMTS activity should enable diagnostic imaging of early osteoarthritic processes as well as provide more effective tracking of disease progression. The current weight of evidence points towards ADAMTS-5 as the major aggrecanase in osteoarthritis (117, 119-121). The primary aim of this thesis was thus to develop a selective substrate for monitoring ADAMTS-5 activity in osteoarthritis and potentially other disease settings.

1.5. Combinatorial chemistry

1.5.1. Introduction

Combinatorial chemistry refers to the synthesis of a large number of different compounds (collectively known as a “library”) in a more rapid and facile manner than if each compound were synthesised in a linear fashion (141). Historically, one of the bottle necks in the discovery of novel ligands in drug discovery has lain in the problem of synthesising large libraries of compounds for biological screening (142). For example, 9-fluorenylmethoxycarbonyl-based solid-phase peptide synthesis (Fmoc-SPPS) (143) of all 3.2 million pentapeptides from 20 proteinogenic amino acids would take thousands of years to accomplish.⁵ However, a potential solution to this problem was first proposed in 1982 by the Hungarian chemist Furka (144). Furka reasoned that acquiring large numbers of compounds could be achieved rapidly through a “portioning-mixing” method, which later became known as “split-mix” synthesis (145). Figure 1.11 illustrates this method for the synthesis of a 16-mer combinatorial library with three unique amino acids coupled in each step.

⁵ Assuming the successful synthesis of one resin-bound peptide a day, this would take ~8767 years of uninterrupted work.

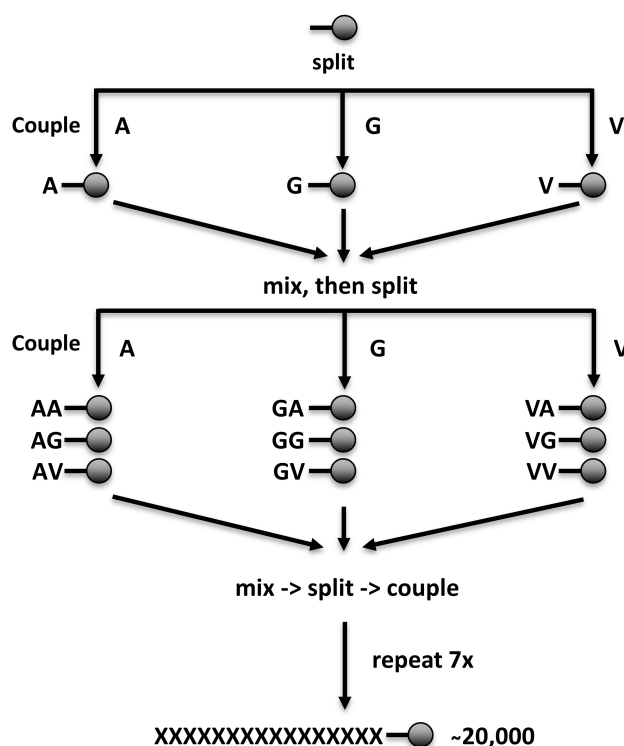


Figure 1.11. Split-mix synthesis facilitates the synthesis of large compound libraries. Representative schematic showing the synthesis of a 16-mer combinatorial library using split-mix synthesis with three amino acid couplings per step. Resin beads are initially split into three equimolar portions, with A, G and V amino acids coupled to each respective portion. Upon completion of coupling, all beads are mixed together, split again, and coupling performed as before. These steps are repeated until the entire library has been synthesised, giving a total of 19,683 unique peptide sequences (3^9).

Initially, resin beads are separated into equimolar portions corresponding to the number of different amino acids that will be coupled during that particular step in the synthesis. In the case shown in Figure 1.11, this is three portions. Coupling of one of the amino acids (*i.e.* A, G or V) to each portion is followed by mixing of all the resin beads and separation into three portions again. Coupling of the same amino acids as before results in the formation of nine unique dipeptides in a total of two reaction steps: AA, GA, VA, AG, GG, VG, AV, GV and VV.

Repetition of the split-couple-mix steps until the library synthesis is complete results in an exponential increase in the total number of peptide sequences in the minimum number of steps. The synthesis of the aforementioned pentapeptide library of millions of beads can

thus take several days as opposed to thousands of years. The application of this method was first described in 1991 for the synthesis of mixtures of tetrapeptides and pentapeptides (145). However, the idea that each bead contained a unique peptide sequence and could be screened against different biological targets to identify novel ligands led to the “one bead one compound” (OBOC) concept (146-148). To illustrate this concept, a pentapeptide library containing ~2 million beads was synthesised using split-mix synthesis and screened against a proteinase-tagged anti- β -endorphin antibody to identify novel ligands through a colorimetric detection technique (146). Following microsequencing of beads, this enabled the identification of a ligand with a binding affinity on par with the native ligand for anti- β -endorphin antibody and > 50-fold better than one identified for the very same antibody through phage display (146). Other than peptide libraries, many different types of OBOC libraries have been developed, including non-peptide oligomeric (*e.g.* peptoids and oligocarbamates) and small molecule (*e.g.* acyclic and heterocyclic) (148, 149). The biggest advantage of the combinatorial method over other library techniques such as phage display is that it enables the use of un-natural amino acids, thereby allowing coverage to extend beyond natural chemical space and increase molecular diversity. More specifically for this thesis, the ability to use un-natural amino acids for the purposes of developing proteinase substrates allows each peptide substrate to explore pockets within the active site of the proteinase more deeply through additional electrostatic interactions of modified side-chain functional groups. This could theoretically enable a higher level of substrate selectivity to be achieved for a particular proteinase. As a result of these advantages, the combinatorial library was chosen as the preferred technique to screen for a novel ADAMTS-5 FRET peptide substrate.

1.5.2. Screening substrate libraries requires specialised resin beads

In order to identify novel substrates for ADAMTS-5 through combinatorial library screening, it is important that the proteinase can diffuse into the polymer network of the bead and cleave the substrate. Standard polystyrene resins are not sufficiently high swelling to enable proteinases to diffuse into them (150). This problem was resolved in the early 1990s by Morten Meldal, who developed a new type of poly(ethylene glycol)-polydimethyl acrylamide (PEGA) resin bead. This had high swelling properties and was suitable for use in peptide synthesis and in the aqueous environment of proteinase screening (142, 151-155). Thus, a series of PEGA resins of different porosity were designed that enabled proteinases with different molecular weights to enter, *e.g.* PEGA₁₉₀₀ resins allowed those up to 60 kDa to enter, whereas PEGA₄₀₀₀ allowed those up to 120 kDa to enter (156). The structure of both of these resins is shown in Figure 1.12.

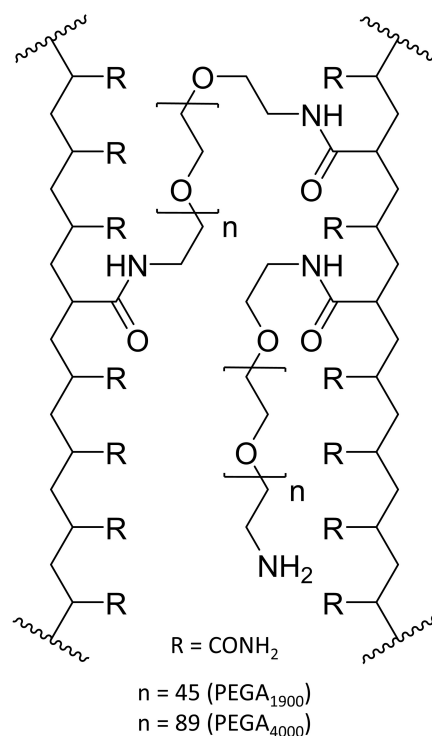


Figure 1.12. The structure of PEGA resins is tailored to proteinase screening assays. PEGA resins are compatible with Fmoc-SPPS and can be used under aqueous conditions for screening against different proteinases. In addition, they can be swelled in a variety of solvents, including water, dimethylformamide, dichloromethane, tetrahydrofuran, toluene and acetonitrile. This figure was derived from Renil *et al* (155).

However, once the proteinase had entered the bead, it was important to determine whether cleavage of the substrate had occurred. Therefore, each substrate on the bead was synthesised as a FRET substrate utilising the ortho-aminobenzoic acid (Abz) and 3-nitro-L-tyrosine (Y(NO₂)) fluorophore/quencher pair (157). These were chosen as they are inexpensive and readily available, thereby allowing large libraries to be synthesised at low cost. In addition, they are both small innocuous units that can be introduced into the peptide sequence without fear of disrupting substrate cleavage. Screening resin-bound FRET substrates has been successfully utilised to map the subsite specificity of the active sites of numerous proteases, such as papain (158), α -amylase (159), subtilisin A (160), cruzipain (161), EP-A and EP-B (162), protein disulphide isomerase (163), a specific cysteine protease

CPB2.8 (164, 165), caspases 1, 3, 6, 7 and 8 (166) and glutamyl endopeptidases from bacteria including *Staphylococcus aureus*, *Bacillus licheniformis* and *Streptomyces griseus* (167).

Once hit beads have been isolated from the library manually or through the use of an automated bead sorter, there are two methods for identifying peptide sequences on resin beads: decoding of optically encoded beads or matrix-assisted laser desorption ionisation (MALDI) mass spectrometry coupled with tandem mass spectrometry. It is important to understand the advantages and disadvantages of these methods before the most appropriate technique is chosen for sequencing the peptide hits.

1.5.3. Identification of peptide hits through optical encoding

Optical encoding is a technique that allows resin beads to be chemically or physically modified so that an optical “barcode” is generated. If a peptide sequence is bound to the bead it can be identified by “reading” of this barcode. Numerous methods have been used to encode beads, of which a few will be described here to highlight the variety of approaches taken (168-171).

Battersby and co-workers employed “colloidal barcoding” (168). Prior to coupling of an amino acid during combinatorial library synthesis, one or more combinations of different coloured fluorescent colloidal particles (so called “reporters”) were physically attached to resin beads. Following completion of library synthesis, each bead would contain a series of reporters, which would enable the peptide sequence to be identified by analysing the reporters under a fluorescent microscope with a variety of different filters (168).

In contrast, Fenniri *et al* chemically modified polystyrene resin beads through copolymerisation with a variety of different styrene comonomers and used the unique spectroscopic infrared (IR) and Raman barcodes to encode the beads (169). This technique

was developed to assist dual recursive deconvolution of combinatorial libraries in identifying peptide sequences (169).

Literal barcoding of beads (like those found on everyday commercial products) was described by Braeckmans and co-workers (170). These were generated by photobleaching fluorescently dyed microspheres in specific locations on resin beads. A key component of the system was to ensure that fluorescent particles within the spheres were fixed in place to prevent diffusive mixing with the photobleached particles (170). However, it is uncertain how this technique could be used for identifying peptides during combinatorial library synthesis.

On the other hand, an innovative and practical method for encoding beads for library synthesis and screening was reported by Meldal and Christensen (171). This technique, termed microparticle matrix (MPM) encoding, involved the generation of PEGA beads containing fluorescent tris(dialkylamino)-trioxatriangulenium ion microparticles (171). Each PEGA contained an average of eight randomly oriented microparticles (171). The distances and angles between the particles could be viewed as a rotation and translation independent 3D matrix, uniquely “barcoding” each individual bead (171). Therefore, during library synthesis, each bead was imaged using three orthogonal charge-coupled device (CCD) cameras on a rotating disk prior to coupling of an amino acid to generate the 3D coordinates of the matrix barcode (171). Once library synthesis had been completed, the library beads were incubated with fluorescently labelled ligands (171). The barcode of any fluorescent hit bead was imaged, and the peptide sequence on the bead was identified using a computational algorithm that matched the image of the barcode to images of every barcoded bead acquired during each coupling step of the library synthesis (171). This technique was successfully utilised to generate small molecule ligands for affinity purification of human growth hormone (171, 172).

1.5.4. Identification of peptide hits through MALDI and tandem mass spectrometry

Peptide sequences can also be identified on hit beads through a combination of MALDI and tandem mass spectrometry. Following cleavage of the linker that connects the peptide to the solid support (*i.e.* the resin bead), a MALDI spectrum is obtained to identify the mass of the peptide (M) in a singly protonated state; known as the quasi-molecular ion, $[M + H]^+$. Fragmentation of this ion results in the formation of a series of different fragment ions. Depending on which bond in the peptide is broken, a total of six different fragment ions may be formed (Figure 1.13). By piecing together the different fragment ions formed, the final peptide sequence can be elucidated. However, it is important to note that the greater the number of possible sequences, the harder it becomes to elucidate the exact peptide sequences from the mass spectra. This is due to the increasing number of permutations for amino acids within each position of the peptide. In practise therefore, combinatorial libraries are typically restricted to those with < 50,000 unique peptide sequences.⁶ The MALDI method was chosen preferentially over the MPM encoding method to identify the hit peptide sequences described in Chapter 5. The reason for this will be explained in the discussion Chapter 6.

⁶ Personal communication with Professor Morten Meldal.

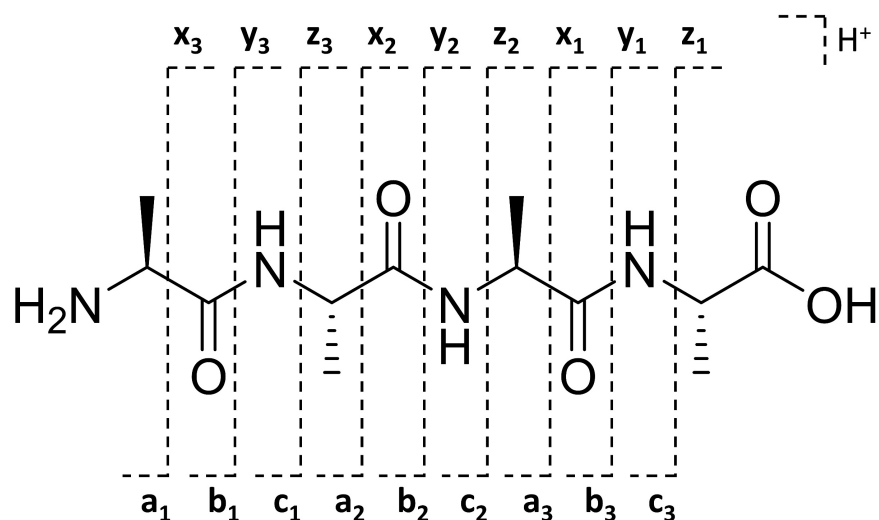


Figure 1.13. Peptides show unique fragmentation patterns during tandem mass spectrometry. The singly charged peptide sequence can fragment in multiple ways, depending on whether the carbon-carbon, carbon-nitrogen or nitrogen-carbon bond is cleaved. Any cleavage results in the formation of two fragment ions, which are assigned one of the following letters: a, b, c, x, y or z. The numbering for each letter refers to the number of side chains to the right or left of the cleaved bond and can be used to assign fragment ion peaks in mass spectra.

1.6. Hypothesis

The central hypothesis of this thesis is that the development a FRET substrate to monitor ADAMTS-5 activity should enable earlier diagnostic imaging of osteoarthritic processes as the cleavage of aggrecan by aggrecanases is thought to arise before the degradation of type II collagen by collagenases.

1.7. Aims

To summarise, the primary aim of this project is to obtain a selective ADAMTS-5 FRET peptide substrate for monitoring proteinase activity in different *in vitro*, *ex vivo* and potentially *in vivo* OA settings. Furthermore, it could also be used in other disorders typified by heightened ADAMTS-5 activity. Towards this goal, the following objectives will be met and covered in the results Chapters 3–5:

- (a)** The purification and characterisation of ADAMTS-4 and ADAMTS-5 proteinases (Chapter 3).
- (b)** The design, synthesis and screening of a combinatorial library of FRET peptide substrates (Chapter 4).
- (c)** The sequencing of hit substrates from combinatorial library screening and the determination of their selectivity and sensitivity for ADAMTS-5 (Chapter 5).

Chapter 2. Materials and methods

2.1. Molecular biology and biochemistry

2.1.1. Materials

For each section, materials were obtained from the sources listed.

2.1.1.1. Tissue culture

Falcon cell culture dishes and 0.22 µm bottle-top vacuum filters from Corning (Thermo Fisher Scientific, Loughborough, UK); hygromycin B (50 mg/ml) from Calbiochem (Merck, Feltham, UK); penicillin-streptomycin (PS) and Dulbecco's modified eagle's medium (DMEM) with glucose (4.5 g/L) and L-glutamine (0.584 g/L) from Lonza (Slough, UK); fetal bovine serum (FBS) from Gibco (Thermo Fisher Scientific, Loughborough, UK); lactalbumin enzymatic hydrolysate (LEH), Virkon and dimethyl sulphoxide (DMSO) from Sigma (Merck, Gillingham, UK); sodium azide (NaN₃) from VWR International (Lutterworth, UK); and cryovials from Thermo Fisher Scientific (Loughborough, UK). The human embryonic kidney (HEK)293 cells transfected with TS4-4, TS5-5 and proMMP-2ΔC domain deletion mutants were provided by Dr. Ngee Han Lim and Dr. Linda Troeberg and were derived from the work of Kashiwagi *et al*, Gendron *et al* and Troeberg *et al*, respectively (16, 17, 62).

2.1.1.2. Protein purification

Anti-flag M2 affinity gel, calcium chloride (CaCl₂), DMSO, tris base and sodium chloride (NaCl) from Sigma (Merck, Gillingham, UK); Gelatin-Sepharose 4B affinity resin, Sephacryl S-200 high resolution resin, Vivaspin 20 and Vivaspin 500 protein spin concentrators (molecular weight cut-off = 3 kDa) from Cytiva (Little Chalfont, UK); glycine and NaN₃ from VWR International (Lutterworth, UK) and 0.22 µm bottle-top vacuum filters from Corning (Thermo Fisher Scientific, Loughborough, UK).

2.1.1.3. SDS-PAGE

Sodium dodecyl sulphate (SDS) and β -mercaptoethanol from MP Biomedicals (Thermo Fisher Scientific, Loughborough, UK); 30% (w/v) acrylamide bis-acrylamide solution from Severn Biotech (Kidderminster, UK); bromophenol blue, tris base, ammonium persulphate and *N,N,N',N'*-tetramethylethane-1,2-diamine (TEMED) from Sigma (Merck, Gillingham, UK); and glycerol from VWR International (Lutterworth, UK).

2.1.1.4. Coomassie staining

Coomassie brilliant blue (CBB) R-250 from Sigma (Merck, Gillingham, UK), methanol and formic acid from Thermo Fisher Scientific (Loughborough, UK); gel-drying film from Promega (Southampton, UK); and glacial acetic acid (AcOH) from VWR International (Lutterworth, UK).

2.1.1.5. Silver staining

Sodium thiosulphate from Sigma (Merck, Gillingham, UK); AcOH, formaldehyde and sodium carbonate from VWR International (Lutterworth, UK); gel-drying film from Promega (Southampton, UK); and silver nitrate and methanol from Thermo Fisher Scientific (Loughborough, UK).

2.1.1.6. Western blotting

Mouse monoclonal anti-flag M2 alkaline phosphatase antibody, NaCl and tris base from Sigma (Merck, Gillingham, UK); immunoblot 0.2 μ m polyvinylidene difluoride (PVDF) membranes from Bio-Rad (Watford, UK); Western Blue substrate from Promega (Southampton, UK); skimmed milk powder from Tesco (Oxford, UK); methanol and Tween 20 from Thermo Fisher Scientific (Loughborough, UK); and potassium chloride (KCl) and glycine from VWR International (Lutterworth, UK).

2.1.1.7. Gelatin zymography

Triton X-100 from Thermo Fisher Scientific (Loughborough, UK); and tris base, CaCl₂, zinc chloride (ZnCl₂) and porcine gelatin from Sigma (Merck, Gillingham, UK).

2.1.1.8. Enzyme kinetics

Eppendorf tubes from STARLAB (Milton Keynes, UK); 96-well black microplate from Greiner bio-one (Stonehouse, UK); 96-well microplate with clear bottom and low evaporation lid from Corning (Thermo Fisher Scientific, Loughborough, UK); Abz-TESE~SRGAIY-Dpa-KK-NH₂ substrate **2** from GL Biochem (Shanghai, China) and Severn Biotech (Kidderminster, UK); library parent compound **5** and hit substrates **31–39** from GL Biochem (Shanghai, China); Mca-PLGL-Dpa-AR-NH₂ from Bachem (Bubendorf, Switzerland); 5-FAM-AE~LQGRPISIAK-TAMRA substrate **1** from Bachem (Bubendorf, Switzerland) and CT-1746 from Celltech Therapeutics (UCB, Brussels, Belgium). For selectivity studies with **1**, the proteinases MMP-1ΔC, MMP-2ΔC, MMP-3ΔC, MMP-9 and MMP-13ΔC and N-TIMP-3 were kindly gifted by the laboratory of Professor Hideaki Nagase. For selectivity studies with **2**, the proteinase MMP-2ΔC was purified in Chapter 5, MMP-13ΔC was from Enzo Life Sciences (Exeter, UK) and MMP-1, MMP-3 and MMP-9ΔC were generous retirement gifts from the laboratory of Professor Gillian Murphy (Cambridge University, UK).

2.1.2. Methods

2.1.2.1. Tissue culture

An aliquot (1 ml) of HEK293 cells (TS4-4/TS5-5/MMP-2ΔC) was thawed from a cryogenic freezer (-169 °C) and rapidly warmed to 37 °C. DMEM with serum containing 10% (v/v) FBS and 1% (v/v) PS was warmed to 37 °C and added to the cells in a tissue culture dish (10 cm, 20 ml). This dish was left overnight at 37 °C in a Heracell 150i incubator with 5% CO₂ (Thermo Fisher Scientific, Loughborough, UK). The next day, medium was replaced with

DMEM (20 ml) and cell confluence checked using an Olympus CKX31 light microscope (Olympus, Southend-on-Sea, UK). Cells were then cultured according to the desired domain deletion mutant.

TS4-4 domain deletion mutant

After three days, fully confluent cells were seeded in new wells (2 x 2 ml), which were treated with 100 µg/ml of hygromycin B and regularly passaged for ~6 weeks. Antibiotic treatment was then stopped, and sufficient wells were seeded (10 x 2 ml) for expansion to intermediate (10 x 10 cm, 20 ml) and eventually large (10 x 15 cm, 50 ml) tissue culture dishes. Fully confluent cells from five tissue culture plates (10 cm) were re-suspended in a solution (5 ml) of 40% (v/v) FBS, 10% (v/v) DMSO and 50% (v/v) DMEM (10% FBS + 1% PS), transferred to cryovials (5 x 1.0 ml aliquots) and frozen down to -80 °C in a “Mr Frosty” box (Thermo Fisher Scientific, Loughborough, UK). After 1–3 days, the cryovials were transferred to a cryogenic freezer for long term storage at -169 °C. After complete confluence had been achieved for cells in the large 10 x 15 cm culture plates, media was changed to serum-free DMEM containing sterile-filtered 0.2% (w/v) LEH in distilled water (0.45 µm filter used) and 1% (v/v) PS (10 x 50 ml). Approximately 500 ml of media was collected once a week for a fortnight and then every four days a further four times. It was then centrifuged (1146 x g, 10 min, 4 °C) in a Hettich Rotina 420 R centrifuge (Hettich, Salford, UK). The supernatant was decanted and filtered using a 0.22 µm vacuum filter and stored at 4 °C overnight prior to next-day purification by affinity chromatography. Once sufficient media containing TS4-4 had been collected (2.5 L), all 15 cm plates were treated with 1% (v/v) Virkon in water (10 min) and the resultant solution containing dead cells disposed of down the drain.

TS5-5 domain deletion mutant

This was performed in a similar fashion to TS4-4, except that cells were treated with 800 µg/ml hygromycin B with regular passaging for one week before expansion to intermediate (10 x 10 cm) and large (10 x 15 cm) tissue culture dishes. Media was changed to serum-free and ~250–500 ml of conditioned medium was collected after 7 days and processed as described for TS4-4, except that 0.02% (w/v) NaN₃ was added before freezing down to -20 °C. When needed, cells were frozen down to -80 °C as described for TS4-4. When collection had finished (a total of 6.3 L of conditioned media was collected), cells were lysed and disposed of as described for TS4-4.

ProMMP-2ΔC domain deletion mutant

This was carried out using the same procedure as for TS5-5, except that cells were treated with 800 µg/ml hygromycin B with regular passaging for 2 weeks before expansion to intermediate (10 x 10 cm) and large (10 x 15 cm) tissue culture dishes. Media was changed to serum-free, and ~500 ml of conditioned medium was collected every 3-5 days and processed as described for TS5-5. When needed, cells were frozen down to -80 °C as described for TS4-4. When collection had finished (a total of 2 L of conditioned media was collected), cells were lysed and disposed of as described for TS5-5.

2.1.2.2. Purification of TS4-4 and TS5-5 by FLAG affinity chromatography

The following buffers were used: equilibration buffer consisting of 50 mM tris-HCl, pH 7.5, and 500 mM NaCl; wash buffers of 0.2 M glycine-HCl pH 6.0 and pH 5.0; an elution buffer of 0.2 M glycine-HCl pH 3.0 and a neutralisation buffer of 1M or 2M Tris-HCl, pH 7.5. All buffers were adjusted to pH with a HI3220 pH/ORP meter from Hanna Instruments (Leighton Buzzard, UK). Glycine buffers were filtered through a 0.22 µm vacuum filter and stored at 4 °C. A slurry of anti-flag M2 affinity gel (2 column volumes–CVs, where 1 CV = 1.5 ml) was

packed into a protein purification C column (Cytiva, Little Chalfont, UK) using equilibration buffer (2.5 ml/min) on an ÄKTA prime 2.0 system (Cytiva, Little Chalfont, UK). After washing with equilibration buffer (13 CVs, 2.0 ml/min), collected media (~500 ml) was thawed from -20 °C to room temperature (TS5-5) or used directly (TS4-4) and loaded onto the column while on ice (2.0 ml/min). This was sequentially washed with 0.2 M glycine-HCl pH 6.0 and pH 5.0 (13 CVs each, 2.0 ml/min) with collection of 4 x 5.0 ml fractions from each wash. The bound proteinase was eluted with 0.2 M glycine-HCl pH 3.0 and 28 x 1.0 ml fractions were collected and immediately neutralised with 1M or 2M Tris-HCl, pH 7.5 (200 µl or 100 µl). A final column wash was performed with equilibration buffer at the end of each purification (13 CVs, 2.0 ml/min). When not in use, affinity gel was stored at 4 °C in a buffer consisting of 50 mM tris-HCl, pH 7.5, 150 mM NaCl, 10 mM CaCl₂ and 0.02% (w/v) NaN₃ (henceforth defined as TNC buffer).

2.1.2.3. Purification of proMMP-2ΔC by gelatin-sepharose affinity chromatography

Purification was performed in the cold room at 4 °C. The following buffers were used: equilibration buffer consisting of TNC; dilution buffer comprising 100 mM tris-HCl, pH 8.0, with 0.02% (w/v) NaN₃; and 1% (v/v) DMSO and 5% (v/v) DMSO TNC, pH 7.5, wash and elution buffers, respectively. All buffers were adjusted to pH with a HI3220 pH/ORP meter from Hanna Instruments (Leighton Buzzard, UK). A slurry of Gelatin-Sepharose 4B affinity resin (1 CV, where 1 CV = 25 ml) was packed into an XK 26/20 column (Cytiva, Little Chalfont, UK) using de-ionised water (5 CVs, 7.0 ml/min) on an ÄKTA prime 2.0 system (Cytiva, Little Chalfont, UK). The column was washed with 5% (v/v) DMSO TNC buffer, pH 7.5 (2 CVs, 2.0 ml/min) and equilibrated with TNC buffer (4 CVs, 2.0 ml/min). Media containing proMMP-2ΔC was thawed to room temperature (RT) and diluted 1:1 with the dilution buffer and loaded onto the column (2.0 ml/min). The column was washed with 1% (v/v) DMSO TNC (2

CVs, 2.0 ml/min). Bound proMMP-2ΔC proteinase was eluted with 5% (v/v) DMSO TNC buffer, pH 7.5 and 25 x 5 ml fractions collected alongside smaller aliquots for each fraction (40 μl). Both of these sets of fractions were frozen at -20 °C. A final column wash was performed with equilibration buffer at the end of each purification (2 CVs, 2.0 ml/min). Small aliquots were also collected from the flow-through at each step and frozen at -20 °C for gelatin zymography (20 μl). The column was left at 4 °C when not in use.

2.1.2.4. Removal of DMSO from collected fractions by native sephacryl affinity chromatography

A slurry of sephacryl S-200 high resolution resin (1.5 CV, where 1 CV = 500 ml) was packed into an XK 26/100 column (Cytiva, Little Chalfont, UK) using equilibration buffer (5 ml/min) on an ÄKTA fast protein liquid chromatography (FPLC) machine (UPC-900 monitor and P-920 pump; Cytiva, Little Chalfont, UK). The proteinase fractions were thawed from -20 °C, pooled and initially concentrated to 45 ml in a Vivaspin 20 protein spin concentrator (1146 x g, 4 °C). This sample was split into three fractions (15 ml each), two of which were stored at -20 °C (~ 3 weeks) and one of which was stored at 4 °C (~ 3 weeks). One fraction from each temperature was concentrated to 5 ml (1146 x g, 4 °C) and an aliquot (150 μl) frozen to -20 °C for gelatin zymography. After an absorbance reading at 280 nm (A_{280}) had been taken for both fractions (5 ml), they were both stored at 4 °C. After two days, the fraction initially stored at 4 °C was loaded onto the column (2.0 ml/min). The purification was performed with equilibration buffer (2.0 ml/min) and progress monitored at a wavelength of 280 nm using Unicorn 5.2 software (Cytiva, Little Chalfont, UK). After an initial UV reading was recorded (0.29 CVs, 2.0 ml/min), 3 x 10 ml fractions (B1–B3) were collected and frozen at -20 °C. After a second UV reading was recorded (0.53 CVs, 2.0 ml/min), 9 x 10 ml fractions (C1–C9) were collected and frozen at -20 °C. In each case, small aliquots of each B and C fraction (1 x 150 μl) were sampled from the fractions and frozen at -20 °C for gelatin zymography. After 1.2

CVs had passed through the column, no further UV signals were recorded and the purification was stopped. This purification was repeated as before for the second 5 ml sample initially stored at -20 °C.⁷

2.1.2.5. SDS-PAGE

SDS-PAGE was carried out using the Laemmli method (25). Polyacrylamide gels (containing 4% total acrylamide for the stacking gel and 15% total acrylamide for the resolving gel) were cast by using either an SE215 multi-gel caster or an SE245 dual-gel caster (Hoefer, Thermo Fisher Scientific, Loughborough, UK) (173). Loading buffer (Laemmli, 2X), consisting of 0.5 M tris-HCl, pH 6.8, 4% (v/v) SDS, 0.005% (v/v) bromophenol blue, 20% (v/v) glycerol and 5% (v/v) β -mercaptoethanol was added to an equal volume of sample and the resultant mixture boiled at 90 °C (10 min). After centrifugation (12,519 x g, 10 min, RT) samples were loaded onto the gel in running buffer (1X), comprising 0.02 M tris base, pH 8.3, 0.2 M glycine and 3.5 mM SDS. The gel was initially run at 100 V (20 min) and then at 150 V until the tracking dye (bromophenol blue) had reached the bottom of the gel.

2.1.2.6. Coomassie staining

Gels that had been run by SDS-PAGE were stained using a solution of 0.1% (w/v) CBB, 50% (v/v) methanol, 20% (v/v) glacial acetic acid and 30% (v/v) de-ionised water for 20 minutes. De-staining was performed using 30% (v/v) methanol, 1% (v/v) formic acid and 69% (v/v) de-ionised water until a clear background was obtained. Images were acquired under manual capture settings using a G-box XX6 gel imager (Syngene, Cambridge, UK) and gels were dried overnight and stored long-term in gel drying film (Promega, Southampton, UK).

⁷ The un-unused fraction at -20 °C (15 ml) was meant to be used for later purification but was not required in the end.

2.1.2.7. Silver staining

Gels were stained as reported by Shevchenko (174). All solutions were made up in de-ionised water unless otherwise stated. Initially, gels were fixed in 50% (v/v) methanol and 5% (v/v) AcOH at 4 °C (20 min or overnight). Sequential washing with 50% (v/v) methanol (10 min) and de-ionised water (10 min) was followed by the addition of 0.02% (w/v) sodium thiosulphate sensitiser (1 min). Further washing (2 x 1 min, de-ionised water) and incubation in cold 0.1% (w/v) silver nitrate (4 °C, 20 min or overnight) ensued. After washing with de-ionised water (2 x 1 min), development was performed with 2% (v/v) sodium carbonate containing 0.4% (v/v) formalin (2 x 5–7 min). After removal of this solution, development was stopped by the addition of 5% (v/v) AcOH (5 min) followed by washing with de-ionised water (2 x 30 min). Images were acquired and gels dried and stored as described in section 2.1.2.6.

2.1.2.8. Western blotting

Proteins on SDS-PAGE gels were transferred onto methanol-activated PVDF membranes at 70 V (90 min) in a transfer tank from Peqlab Biotechnologie GmbH (VWR International, Lutterworth, UK) containing Towbin transfer buffer comprising 25 mM tris base, 192 mM glycine and 20% (v/v) methanol in de-ionised water. The membranes were blocked with 5% (w/v) skimmed milk in TBST buffer consisting of 19 mM tris base, 137 mM NaCl, 2.7 mM KCl and 0.05% (v/v) Tween 20 at room temperature (60 min) or at 4 °C overnight. This was followed by two rinses (10 s) and two washes (5 min) with TBST buffer, and incubation with a 1:1000 dilution of mouse monoclonal anti-flag M2 alkaline phosphatase antibody in 5% (w/v) skimmed milk TBST (60 min). After five washes with TBST buffer (5 min), membranes were developed with Western Blue substrate (5–10 ml), rinsed with de-ionised water (5 x 10 s) and imaged under automatic capture settings using a G-box XX6 gel imager (Syngene, Cambridge, UK).

2.1.2.9. Gelatin zymography

Gelatin zymography was performed according to an Abcam protocol (175). Polyacrylamide gels (containing 4% total acrylamide for the stacking gel and 15% total acrylamide for the resolving gel) were cast using a SE215 multi-gel caster (Hoefer, Thermo Fisher Scientific, Loughborough, UK) with 4 mg/ml (w/v) gelatin. SDS-PAGE was run as described in section 2.1.2.5 but under non-reducing conditions and without boiling of samples. Gels were sequentially washed with washing buffer (1 x 15 min, 1 x 45 min) comprising 50 mM tris-HCl, pH 7.5, 5 mM CaCl₂, 1 μM ZnCl₂ and 2.5% (v/v) Triton X-100 in water; and incubation buffer (5 min, room temperature) containing 50 mM tris-HCl, pH 7.5, 5 mM CaCl₂, 1 μM ZnCl₂ and 1% (v/v) Triton X-100 in water. After removal of the latter solution, fresh incubation buffer was added and the gels allowed to stand at room temperature to enable gelatin hydrolysis to proceed (20 h). The gels were then analysed by Coomassie staining as described in section 2.1.2.6.

2.1.2.10. Concentration and long-term storage of metalloproteinases

The membranes of Vivaspin 20 protein spin concentrators (molecular weight cut-off = 3 kDa) were wetted with TNC buffer (10 ml) and centrifuged (1146 x g, 10 min, RT). Fractions containing purified metalloproteinases (TS4-4, TS5-5 or MMP-2ΔC) were pooled and concentrated to lower volumes (2–20 ml, depending on the amount of purified proteinase obtained) in these Vivaspin 20 tubes (1146 x g, RT). The absorbance (A_{280}) of each sample was measured using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Loughborough, UK). This was used to determine the initial concentration of each proteinase from the equation for the Beer-Lambert-Bouguer law (176). The following molar extinction coefficients were computed from the ExPASy ProtParam tool assuming unreduced cysteine

residues: 27,470 M⁻¹ cm⁻¹ (TS4-4), 29,420 M⁻¹ cm⁻¹ (TS5-5) and 73,060 M⁻¹ cm⁻¹ (MMP-2ΔC).⁸ Each proteinase was frozen down to -80 °C for long-term storage in 0.5 ml (MMP-2ΔC) or 1 ml (TS4-4/TS5-5) fractions in TNC buffer containing 10% (v/v) glycerol and 0.05% (v/v) Brij-35.

2.1.2.11. Fluorophore calibration curves

Calibration curves were performed in TNC buffer under six different concentrations of Brij-35 surfactant: 0%, 0.0005%, 0.005%, 0.02%, 0.05% and 5% (v/v). The only exception was for the Mca fluorophore, where only 0.005% (v/v) Brij-35 was used. The calibration curves for each fluorophore can be found in the Appendix section of Chapter 4.

5-Carboxyfluorescein (5-FAM)

A stock solution of 5-carboxyfluorescein in DMSO (1 M) was serially diluted to 1 mM in DMSO, before further dilution to an initial concentration of 4.0 μM using TNC buffer containing Brij-35. This solution was diluted in 1.5-fold increments to 9.1 nM in TNC Brij-35 buffer. Each concentration in the 4.0 μM–9.1 nM range (inclusive) was added to wells (50 μl) in a 96-well black microplate (Greiner bio-one, Stonehouse, UK) in triplicate. A background control consisting of TNC buffer with Brij-35 (50 μl) was also added to wells in triplicate. The relative fluorescence intensity was measured at 37 °C on a SpectraMax M5 spectrophotometer (Molecular Devices, San Jose, USA) using an excitation wavelength (λ_{ex}) = 485 nm and emission wavelength (λ_{em}) = 538 nm. A plot of relative fluorescence units (RFU) against concentration (μM) followed by linear regression analysis with Prism 9 software (GraphPad, La Jolla, USA) gave the gradient in RFU μM⁻¹. This was used to convert units from RFU s⁻¹ to μM s⁻¹ in Chapters 4 and 5.

⁸ Sequence data for each proteinase was obtained from the UniProtKB database and domain cut-offs from relevant literature of Kashiwagi *et al* (16) Gendron *et al* (17) and Troeberg *et al* (62).

2-Aminobenzoic acid (Abz)

The fluorophore dilution series was prepared as described for 5-FAM, except that the concentration range and wavelengths were: 20 μM –0.10 μM (inclusive) and λ_{ex} = 300 nm and λ_{em} = 430 nm, respectively.

7-Methoxycoumarin-4-acetic acid (Mca)

The fluorophore dilution series was performed as described for 5-FAM, except that only 0.005% (v/v) Brij-35 was used. The concentration range was 20 μM –0.10 μM (inclusive) and the wavelengths were λ_{ex} = 328 nm and λ_{em} = 420 nm.

2.1.2.12. General information on kinetic studies with ADAMTSs and MMPs

All proteinase assays were performed in triplicate at 37 °C in a SpectraMax M5 spectrophotometer (Molecular Devices, San Jose, USA) unless otherwise stated. The following substrates and parameters were used for ADAMTS and MMP proteinases: 5-FAM-AE~LQGRPISIAK-TAMRA (**1**) for TS4-4, with λ_{ex} = 485 nm and λ_{em} = 538 nm; Abz-TESE~SRGAIY-Dpa-KK-NH₂ (**2**) for TS5-5, with λ_{ex} = 300 nm and λ_{em} = 430 nm; Mca-PLG~L-Dpa-KK-NH₂ for all MMPs, with λ_{ex} = 328 nm and λ_{em} = 420 nm; and library parent compound **5** with hit substrates **31**–**39** for TS4-4, TS5-5, MMP-2 ΔC and MMP-9 ΔC , with λ_{ex} = 300 nm and λ_{em} = 430 nm. All analyses were carried out using Prism 9 software (GraphPad, La Jolla, USA).

2.1.2.13. Linear time frame of substrate hydrolysis by TS4-4 and TS5-5

The linear time frame of 5-FAM-AE~LQGRPISIAK-TAMRA substrate hydrolysis by TS4-4 and Abz-TESE~SRGAIY-Dpa-KK-NH₂ substrate hydrolysis by TS5-5 was determined in three independent experiments at a final proteinase concentration of 0.5 nM. A solution (180 μl) of 5-FAM-AE~LQGRPISIAK-TAMRA (0.56 μM) and Abz-TESE~SRGAIY-Dpa-KK-NH₂ (22.22 μM) in freshly prepared TNC buffer with 0.005% (v/v) Brij-35 was added to Eppendorf tubes.

Control tubes containing substrate alone were also prepared. Each substrate was incubated at 37 °C in a dry incubator (30 min) before the addition of TS4-4, TS5-5 (both 5 nM) to sample tubes and buffer alone to control tubes (all 20 µl). All Eppendorf tubes were incubated at 37 °C until the endpoint RFU of each solution was measured in a 96-well plate at the following time points: 0 h, 2 h, 4 h, 18 h, 20 h, 22 h, 24 h in the first experiment and 0 h, 4 h, 8 h in the second experiment (which was performed twice). When solutions were transferred back after each fluorescence reading, Eppendorf tubes were centrifuged (12,590 x g, 1 min, RT) to settle the liquid to the bottom. The data for these experiments can be seen in Figure 3.5 of Chapter 3.

2.1.2.14. Activity studies of ADAMTSs and MMPs

Activity studies are sub-divided by proteinase family.

ADAMTSs

Three concentrations of TS4-4 and TS5-5 proteinases were assayed for each substrate. A solution (180 µl) of 5-FAM-AE~LQGRPISIAK-TAMRA (0.56 µM) or Abz-TESE~SRGAIY-Dpa-KK-NH₂ (22.22 µM) in freshly prepared TNC buffer containing 0.02% (v/v) Brij-35 was added to a 96-well black microplate for each proteinase concentration alongside a background control to which buffer alone would be added. The resultant solutions were incubated at 37 °C (30 min). TS4-4 or TS5-5 proteinase in TNC buffer with 0.02% (v/v) Brij-35 (20 µl) was added to wells to give final proteinase concentrations of 20 nM, 10 nM and 5 nM (TS4-4) or 80 nM, 40 nM and 20 nM (TS5-5). TNC buffer with 0.02% (v/v) Brij-35 (20 µl) was also added to background control wells. The change in RFU was measured at 37 °C for 1 h, with readings every 20 s. Following subtraction of background fluorescence intensity, linear regression analysis of the linear part of each curve was used to determine the cleavage rate, which was expressed as the mean ± standard error of the mean (SEM) in RFU s⁻¹. The data

for these experiments can be seen in Figure 3.3 of Chapter 3. For cases where a single concentration of TS4-4 and TS5-5 was assayed, the same procedure was used (*e.g.* Figure 8.9, Chapter 4 Appendix).

MMPs

A solution (180 μ l) of Mca-PLG~L-Dpa-KK-NH₂ (1.67 μ M) in freshly prepared TNC buffer with 0.005% (v/v) Brij-35 was added to a 96-well black microplate for each MMP to be assayed along with a background control to which buffer alone would be added. The solutions were incubated at 37 °C (30 min). A matrix metalloproteinase (one of MMP-1, MMP-1 Δ C, MMP-2 Δ C, MMP-3, MMP-3 Δ C, MMP-9, MMP-9 Δ C, MMP-13 Δ C) was added to wells at a specific concentration in the 0.5–140 nM range (20 μ l). TNC buffer with 0.005% (v/v) Brij-35 (20 μ l) was added to background control wells. The relative fluorescence intensity was measured at 37 °C for 1 h with readings every 20 s. Following subtraction of background fluorescence intensity, linear regression analysis of the linear part of each curve was used to determine the cleavage rate, which was expressed as the mean $\pm$ SEM in RFU s⁻¹. This was divided by the final concentration of proteinase (nM), then divided by the gradient of the calibration curve for Mca (expressed in RFU nM⁻¹) and finally multiplied by the concentration of substrate (nM) to give the rate in nM s⁻¹. The data for these experiments can be seen in Figure 8.9 (Appendix Chapter 4).

2.1.2.15. Surfactant activity studies of TS4-4 and TS5-5

A 5% (v/v) Brij-35 solution in TNC buffer was diluted with TNC to 0.05%, 0.02%, 0.005% and 0.0005% (v/v) Brij-35 surfactant. For 5% and 0% (v/v) Brij-35, buffer or TNC alone were used, respectively. Activity assays were performed for each surfactant concentration using the method described in the previous section, except that a single concentration of TS4-4 or TS5-5 was used (20 nM). For each surfactant concentration, experiments were

repeated 3–9 times for TS4-4 and 3–7 times for TS5-5 in triplicate or quadruplicate, with final rates expressed as the mean $\pm$ standard deviation (SD) in RFU nM⁻¹. The data for these experiments can be found in Figure 4.1 of Chapter 4.

2.1.2.16. Active site titrations of ADAMTSs and MMPs with N-TIMP-3 and CT-1746

Titration are sub-divided here based on the type of inhibitor used (N-TIMP-3 or CT-1746). The data for these experiments can be found in Figure 3.4 as well as the Appendices of Chapters 3 and 4.

N-TIMP-3

A stock solution of N-TIMP-3 (5.6 μ M or 6.3 μ M) was thawed from -20 °C and serially diluted to 320 nM, 160 nM, 80 nM, 40 nM, 20 nM, 10 nM, 5 nM and 2.5 nM concentrations in freshly prepared TNC with 0.02% (v/v) Brij-35. Each concentration (50 μ l) was added to a 96-well black microplate. Proteinases were thawed from -20 °C or -80 °C, prepared at an initial concentration in TNC buffer with 0.02% (v/v) Brij-35, and added to each well (50 μ l). The initial concentrations for TS4-4, TS5-5 and selected MMPs (MMP-1 Δ C, MMP-2 Δ C, MMP-3 Δ C and MMP-9) were 10 nM, 27 nM and 40 nM, respectively. For TS4-4, TS5-5 and MMP-2 Δ C purified in Chapter 5, the concentration was based on initial A₂₈₀ readings, whereas for the other MMPs the concentration was based on previous active-site titrations performed by the Nagase laboratory. Separate wells containing TS4-4/TS5-5/MMP (50 μ l) with TNC containing 0.02% (v/v) Brij-35 (50 μ l)—as well as those with TNC containing 0.02% (v/v) Brij-35 alone (100 μ l)—were set up as positive and negative controls, respectively. The mixtures were incubated at 37 °C for 1 h. A solution (100 μ l) of 5-FAM-AE~LQGRPISIAK-TAMRA substrate (1 μ M) or Abz-TESE~SRGAIY-Dpa-KK-NH₂ substrate (40 μ M) or Mca-PLG~L-Dpa-KK-NH₂ substrate (3 μ M) in TNC with 0.02% (v/v) Brij-35 was added to wells and the relative

fluorescence intensity measured at 37 °C for 1 h with readings every 20 s. The rates from the different concentrations of N-TIMP-3 were determined by linear regression and expressed in terms of the mean percentage activity $\pm$ SEM by dividing by the rate in the absence of inhibitor. This data was then fitted to the equation for tight-binding inhibition (1) (177):

$$a = 100 \left(1 - \frac{([E] + [I] + K_i) - \{([E] + [I] + K_i)^2 - 4[E][I]\}^{1/2}}{2[E]} \right) \quad (1)$$

where “a” is the percentage activity (proteolysis with inhibitor divided by proteolysis without inhibitor); [E] the concentration of active proteinase; [I] the concentration of the inhibitor, and $K_{i(app)}$ the apparent inhibition constant of the inhibitor. Extrapolation of the linear section of the curve to the abscissa provided the concentration of active proteinase. If the initial concentration of TS4-4/TS5-5/MMP-2 Δ C determined by A_{280} differed from that calculated by active-site titration, the assay was repeated until a consistent proteinase concentration was obtained.

CT-1746

For titrations using CT-1746, the same procedure was employed except that the following initial proteinase concentrations were used: TS4-4/TS5-5 (480 nM), MMP-1 and MMP-9 Δ C (40 nM) and MMP-3 (4 nM). For the inhibitor dilution series, either identical concentrations were used as for N-TIMP-3, or eight concentrations from the following series: 4 μ M, 2 μ M, 1 μ M, 500 nM, 250 nM, 125 nM, 63 nM, 31 nM, 16 nM, 8 nM, 2 nM and 1 nM. The choice of the inhibitor dilution series depended on the initial concentration of the ADAMTS or MMP proteinase. Please note that titration was not performed for both batches of MMP-13 Δ C as insufficient quantities were available.

2.1.2.17. Determination of CT-1746 inhibition constants for TS4-4/TS5-5

For assays with TS4-4, a stock solution of CT-1746 inhibitor in DMSO (1 mM) was thawed from -20 °C and serially diluted to 250 nM, 125 nM, 63 nM, 31 nM, 16 nM and 8 nM concentrations using freshly prepared TNC with 0.02% (v/v) Brij-35. For assays with TS5-5, a two-fold dilution series with concentrations in the 1 µM–8 nM range (inclusive) was generated. Each inhibitor concentration was added to Eppendorf tubes (50 µl). Proteinases were thawed from -20 °C or -80 °C and prepared at a specific concentration in the 1.6 nM–20 nM range (TS4-4) or 10 nM–60 nM range (TS5-5) in TNC buffer with 0.02% (v/v) Brij-35.⁹ TS4-4/TS5-5 was added to each tube containing inhibitor (50 µl). Separate tubes containing TS4-4/TS5-5 (50 µl) with TNC containing 0.02% (v/v) Brij-35 (50 µl)—as well as those with TNC containing 0.02% (v/v) Brij-35 alone (100 µl)—were set up as positive and negative controls, respectively. The mixtures were incubated at 37 °C (1 h). A solution (100 µl) of 5-FAM-AE~LQGRPISIAK-TAMRA substrate (1 µM) or Abz-TESE~SRGAIY-Dpa-KK-NH₂ substrate (40 µM) was added to the tubes and the resultant mixture left in a dry incubator at 37 °C. Endpoint relative fluorescence intensity was then measured at a single time point after incubation between 1–6 h.¹⁰ Data was fitted to the equation for tight-binding inhibition (1) (177) and the apparent inhibition constant $K_{i(app)}$ determined. This was then used to determine the actual K_i value using the following equation (2):

$$K_i = \frac{K_{i(app)}}{1 + \frac{[S]}{K_m}} \quad (2)$$

where K_m is the Michaelis constant—15 µM for TS4-4 (63) and 76 µM for TS5-5 (Troeborg, 2009 #205)—and $[S]$ the substrate concentration used in the experiment. The experiments

⁹ Please refer to Table 3.1 and Table 3.2 of Chapter 3 for final concentration values for TS4-4 and TS5-5.

¹⁰ For TS4-4 K_i determinations outside the linear range of substrate hydrolysis, incubation for 18-22 h was performed.

were repeated several times using different purification batches, concentrations, incubation times and initial storage temperatures to ensure that a consistent K_i value was obtained. The final data for TS4-4 and TS5-5 can be found in Tables 3.1 and 3.2 of Chapter 3.

2.1.2.18. Selectivity studies with ADAMTSs and MMPs

Selectivity studies are sub-divided here based on one procedure used for **1** and another for **2**, the library parent compound **5** and hit peptides **31–39**.

5-FAM-AE~LQGRPISIAK-TAMRA (1)

A solution (180 μ l) of **1** (0.56 μ M) in freshly prepared TNC buffer with 0.005% (v/v) Brij-35 was added to wells in a 96-well black microplate and incubated at 37 °C (30 min). An ADAMTS or MMP proteinase was added to substrate wells (20 μ l) to give a final proteinase concentration of either 50 nM (TS4-4/TS5-5), 2 nM (MMP-13 Δ C), 10 nM (MMP-9) or 50 nM (MMP-1 Δ C/-2 Δ C/-3 Δ C). TNC buffer with 0.005% (v/v) Brij-35 was added to substrate-only control wells (20 μ l). The relative fluorescence intensity was measured at 37 °C for 1 h with readings every 40 s. For proteinases where substrate cleavage was not observed after one hour, the same assay was performed using Eppendorf tubes. Here, incubation was carried out at 37 °C in a dry incubator and end-point fluorescence intensity measured at 0 h, 4 h, 8 h and 24 h time points. For each MMP or ADAMTS proteinase, the assay was performed 2–8 times and rates of substrate cleavage expressed in μ M s⁻¹ as described in the activity studies (section 2.1.2.14). Final rates were reported as the percentage of substrate cleavage by TS4-4 (mean $\pm$ SEM). The data for these experiments can be found in Figure 4.2 of Chapter 4.

Abz-TESE~SRGAIY-Dpa-KK-NH₂ (2) and peptide substrates 5 and 31–39

This was carried out using the 1 h procedure for **1**, except that **2**, the library parent compound **5** and hit peptides **31–39** were used at a concentration of 22.22 μ M. A final volume of 100 μ l was used to conserve substrate and proteinase. Final proteinase

concentrations were 50–90 nM (TS4-4), 25–200 nM (TS5-5), 5–10 nM (MMP-2ΔC), 30 nM (MMP-9ΔC), 92 nM (MMP-1) and 3 μM (MMP-3). Activity was read from underneath a 96-well microplate with a clear bottom covered with a low evaporation lid. However, if no cleavage was observed after 1 h, fluorescence intensity was read continuously at 37 °C for 23 h every 5 min. For each ADAMTS and MMP proteinase, the assay was performed 2–4 times. The only exception was MMP-1, where no cleavage was observed from a single experiment at a final concentration of 92 nM after 72 h. Final rates were expressed as a percentage of substrate cleavage by TS5-5 (mean ± SEM). The data for these experiments can be found in Figure 4.2 of Chapter 4 and Figure 5.7 of Chapter 5.

2.1.2.19. Determination of kinetic parameters for substrate cleavage by TS5-5: k_{cat} and K_m

A solution of **2/5/39** (5 mM) or **32/36** (1 mM) in DMSO underwent a two-fold serial dilution with freshly prepared TNC buffer containing 0.005% (v/v) Brij-35. This gave initial concentrations in the following range: 222.22 μM – 0.87 μM (Abz-TESE~SRGAIY-Dpa-KK-NH₂/**5/36/39**) and 177.78 μM – 1.39 μM. Eight substrate concentrations (180 μl) within these ranges (including outer numbers) were added to wells in a 96-well black microplate and incubated at 37 °C (20 min). Control wells for each substrate concentration (180 μl) were also added, but in single replicate. TS5-5 proteinase was thawed from -80 °C or -20 °C, made up in TNC buffer with 0.005% (v/v) Brij-35 (300 nM) and added to substrate wells (20 μl). TNC buffer was added to substrate-only control wells (20 μl). Fluorescence intensity was measured at 37 °C for 1 h with readings every 20 s. The rate of substrate cleavage was determined at each substrate concentration and expressed as the mean ± SEM in μM s⁻¹. The subsequent data was fitted to the Michaelis-Menten equation (3):

$$v = V_{max} \frac{[S]}{K_m + [S]} \quad (3)$$

Where v is the rate of the reaction, $V_{\max}$ the maximal rate of the reaction, K_m the Michaelis constant and $[S]$ the substrate concentration. This equation enabled the K_m value to be determined. The catalytic rate constant k_{cat} was determined from equation (4):

$$V_{\max} = k_{\text{cat}}[E] \quad (4)$$

Where $[E]$ is the concentration of proteinase, and k_{cat} and $V_{\max}$ are as defined previously. These initial K_m and k_{cat} parameters were determined from $n = 3-4$ independent experiments and the results are shown in the Appendix of Chapter 5. Final values for these parameters were reported after a correction to account for the inner-filter effect, which is described in the next section.

2.1.2.20. Correcting kinetic parameters for the inner-filter effect

This was performed based on the literature protocol by Liu *et al* (178). Two sets of initial substrate concentrations were prepared as described in the previous section. Both sets of substrate concentrations (180 μl) were added to wells in a 96-well black microplate. Two buffer-alone controls consisting of TNC buffer containing 0.005% (v/v) Brij-35 (180 μl) were also added to wells. All solutions were incubated at 37 °C (20 min). Abz fluorophore (1 M) in DMSO was serially diluted to 1 mM in DMSO, before further dilution to an initial concentration of 100 μM using TNC buffer containing 0.005% (v/v) Brij-35. TNC buffer containing 0.005% (v/v) Brij-35 (20 μl) was added to the first set of substrate concentrations as well as to the first buffer-alone control. Abz fluorophore (20 μl , 100 μM) was added to the second set of substrate concentrations as well as the second buffer-alone control. The end-point fluorescence intensity was measured at 37 °C using $\lambda_{\text{ex}} = 300 \text{ nm}$ and $\lambda_{\text{em}} = 430 \text{ nm}$. At each substrate concentration, the RFU with Abz fluorophore (RFU + Abz) was subtracted from the RFU in the absence of Abz fluorophore (RFU) to give the actual fluorescence of the Abz fluorophore (RFU_{Abz}). These values were divided by RFU_{Abz} at a substrate concentration of 0

μM to give the correction factor for each concentration. The rates for each substrate concentration described in the previous section were divided by their respective correction factor to deliver the inner-filter corrected K_m and k_{cat} values.

2.1.2.21. Evaluating the rate of cleavage of library parent compound 5 and hit peptide 39 at different substrate concentrations

This was performed according to the 1 h procedure for **2** (section 2.1.2.18) except that a range of substrate concentrations was used for **5** and **39** (20 μM – 0.63 μM). Additionally, control wells for each substrate concentration were used in single replicate. The final concentration of TS5-5 was 57 nM and readings were performed every 20 s.

2.1.2.22. Determination of the lowest concentration of TS5-5 that can be detected

Literature substrate **1**, library parent compound **5** and hit peptide **39** (5 mM) were each diluted to two concentrations (5.56 μM and 2.78 μM) in freshly prepared TNC buffer with 0.005% (v/v) Brij-35. The two concentrations for each substrate (90 μl) were added to wells in a 96-well black microplate. A control (90 μl), which would contain substrate alone was also included. All wells were incubated at 37 °C (20 min). TS5-5 proteinase was thawed from -20 °C and made up at the following concentrations in TNC buffer with 0.005% (v/v) Brij-35: 10 nM, 5 nM, 2.5 nM and 1.25 nM. Each concentration of proteinase (10 μl) was added to the substrate wells, with buffer (10 μl) added to the control wells containing substrate alone. The relative fluorescence intensity was measured at 37 °C for 24 h, with readings every 2 min. To assess whether substrate cleavage was sufficiently above background to detect TS5-5 activity, the RFU of each substrate and proteinase concentration was compared to the substrate controls. The data for these experiments can be found in Figures 5.9 and 5.10 of Chapter 5.

2.2. Computational Chemistry

2.2.1. Modelling of peptide substrates into the active site of ADAMTS-5

The crystal structure of the combined catalytic and disintegrin domain of ADAMTS-5 (*i.e.* TS5-5, residues 264-555) (179) was retrieved from the RCS protein data bank (2RJQ, www.rcsb.org) and analysed by Molecular Operating Environment (MOE) software version 2018.12 (Chemical Computing Group, Montreal, Canada). Initially, the literature substrate Abz-TESE~SRGAIY-Dpa-KK-NH₂ was constructed using the builder tool, partial charges corrected using the Protonate 3D tool (< 5 min) and the lowest energy conformation obtained through energy minimisation using the Amber12:EHT forcefield with a gradient of 0.1 (30 min). The sequence was docked into the active site of TS5-5 and anchored in position by fixing the distance between the carbonyl group of Glu in P1 and the catalytic Zn²⁺ ion to ~2 Å. The substrate-TS5-5 complex was then solvated using a spherical droplet solvation shell of water with a margin of 4.0. The overall structure was minimised as before (30 min) and molecular dynamics simulations performed with rigid water molecules (24 h, 300 K) to give the starting point for the design process. From here, the initial P4' fluorophore substrate Y(NO₂)TESESRGK(Abz)IYYKKG was modelled according to Scheme 4.2 (Chapter 4) with a single change made at a time. For each change, a charge correction was performed (< 5 min), followed by an energy minimisation (30 min) and molecular dynamics simulations (30 min) using the methods described previously. Amino acids at specific positions were mutated directly to the desired residue using the protein builder tool. The nitro group (NO₂) and Abz moiety were introduced into the substrate by generating the structure through the builder tool and “stitching” this on through bond formation. Both parent compounds for the P7' and P9' fluorophore substrates were modelled from the parent P4' structure (see Schemes 4.3–4.4, Chapter 4). For the P4', P7' and P9' fluorophore series shown in Tables 4.1–4.3 of

Chapter 4, substrates were modelled starting from the relevant parent compound each time using the same general procedure described for the parent P4' substrate. A total of 17 substrates were designed and analysed.

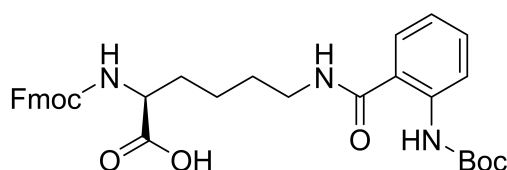
2.3. Combinatorial chemistry and chemical synthesis

2.3.1. General information

This was performed at the Center for Evolutionary Chemical Biology (CECB) at the University of Copenhagen, Denmark. All reagents were obtained from commercial suppliers and used as received without further purification. The only exception was PEGA₁₉₀₀ resin, which came from the now defunct VersaMatrix A/S. Peptides were synthesised in a custom 20-well Teflon synthesis block (180) using Fmoc-SPPS or split-mix combinatorial synthesis coupled with Fmoc-SPPS. Dimethylformamide (DMF) and 20% (v/v) piperidine/DMF were dispensed into the Teflon synthesis block using a bottle-top dispenser from Socorex Isba S.A. Shaking was carried out using an IKA-Vibrax-VXR or IKA KS 130 basic and vortexing with an IKA MS2 Minishaker. A strip heater was used to heat the IKA KS 130 basic shaker during the Kaiser test. TELOS SPE filtration columns with a polyethylene frit (20 µm pore size) were used for amino acid couplings outside of the Teflon synthesis block. Buffers were adjusted to the correct pH using a VWR pHenomenal pH 1100L pH meter with distilled water from a Siemens Ultra-Clear TWF water purification system. Thin layer chromatography was performed using 60F₂₅₄ silica plates from Merck and analysed under a UV lamp (254 nm) from Konrad Benda. Analytical liquid chromatography runs (13 min) were performed on an Agilent 1100 series HPLC using a Waters X-Bridge C18 5 µm column (4.6 x 100 mm) and solvent system of 90% acetonitrile (MeCN)/water + 0.1% trifluoroacetic acid (TFA) and 100% water + 0.1% TFA. After ten-fold sample dilution, peptide mass was confirmed by an analytical run (5 min) on a Waters Acquity H-Class UPLC with a XEVO G2-S QToF Zspray mass spectrometer (ESI-MS)

using a BEH C-18 1.7 μm column (2.1 x 50 mm) and a solvent system of water + 0.1% formic acid (FA) and MeCN + 0.1% FA. MALDI and tandem mass spectrometry were carried out on a Bruker solariX XR instrument, with data analysed using Bruker Compass DataAnalysis 5.1. ^1H -NMR spectroscopy was performed on a Bruker BioSpin at 400 MHz. Chemical shifts are referenced to residual solvent and reported in ppm with coupling constants listed in Hertz (Hz).

2.3.2. Synthesis of Fmoc-Lys(*retro*-Abz-Boc)-OH (22)



To a solution of *tert*-butyloxycarbonyl (Boc)-Abz-OH (3.1 g, 13 mmol., 1.2 equiv.), (1-Cyano-2-ethoxy-2-oxoethylidenaminoxy)dimethylamino-morpholino-carbenium hexafluorophosphate (COMU, 5.6 g, 13 mmol., 1.2 equiv.) in DMF (94 ml) was added *N,N*-diisopropylethylamine (DIPEA, 4.6 ml, 27 mmol., 2.5 equiv.) and the subsequent solution stirred at room temperature (30 min). Fmoc-L-lysine (4.0 g, 11 mmol., 1.0 equiv.) was added and the subsequent cloudy orange solution stirred overnight (21 h). HPLC and ultra-high performance liquid chromatography (UHPLC)-MS of the reaction mixture revealed the presence of product mass. Gravity filtration of the reaction mixture was followed by the addition of ethyl acetate (EtOAc, 100 ml) and brine (200 ml). The aqueous layer was extracted with EtOAc (3 x 100 ml), and the combined organic layers washed sequentially with saturated NaCl, 1M H_2SO_4 , 5% (w/v) sodium hydrogen carbonate (NaHCO_3), 1M sulphuric acid (H_2SO_4) and water (2 x 100 ml each). Drying with sodium sulphate, filtration and concentration *in vacuo* afforded a pale yellow solid. This was purified using silica gel column chromatography using a gradient system of 1-3% MeOH/dichloromethane (DCM) and dried under high

vacuum to deliver the title compound as an off-white solid with a purity $\geq 99\%$ by HPLC (1.21 g, 19%), which was used in the synthesis of ADAMTS-5 substrates. $R_f = 0.2$ (9:1 v/v) DCM/MeOH); $^1\text{H-NMR}$ (400 MHz, CDCl_3): δ 10.00 (s, 1H, OH), 8.16 (d, $J = 8.4$ Hz, 1H, ArH), 8.10 – 7.72 (m, 1H, NHCOO), 7.62 (d, $J = 7.5$ Hz, 2H, ArH), 7.43 (dd, $J = 7.7, 3.9$ Hz, 2H, ArH), 7.33 – 7.20 (m, 4H, ArH), 7.16 (t, $J = 7.8$ Hz, 2H, ArH), 6.75 (t, $J = 7.6$ Hz, 1H, ArH), 6.73 – 6.63 (m, 1H, CH_2NHCO), 5.70 – 5.61 (m, 1H, OCONH), 4.27 – 4.15 (m, 3H, CH_2O , Lys H_α), 4.04 (t, $J = 7.2$ Hz, 1H, CHCH_2O), 3.30 – 3.16 (m, 2H, Lys H_ϵ), 1.82 – 1.69 (m, 1H, Lys H_β), 1.65 – 1.53 (m, 1H, Lys H_β), 1.52 – 1.45 (m, 2H, Lys H_δ), 1.40 (s, 9H, CH_3), 1.35 – 1.23 (m, 2H, Lys H_γ) ppm. **LRMS:** LC-MS (ESI) m/z 588.1 $[\text{M} + \text{H}]^+$.

2.3.3. Linear synthesis of peptide substrates on PEGA₁₉₀₀ resin

PEGA₁₉₀₀ resin (8.4 g, 1.0 mmol., 0.12 mmol/g loading) was swelled in DMF (14 ml) at room temperature (10 min) in a TELOS column and solvent removed. Treatment with 20% (v/v) piperidine/DMF (14 ml, 20 min) was followed by washing with DMF (5 x 14 ml). To Fmoc-Gly-OH (1.2 g, 4.0 mmol., 4.0 equiv.) in DMF (14 ml) was added 4-Ethylmorpholine (NEM, 497 μl , 3.9 mmol., 3.9 equiv.) and the subsequent solution allowed to stand (≥ 1 min) at room temperature. After the addition of *N,N,N',N'*-Tetramethyl-O-(benzotriazol-1-yl)uronium tetrafluoroborate (TBTU, 1.3 g, 4.0 mmol., 4.0 equiv.) the mixture was added to the resin, shaken, and allowed to stand at room temperature (1.5 h). Resin was washed (DMF, 4 x 14 ml) and deprotection performed with 20% (v/v) piperidine/DMF (14 ml, 20 min). Coupling of the 4-hydroxymethylbenzoic acid linker (HMBA, 0.61 g, 4.0 mmol., 4.0 equiv.) was then carried out as described for Fmoc-Gly-OH (2 h). After washing with DMF (4 x 14 ml), the resin was sequentially washed with dry MeCN and dry DCM (both 5 x 14 ml). The first amino acid of the peptide sequence; Fmoc-Gly-OH (1.2 g, 4.0 mmol., 4.0 equiv.), was coupled in dry DCM (14 ml) using 1-Methylimidazole (1-Melm, 370 μl , 4.6 mmol., 4.6 equiv.) and 1-(Mesitylene-

2-sulfonyl)-3-nitro-1H-1,2,4-triazole (MSNT, 1.2 g, 4.0 mmol., 4.0 equiv.). After one hour, resin was washed with dry DCM (4 x 14 ml) and the coupling step repeated a second time. This was followed by deprotection with 20% (v/v) piperidine/DMF (5 min) and washing with DMF (4 x 14 ml). The resin was transferred to the custom 20-well Teflon synthesis block and evenly re-distributed across wells with DMF (14 ml). Amino acids were then coupled with TBTU for one hour up to the first six residues, after which couplings were performed for two hours using the same conditions and equivalents as for Fmoc-Gly-OH (0.05 mmol. scale/well). The only exceptions were non-canonical amino acids (which were all coupled for 3 h) and Fmoc-Lys(*retro*-Abz-Boc)-OH, which was coupled using 3.0 equiv. of amino acid, 3.0 equiv. of TBTU and 3.9 equiv. of NEM. Completion of coupling reactions and successful deprotections were confirmed using the Kaiser test (181) (section 2.3.4) with un-modified PEGA₁₉₀₀ resin acting as a positive control. Final Fmoc deprotection was carried out twice using 20% (v/v) piperidine/DMF (14 ml, 1 x 2 min and 1 x 20 min) followed by washing with DMF (6 x 14 ml) and drying under low vacuum (0.5 h). Protecting groups were removed using a cleavage cocktail consisting of 95% TFA: 2.5% triisopropylsilane: 2.5% water (14 ml). After initial cleavage (10 min), solvent was removed and a second cleavage performed under the same conditions (14 ml, 3 h). After removal of the second cleavage cocktail, the resin was sequentially washed with DCM (2 x 14 ml), DMF (2 x 14 ml), 5% (v/v) DIPEA/DMF (5 x 14 ml), DMF (5 x 14 ml), DCM (5 x 14 ml), MilliQ water (3 x 14 ml), TNC buffer (3 x 14 ml) and then dried under low vacuum (1 h) and air-dried overnight. The resultant resin-bound peptides were taken forward to test for cleavage by TS5-5 (section 2.3.7).

2.3.4. Kaiser Test

To a small number of resin beads (< 1 mg) was sequentially added 2–3 drops of phenol: ethanol (80:20, w/v), 1 mM potassium cyanide: pyridine (2:98, v/v) and 5% (w/v)

ninhydrin in ethanol. The beads were heated to 110-120 °C and the colour change monitored visually. The presence of free amine groups was indicated by dark blue beads, whereas clear beads indicated that successful coupling had occurred (181).

2.3.5. Cleavage of peptide substrates from PEGA₁₉₀₀ resin

To release peptide substrates from PEGA₁₉₀₀ resin, the base-labile HMBA linker was cleaved using 0.1 M NaOH (17 x 1 ml, 1 h). This solution was transferred to a TELOS column containing 0.1 M HCl (17 x 1 ml). The subsequent neutralised solution was diluted 65-fold with MilliQ water or MeCN and an analytical UHPLC run performed to determine peptide purity (see Chapter 4, Table 4.4). Accurate peptide mass was then ascertained *via* MALDI-Time-Of-Flight Mass Spectrometry (MALDI-TOF-MS, section 2.3.6).

2.3.6. MALDI-TOF-MS of peptide substrates

A solution of α -cyano-4-hydroxycinnamic acid (1 μ l, 10 mg/ml) in 70% (v/v) MeCN/water was spotted onto a MTP 384 polished steel target plate (Bruker, Coventry, UK). Peptide dissolved in water/MeCN (1 μ l) from the UHPLC analytical sample was also added. After sample crystallisation, the plate was read in a Bruker solariX XR instrument and data analysed using Bruker Compass DataAnalysis 5.1.

2.3.7. On-resin screening of modelled peptide substrate against TS5-5

For each peptide substrate to be assayed, TNC buffer containing 0.005% (v/v) Brij-35 (50 μ l) was added to wells in a 96-well glass-bottomed black microplate (PerkinElmer, Seer Green, UK) in single replicate. A background control well with the same volume was also prepared. PEGA₁₉₀₀ beads containing a bound peptide substrate from section 2.3.3 were added to wells, with unbound PEGA₁₉₀₀ starting material beads added to the control well (between 10–30 beads were used in each case). All wells were incubated at 37 °C (15 min). The initial fluorescence intensity of one bead from each substrate was measured using an

IX73 inverted fluorescence microscope with an XC10 magnifier (Olympus, Southend-on-Sea, UK) under UPlanSAPO 10x magnification (30 ms exposure) and subtracted from the fluorescence intensity of the background solution. The beads were then imaged using cellSens dimension 1.13 software (Olympus, Southend-on-Sea, UK). After further incubation (37 °C, 15–30 min), TS5-5 (50 µl, 1 µM) in TNC buffer with 0.005% (v/v) Brij-35 was dispensed into all wells to give a final concentration of 500 nM. The microplate was incubated again at 37 °C and the fluorescence intensity of the beads measured at 1 h and 24 h time points as before.

2.3.8. Synthesis of a combinatorial library on PEGA₁₉₀₀ resin

To enable the rapid synthesis of a combinatorial library of peptide substrates using the split-mix method, a custom-made 20-well Teflon synthesis block was used (180). Initially however, uniformly sized PEGA₁₉₀₀ resin beads in water (18 ml) (182) were swollen in a TELOS column and solvent removed. The resin beads were weighed (8.28 g, 1.0 mmol., 0.1 mmol/g loading), washed (4 x 20 ml) and swelled in DMF (10 min, 20 ml). After removal of solvent, the resin was deprotected using 20% (v/v) piperidine/DMF (20 ml, 20 min) and washed with DMF (5 x 20 ml). Synthesis was then performed as described in section 2.3.3, with the following differences: i) the library was mixed through vigorous manual shaking (2 x 20 s) after the completion of each coupling step (and prior to Fmoc-deprotection) to ensure that each well contained a mixture of all coupled amino acids; and ii) the coupling of Fmoc-Lys(*retro*-Abz-Boc)-OH was performed overnight using 3.0 equiv. of amino acid, 3.0 equiv. of TBTU and 2.93 equiv. of NEM in DMF (14 ml). Removal of N-terminal and side chain protecting groups and subsequent washing steps were performed as described in section 2.3.3. The library was dried under low vacuum (2 h), transferred out of the 20-well Teflon synthesis block into a TELOS column using water and stored at 4 °C overnight.

2.3.9. Incubation of the combinatorial library with TS5-5 and isolation of hits

The combinatorial library was suspended in TNC Brij-35 0.005% (v/v) (31 ml) and beads dispersed through gentle mixing and manual shaking. Library beads (1 ml) were then evenly distributed throughout 31 individual wells across Corning 6-well tissue culture plates (Thermo Fisher Scientific, Loughborough, UK) to give a single layer of ~3500 beads/well. A well containing unbound PEGA₁₉₀₀ starting material beads (1 ml) was used as a negative control. After incubation of the library at 37 °C (30 min), TS5-5 proteinase in TNC Brij-35 0.005% (v/v) was added to each well (1 ml, 1 µM) to give a final concentration of 500 nM and the plates incubated at 37 °C. Library beads were analysed for signs of fluorescence every 30 minutes. After 1 hour, fluorescence was observed for ~80 beads across all wells. Library beads were therefore transferred back into a TELOS column and washed with water (2 x 20 ml). TS5-5 proteinase was deactivated with 2% (v/v) TFA in water (2 x 20 ml), the beads washed (water, 2 x 20 ml and basified with 2% (w/v) NaHCO₃ (2 x 20 ml). The beads were then washed a final time with water (3 x 20 ml) prior to being dried under low vacuum (30 min) and left overnight at 4 °C. The library was re-suspended in water (20 ml) and transferred back to 6-well tissue culture plates (1 ml/well). Beads were analysed under a fluorescent microscope using the same settings as described previously (section 2.3.7). Twenty fluorescent beads were removed manually using a pipette tip (1 ml) and peptides were cleaved from PEGA₁₉₀₀ resin as described in the next section.

2.3.10. Cleavage of hit peptide substrates from PEGA₁₉₀₀ resin

Twenty fluorescent beads were transferred to individual Kinesis LC-MS vials (Cole-Palmer, St. Neots, UK) and each bead treated with a cleavage solution of 5% (v/v) triethylamine (TEA) in water (50 µl, 2.5 h). To evaporate off the cleavage solution, vials were left at room temperature overnight in a fumehood set to an increased flow rate. The beads

were then washed with 70% (v/v) MeCN/water (3 x 20 μ l). An aliquot of each solution was taken (1 μ l) and the sample prepared for MALDI as described in section 2.3.6.

2.3.11. MALDI and tandem mass spectrometry of hit peptide substrates

The MTP 384 polished steel target plate containing hit substrate samples from library screening was read on a Bruker solariX XR instrument. An initial MALDI spectrum was obtained in the 150 – 2500 m/z range using the following parameters: a laser frequency = 200 Hz and collision energy = 14.0 V combined with a raster smart walk pattern. This provided the mass of the quasi-molecular ion; namely the singly protonated intact peptide $[M + H]^+$. Using collision-induced dissociation (CID), the quasi-molecular ion was fragmented with a laser frequency = 200 Hz, collision cell radiofrequency (RF) = 1.6 Vpp and collision energy = 80.0 V using a raster smart walk pattern. This generated a tandem mass spectrum also in the 150 – 2500 m/z range. The data for each hit peptide was collected and analysed using Bruker Compass DataAnalysis 5.1.

2.3.12. Analysis of mass spectra with computational software

Bruker Compass DataAnalysis was used to analyse both MALDI and tandem mass spectra, while LibMSCalc was used to score each peptide hit (183). To begin with, Bruker Compass DataAnalysis established the mass of the quasi-molecular ion from the initial MALDI mass spectrum. This mass was that of the singly protonated intact peptide and is denoted $[M + H]^+$. A subsequent tandem mass spectrum of the quasi-molecular ion (or a closely related high molecular mass ion) generated a series of fragment ions (y ions or b ions), which made up the structure of the final peptide. To identify the ion fragments within each series, a database of amino acids was generated. Natural amino acids were already present, but unnatural amino acids (e.g. β -cyclopropyl-alanine) had to be entered manually by residue mass. To avoid confusion, amino acids that were not used in combinatorial library synthesis were

removed from the database. A tool was then used to identify these fragment ions by assigning residues to the differences in masses between the fragment ion peaks. As the amino acids used in the combinatorial library at each position were known, the residues lost from the fragment ion peaks were pieced together to generate an overall fragmentation pattern, which enabled the determination of the final peptide sequence.

Thereon, LibMSCalc was used to “score” each of the determined sequences. The amino acid building blocks of the combinatorial library were entered into the software with the mass of the quasi-molecular ion. The output from the software was an exhaustive list of all possible peptide sequences that matched the mass of the quasi-molecular ion. In addition, a computed mass list of fragment ions with masses and sequences for each of the possible peptides was also listed. An experimental mass list of fragment ions compiled from the tandem mass spectrum of each peptide hit was imported into LibMSCalc. The software matched the experimental mass list to the computed mass list and provided a score for each peptide based on the number of matching fragment ion peaks and their intensity. The higher the score, the higher the likelihood that there was only one peptide sequence that could match the data. Out of the hits that could be conclusively identified, the sequences listed with the highest score were synthesised by Fmoc-SPPS, and purified and characterised by LC-MS.

Chapter 3. The purification and characterisation of ADAMTS-4 and ADAMTS-5 proteinases

3.1. Introduction

As described in the aims section at the end of Chapter I, the first step towards developing a selective ADAMTS-5 FRET peptide substrate was to acquire pure, fully characterised and active proteinase. ADAMTS-5 was chosen as the proteinase to target due to the fact that it is the major aggrecanase in murine osteoarthritis (117) and also appears to be more important in human disease (119-121). Nonetheless, ADAMTS-4 was also purified and characterised in order to determine whether a final probe could distinguish between these two closely related family members.

Firstly, to acquire these proteinases, it was necessary to culture them in an adequate expression system. There were several different systems available: bacterial, mammalian, insect, yeast, algal and cell-free (184). A discussion on the merits of each is beyond the scope of this thesis. However, given that both ADAMTS-4 and ADAMTS-5 are human proteinases, a mammalian expression system was the logical choice, as it delivers the desired protein with the highest level of native structure, post-translational modifications and activity (184). In this process, stable cell lines were generated through transfection with vectors containing the ADAMTS-4 or ADAMTS-5 sequence alongside an antibiotic resistance gene. To select for a high copy number of each vector (and thus high expressors of each proteinase), cells were treated with a high concentration of an antibiotic specific to the antibiotic resistance gene (184). After culture of these treated cells, the resultant medium (containing released proteinase) was collected. Unfortunately, in the case of the full-length forms of both proteinases, this process is difficult and poor yields are reported in the literature, namely 50 µg/L for ADAMTS-4 and 2 µg/L for ADAMTS-5 (16, 17). This is in contrast to the purification of shorter domain deletion mutants such as the combined catalytic and disintegrin domains of both ADAMTS-4 and ADAMTS-5, where higher yields of 1.8 mg/L and 0.25 mg/L were

obtained, respectively (16, 17). Designing a FRET substrate requires milligram quantities of a proteinase for library screening, assessing cleavage and determining kinetic parameters. This favours the shorter forms of ADAMTS-4 and ADAMTS-5, as they can be acquired more easily and rapidly. Therefore, the combined catalytic and disintegrin domains of ADAMTS-4 (*i.e.* TS4-4) and ADAMTS-5 (*i.e.* TS5-5) were chosen for tissue culture in a mammalian expression system.

Secondly, to acquire pure ADAMTS-4 and ADAMTS-5 proteinases, the collected medium was purified through affinity chromatography. As both ADAMTS-4 and ADAMTS-5 are tagged with FLAG peptide (DYKDDDDK), this process separates the proteinases from non-tagged impurities *via* non-covalent interactions with a stationary phase of anti-FLAG affinity gel (185). Initially, non-tagged impurities were washed out with either glycine buffers (pH 6.0 and pH 5.0) or a high salt wash buffer (pH 7.5) (185). Elution of each proteinase can then be performed either by using excess FLAG peptide, or by using an acidic glycine buffer (pH 3.0) (185). If FLAG peptide elution is employed, gel filtration is necessary to remove the peptide from the fractions containing the ADAMTS proteinase (185). If acidic elution is employed, eluting fractions are neutralised with tris buffer (pH 7.5) (185). Elution under acidic conditions was preferred. This is not only because it involves fewer steps, but also because FLAG peptide is expensive and difficult to synthesise due to the presence of multiple aspartate residues in its peptide sequence (186, 187). After elution, the proteinase fractions were analysed through western blotting and then pooled. To fully characterise the pooled ADAMTS-4 and ADAMTS-5 fractions and detect any impurities, Coomassie and silver staining was carried out. Coomassie staining detects concentration bands in a 50–500 ng linear range, while silver staining covers a 1–10 ng linear range (35). The use of both methods in tandem

ensured that impurities across a range of concentrations were detected and enabled more complete characterisation of the final proteinase than if either method was used in isolation.

Thirdly, to assess the level of activity of the purified ADAMTS-4 and ADAMTS-5 proteinases, they were tested against commercial FRET substrates (62, 63). The presence of active proteinase was indicated by the increase in fluorescence that occurred from hydrolysis of these probes by either of the ADAMTSs. Once activity had been detected, it was important to determine the concentration of active proteinase in the sample under investigation. This is because the active proteinase concentration cannot be determined using a simple absorbance reading at 280 nm (A_{280}) in conjunction with the Beer-Lambert-Bouguer Law (176), as this does not enable differentiation between active and inactive proteinase in the sample. To solve this conundrum, active-site titration of ADAMTS-4 and ADAMTS-5 with an endogenous inhibitor (*i.e.* N-TIMP-3) or a small molecule inhibitor (*i.e.* CT-1746) was used to correct for any variations in the concentration determined by the A_{280} reading. Active-site titrations are discussed in more detail in the results (subsection 3.2.4) and discussion (section 3.3).

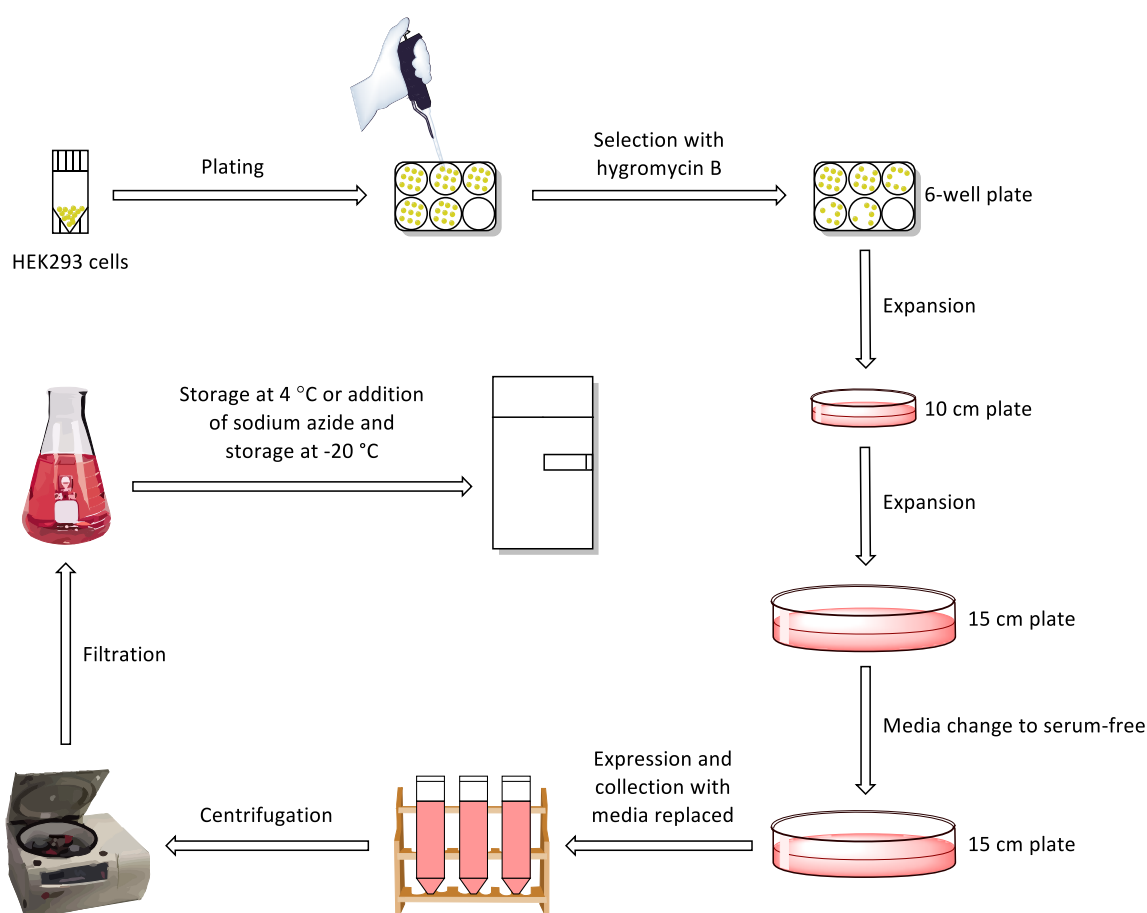
Fourthly, in order to obtain a sufficient quantity of the ADAMTS-5 proteinase for library screening and ADAMTS-4 proteinase for kinetic studies (*i.e.* ≥ 2.0 mg), several batches of medium were collected, purified and characterised.

To summarise, the goal of this chapter is to describe the expression, purification and characterisation (purity and activity) of ADAMTS-4 (*i.e.* TS4-4) and ADAMTS-5 (*i.e.* TS5-5) proteinases, the results thereof and the subsequent discussion.

3.2. Results

3.2.1. Expression and selection of ADAMTS-4 and ADAMTS-5 for purification

A mammalian expression system employing HEK293 cells stably transfected with the pCEP4 vector containing the ADAMTS-4 or ADAMTS-5 sequence was used to obtain ≥ 2.0 mg of proteinase for library screening and kinetic studies. This process is outlined in Scheme 3.1. Cryofrozen cells containing FLAG-tagged TS4-4 or TS5-5 plasmids were cultured in DMEM and treated with different hygromycin B concentrations (50 – 800 $\mu\text{g/ml}$). A final concentration of 100 $\mu\text{g/ml}$ and 800 $\mu\text{g/ml}$ hygromycin B was used for TS4-4 and TS5-5 expressing cells, respectively. A lower concentration was used for TS4-4 expressing cells, as at 800 $\mu\text{g/ml}$ hygromycin B cells were found to detach from plates. Following expansion to larger plates, media was changed to serum-free DMEM augmented with lactalbumin enzymatic hydrolysate to facilitate later purification. The conditioned medium was collected and replaced twice a week for TS4-4 and once a week for TS5-5. Residual cell debris was removed by centrifugation of the conditioned medium, followed by decantation and filtration of the supernatant. The filtered medium was either stored short-term at 4 °C (≤ 24 h) or long-term at -20 °C (≥ 24 h) with the addition of 0.02% (w/v) sodium azide.



Scheme 3.1. Selection of cells, expression of ADAMTS proteinases and collection of conditioned medium for affinity purification. Workflow for obtaining ADAMTS starting material for purification. HEK293 cells containing ADAMTS plasmid were cultured in DMEM on 6-well plates and treated with varying concentrations of hygromycin B (50 – 800 $\mu\text{g}/\text{ml}$) to obtain stable expressors. They were then passaged to 10 cm and thereafter 15 cm tissue culture plates. Once confluence had been achieved, medium was changed to serum-free containing lactalbumin enzymatic hydrolysate. The conditioned ADAMTS medium was collected every 4-7 days prior to being centrifuged and filtered to remove cellular debris. Medium that was not processed within 24 hours was stored at $-20\text{ }^{\circ}\text{C}$ after the addition of 0.02% (w/v) sodium azide.

3.2.2. Purification of ADAMTS-4 and ADAMTS-5 by FLAG affinity chromatography

FLAG-tagged TS4-4 or TS5-5 proteinase was purified from conditioned medium through stepwise pH elution using an anti-FLAG agarose column on an ÄKTA prime FPLC. Figure 3.1A shows the anti-FLAG western blot for detecting these proteinases and the outcome of these steps for TS5-5. After initial equilibration of the column using 50 mM tris-HCl, pH 7.5, 500 mM NaCl, medium was loaded (LD, lane 2), flow-through collected (FT, lane

3), and the column sequentially washed with 0.2 M glycine-HCl pH 6.0 (Fr 1, lane 4) and pH 5.0 (Fr 5, lane 5) to remove untagged contaminants. The proteinase was eluted with 0.2 M glycine-HCl pH 3.0 (Fr 9-11, lanes 6–8) as two bands, one at ~37 kDa and the other at ~41 kDa. The TS5-5 fractions were immediately neutralised with 10% (v/v) 2M tris-HCl, pH 7.5 and frozen at -20 °C to reduce breakdown of active proteinase.

In addition, a small aliquot of each sample was stored at -20 °C for later western blots. This was carried out in order to determine the proteinase-containing fractions that were to be combined. Figure 3.1B shows the western blot of the same process for the catalytic domain of ADAMTS-5 alone (*i.e.* TS5-6), which eluted as a single band at ~31 kDa (Fr 10, lane 7). The difference in molecular weight between TS5-5 and TS5-6 confirms that distinct domain deletion mutants of ADAMTS-5 were purified. A silver stain was used to detect contaminants and determine the effectiveness of the washing steps. Figure 3.1C displays a representative silver stain for TS5-5. Medium was loaded (LD, lane 2), flow-through collected (FT, lane 3) and the column washed with 0.2 M glycine-HCl pH 6.0 (Fr 1 and 3, lanes 4–5) and pH 5.0 (Fr 5 and 7, lanes 6–7). Proteinase was eluted at pH 3.0 (Fr 9, 11 and 13, lanes 8–10). The efficacy of the washing steps for the TS5-5 proteinase at pH 6.0 and pH 5.0 (Fr 1, 3, 5 and 7, lanes 4–7) is apparent, in that contaminants in the load fraction (lane 2) are washed out before elution of pure proteinase (Fr 13, lane 10).

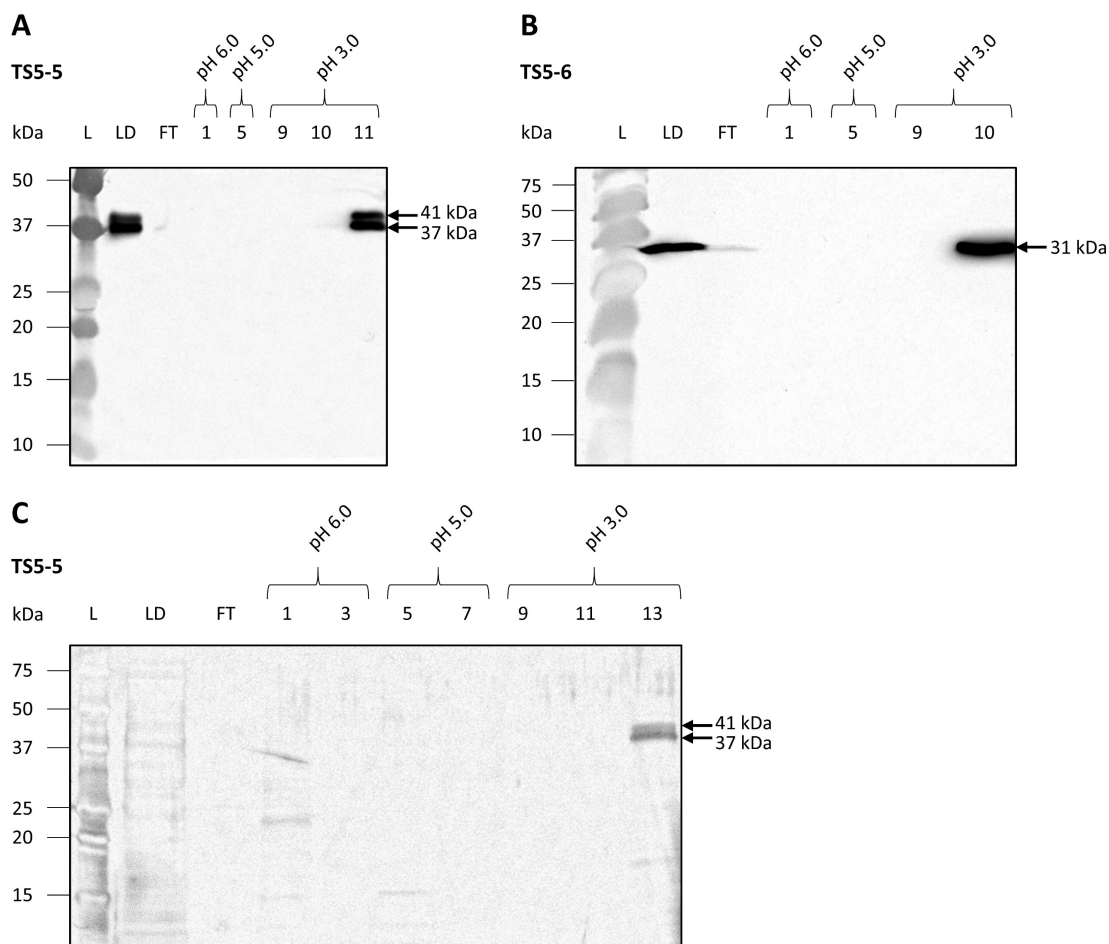


Figure 3.1. Representative affinity purification of ADAMTS proteinases using stepwise pH elution shown for ADAMTS-5. Western blots of eluted (A) TS5-5 and (B) TS5-6 proteinase fractions of different pH detected by mouse monoclonal anti-FLAG M2 alkaline phosphatase antibody. ADAMTS5-FLAG (TS5-5 and TS5-6) was detected at pH 3.0. (C) Silver stain of eluted odd-numbered TS5-5 fractions of different pH. Contaminants were detected at pH 6.0 and pH 5.0 and were washed out prior to elution of proteinase at pH 3.0. L = ladder, LD = load and FT = flow-through.

After western blotting of individual ADAMTS-4 and ADAMTS-5 fractions, samples were thawed, pooled and concentrated in a protein spin concentrator. An A_{280} reading was used to estimate the concentration of proteinase using the Beer-Lambert-Bouguer Law with the suitable molar absorption coefficient. Samples were frozen down in TNC buffer with 0.05% (v/v) Brij-35, 10% (v/v) glycerol and 0.02% (w/v) sodium azide for long-term storage at -80°C . The characterisation data for pooled ADAMTS-4 and ADAMTS-5 fractions is

summarised in Figure 3.2, which shows representative western blots (A), Coomassie stains (B) and silver stains (C) of purified TS4-4 and TS5-5.

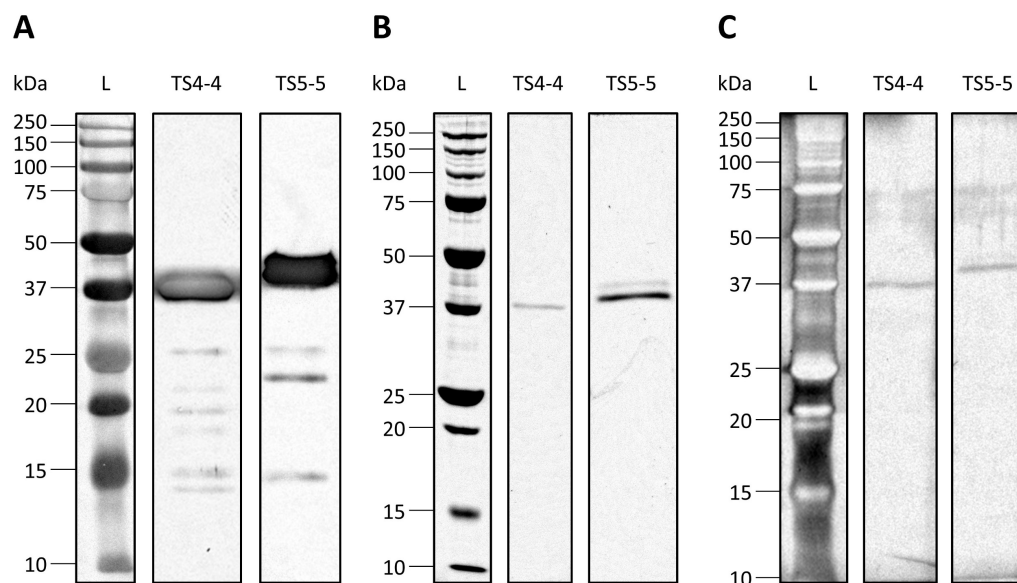


Figure 3.2. Assessment of purity of pooled fractions of ADAMTS proteinases by western blot, Coomassie stain and silver stain. Western blots (A), Coomassie stains (B) and silver stains (C) of pooled ADAMTS-4 (TS4-4) and ADAMTS-5 (TS5-5) proteinase fractions after affinity purification. For western blots and Coomassie stains, pooled fractions were loaded at 1.3 μ g for TS4-4 and 4.5 μ g for TS5-5. For silver stains, pooled fractions were loaded at 0.1 μ g for TS4-4 and TS5-5. L = ladder.

3.2.3. Validation of ADAMTS-4 and ADAMTS-5 activities using FRET substrates

Once the purity of the pooled proteinases had been established, the amount of activity was measured using the commercially available FRET peptide substrates 5-FAM-AE~LQGRPISIAK-TAMRA for ADAMTS-4 and Abz-TESE~SRGAIY-Dpa-KK-NH₂ for ADAMTS-5 (in both cases, the cleavage site is designated with a “~”). Figure 3.3 shows data for the proteolysis of each substrate by different concentrations of TS4-4 (A) and TS5-5 (B) fitted by linear regression. The rate (v) is expressed in terms of relative fluorescent units (RFU) per time (s). In both graphs, an increase in substrate cleavage can be seen with increasing proteinase concentration, an outcome which is dose dependent. For example, as the concentration of TS4-4 proteinase doubles from 5 nM to 10 nM, a doubling of the rate can

be observed (0.036 RFU s^{-1} to 0.074 RFU s^{-1}). This can also be seen for TS5-5, where an increase from 20 nM to 40 nM leads to an increase in rate from 0.022 RFU s^{-1} to 0.039 RFU s^{-1} . These linear changes in reaction velocity with increasing enzyme concentration demonstrate that both enzymes are active and can be used for further experimental studies.

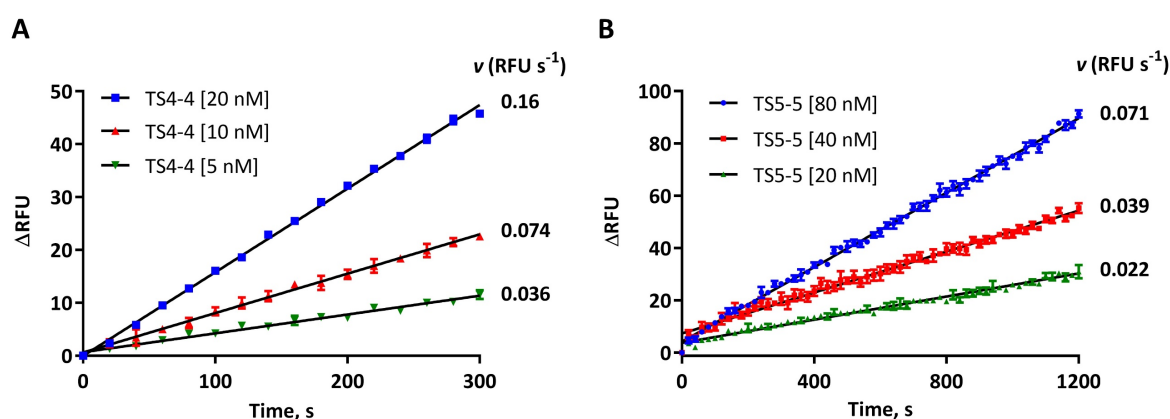


Figure 3.3. Dose-dependent cleavage of FRET substrates by purified ADAMTS-4 and ADAMTS-5 proteinases. (A) Rate data for cleavage of 5-FAM-AE~LQGRPISIAK-TAMRA ($0.5 \mu\text{M}$) by different concentrations of ADAMTS4-4 fitted by linear regression; (B) rate data for cleavage of Abz-TESE~SRGAIY-Dpa-KK-NH₂ ($20 \mu\text{M}$) by different concentrations of ADAMTS5-5 fitted by linear regression. The rate (v) is given in RFU s^{-1} and exhibits a linear increase with each rise in enzyme concentration. Each time point is expressed as ΔRFU ($\text{RFU at time } t - \text{RFU at time } 0$) $\pm$ SEM. For each proteinase concentration, a single experiment with $n = 3$ wells was performed at 37°C . Adapted by permission from Springer Nature: *ADAMTS Proteases: Methods and Protocols, Methods in Molecular Biology*, vol. 2043, pp 75-91, Purification and Activity Determination of ADAMTS-4 and ADAMTS-5 and Their Domain Deleted Mutants by Milan M. Fowkes and Ngee H. Lim © 2020; https://link.springer.com/protocol/10.1007%2F978-1-4939-9698-8_7.

3.2.4. Determination of ADAMTS-4 and ADAMTS-5 concentrations by active-site titration with N-TIMP-3

Having confirmed that both ADAMTS-4 and ADAMTS-5 proteinases were active, the concentration of purified TS4-4 and TS5-5 was determined through active-site titration using an endogenous inhibitor. This process involved measuring the rate of substrate cleavage by each proteinase under different inhibitor concentrations and expressing this as a percentage of substrate cleavage in the absence of inhibitor. A proteinase concentration approximately

ten-fold higher than that of the inhibition constant (K_i) of the inhibitor was utilised, and the resultant data was fitted to the equation for tight-binding inhibition (177). The inhibitor used was N-TIMP-3, with a known concentration and K_i (62). Figure 3.4 displays the fitted curves that resulted from active-site titration of TS4-4 (10 nM, red curve) and TS5-5 (27 nM, blue curve) with N-TIMP-3.

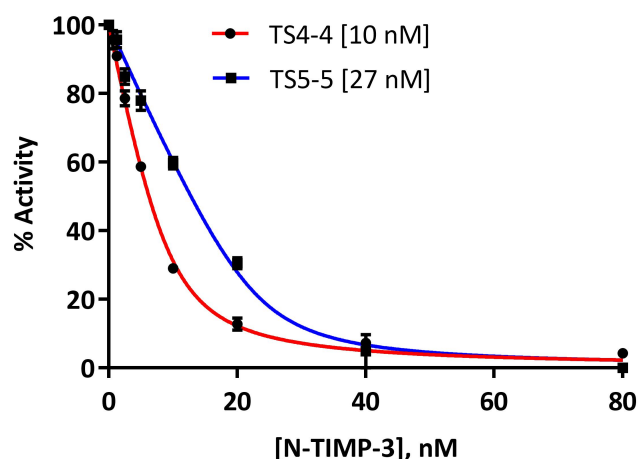


Figure 3.4. Determination of ADAMTS-4 and ADAMTS-5 concentrations through active-site titration with an endogenous TIMP-3 inhibitor. Representative curves for active-site titration of TS4-4 (red) and TS5-5 (blue) with N-TIMP-3. Rates are expressed as a function of percentage activity without inhibitor using the 5-FAM-AE~LQGRPISIAK-TAMRA (0.5 μ M) and Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μ M) substrates for TS4-4 and TS5-5, respectively. Values are given as mean percentage activity $\pm$ SEM. For each proteinase, $n = 3$ wells were used for each N-TIMP-3 concentration at 37 °C.

Extrapolation of the linear segment of each curve to the x-axis delivered the equivalence point and thus the concentration of active proteinase (85, 177). This process always revealed proteinase concentrations higher than those estimated through the initial absorbance measurement (A_{280}). As a consequence of this, the titration was repeated at least once more with the most recent result to validate the final concentration of TS4-4 and TS5-5.

3.2.5. Determination of CT-1746 inhibition constants for ADAMTS-4 and ADAMTS-5

Active-site titration with a small molecule inhibitor (CT-1746) was chosen to further substantiate the final concentration of TS4-4 and TS5-5. However, this inhibitor had an unknown K_i value for both proteinases. This made finding the optimal proteinase concentration for active-site titration difficult. As a result of this, the inhibition constant of CT-1746 for TS4-4 and TS5-5 was determined first. In order to obtain an accurate K_i value, a proteinase concentration lower than the inhibition constant must be used (177). The K_i was initially assumed to be 10–20 nM, with concentrations adopted below this value. The lower the concentration of proteinase, the more accurate the K_i , although this must be balanced against the ability to measure the rate of substrate cleavage. The lowest proteinase concentration which gave detectable substrate cleavage was found to be 0.4 nM.¹¹ Across several experiments, the linear time period of substrate hydrolysis was found to be up to 8 h for both TS4-4 (red) and TS5-5 (blue) at final concentrations of 0.5 nM (Figure 3.5). Having established a minimum concentration and experimental time window, the actual K_i determinations were performed. Data was fitted to the equation for tight-binding inhibition and yielded the representative curves for TS4-4 (red) and TS5-5 (blue) shown in Figure 3.6.

¹¹ This value was determined by Dr. Linda Troeberg with the laboratory microplate reader using ADAMTS-4 and ADAMTS-5 proteinases with the 5-FAM-AE~LQGRPISIAK-TAMRA (0.5 μ M) and Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μ M) substrates, respectively.

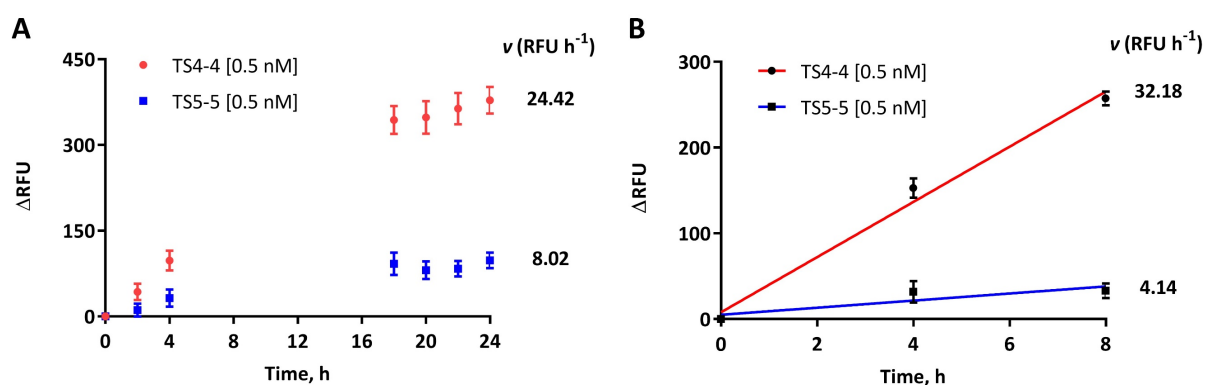


Figure 3.5. FRET substrate hydrolysis by ADAMTS-4 and ADAMTS-5 proteinases over 24 hours. Representative data for cleavage of the 5-FAM-AE~LQGRPISIAK-TAMRA substrate (0.5 μ M, red) by TS4-4 (0.5 nM) and the Abz-TESE~SRGAIY-Dpa-KK-NH₂ substrate (20 μ M, blue) by TS5-5 (0.5 nM) over 24 h (A) and 8 h (B). The rate (v) is given in RFU h^{-1} for 0–4 h (A) and 0–8 h (B). The rate at each time point is expressed as Δ RFU (RFU at time t – RFU at time 0) $\pm$ SEM for $n = 6$ wells from two experiments performed in triplicate (A); and $n = 3$ wells from one experiment performed in triplicate (B) both at 37 °C.

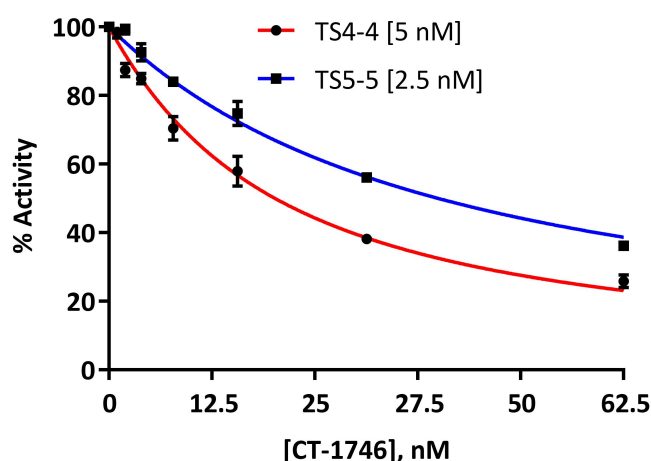


Figure 3.6. Determination of the inhibition constant for a small molecule inhibitor of ADAMTS-4 and ADAMTS-5. Representative curves for determination of the apparent inhibition constant $K_{i(app)}$ for CT-1746 inhibition of TS4-4 (red) and TS5-5 (blue). Rates are expressed as a function of percentage activity without inhibitor using the 5-FAM-AE~LQGRPISIAK-TAMRA (0.5 μ M) and Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μ M) substrates for TS4-4 and TS5-5, respectively. Values are given as mean activity $\pm$ SEM. For each proteinase, $n = 3$ wells were used for each CT-1746 concentration at 37 °C.

These experiments were repeated several times using proteinases stored under disparate temperatures (-20 °C or -80 °C) using different batches, concentrations, and incubation times. This was to ensure that a reliable $K_{i(app)}$ value was obtained. It is important

to note that provided the proteinase concentration is below the $K_{i(app)}$ and the activity can be measured accurately, proteinases stored at temperatures above -80 °C may be used. This is because a decrease in concentration at these temperatures will not considerably affect the $K_{i(app)}$.

However, the apparent inhibition constant $K_{i(app)}$ does not represent the actual K_i value of the inhibitor for the proteinase. This is because it does not take into account the concentration of the substrate or its Michaelis constant (K_m). The actual K_i was therefore determined using equation 2, which takes these parameters into account (see materials and methods section 2.1.2.17) (62). A summary of the $K_{i(app)}$ and K_i values acquired for CT-1746 inhibition of TS4-4 and TS5-5 is shown in Table 3.1 (see Appendix Figure 8.1 and 8.2 for curves). The average K_i value was 16.06 ± 1.71 nM (mean $\pm$ SEM, $n = 5$ individual experiments) for TS4-4 and 20.02 ± 1.21 nM (mean $\pm$ SEM, $n = 3$ individual experiments) for TS5-5. The higher K_i value observed for entry F (TS5-5) could be due to the proteinase concentration (15 nM) being closer to $K_{i(app)}$ than for entries G (10 nM) and H (5 nM), thereby leading to a less accurate value.

The importance of determining the inhibition constant within the linear range of substrate hydrolysis is shown by $K_{i(app)}$ values obtained for CT-1746 inhibition of TS4-4 when incubation is performed for 18–24 h (Table 3.2). Outside the linear range of substrate hydrolysis, $K_{i(app)}$ was determined to be 7.58 ± 0.85 nM (mean $\pm$ SEM, $n = 7$ individual experiments), a value that considerably differs from that obtained within the linear range of substrate hydrolysis: 16.59 ± 1.77 nM (mean $\pm$ SEM, $n = 5$ individual experiments).

Table 3.1. Summary of inhibition constants determined for CT-1746 inhibition of ADAMTS-4 and ADAMTS-5.

ADAMTS proteinase	Entry	Final proteinase concentration, nM	Apparent inhibition constant, $K_{i(app)}$, nM	Inhibition constant, K_i , nM
TS4-4	A	5	17.69 ± 0.23	17.12 ± 0.23
	B	5	15.71 ± 0.29	15.21 ± 0.28
	C	1	22.72 ± 0.04	21.99 ± 0.04
	D	0.8	10.36 ± 0.10	10.03 ± 0.10
	E	0.4	16.46 ± 0.02	15.93 ± 0.02
			16.59 ± 1.77	16.06 ± 1.71
TS5-5	F	15	28.96 ± 0.69	22.98 ± 0.55
	G	10	23.28 ± 0.63	18.48 ± 0.50
	H	2.5	23.46 ± 0.53	18.62 ± 0.42
			25.23 ± 1.52	20.02 ± 1.21

Table 3.2. Summary of inhibition constants determined for CT-1746 inhibition of ADAMTS-4 outside the linear range of substrate hydrolysis.

ADAMTS proteinase	Final proteinase concentration, nM	Apparent inhibition constant, $K_{i(app)}$, nM
TS4-4	0.55	4.21 ± 0.03
	0.97	11.94 ± 0.11
	0.97	5.82 ± 0.04
	0.75	7.59 ± 0.03
	1.00	8.58 ± 0.03
	0.60	6.77 ± 0.06
	0.60	8.17 ± 0.02
		7.58 ± 0.85

3.2.6. Validation of ADAMTS-4 and ADAMTS-5 concentrations by active-site titration with CT-1746

Having determined the K_i values for CT-1746 inhibition of TS4-4 and TS5-5 (Table 3.1), a sample of each proteinase was actively-site titrated with N-TIMP-3 and CT-1746 to confirm whether similar values would be obtained (Table 3.3, see Appendix Figure 8.3 for curves). Active-site titration with CT-1746 and N-TIMP-3 yielded values of 2.88 μ M and 4.81 μ M for TS4-4 and 2.05 μ M and 2.52 μ M for TS5-5, respectively. This data shows that titration with CT-1746 gives a concentration with a confidence interval that overlaps with that of the endogenous inhibitor (cf. 2.18, 4.37 to 3.50, 8.14 for TS4-4; and 1.66, 2.72 to 2.33, 2.76 for TS5-5). This demonstrates that a small-molecule inhibitor with a known K_i value can be successfully employed for active-site titration of proteinases with unknown concentrations.

Table 3.3. Summary of active ADAMTS-4 and ADAMTS-5 concentrations obtained by active-site titration with endogenous and small molecule inhibitors.

Proteinase	Inhibitor	Concentration found, μ M	95% Confidence interval	Inhibition constant, nM
TS4-4	N-TIMP-3	4.81	[3.50, 8.14]	$0.64 \pm 0.15^*$
	CT-1746	2.88	[2.18, 4.37]	16.06 ± 1.71
TS5-5	N-TIMP-3	2.52	[2.33, 2.76]	$1.01 \pm 0.45^*$
	CT-1746	2.05	[1.66, 2.72]	20.02 ± 1.21

* Literature values from L. Troeberg *et al Matrix Biol.* **2009**, 28, 463-469.

3.2.7. Summary of purified amounts of ADAMTS-4 and ADAMTS-5 proteinases

Finally, the amount of TS4-4 and TS5-5 obtained from each purified batch was calculated using the molecular weight of each proteinase. A summary of these quantities (determined through active-site titration with N-TIMP-3) is displayed in Table 3.4 (see Appendix Figure 8.4 for curves).

Table 3.4. Summary of purified amounts of ADAMTS-4 and ADAMTS-5 proteinases determined through active-site titration with N-TIMP-3.

ADAMTS proteinase	Purification batch	Starting material, L	Amount obtained, mg	Purification efficiency, mg/L	Active concentration (%)
TS4-4	1	0.5	0.7	1.4	100
	3	0.5	2.1	4.2	100
	4	0.5	0.6	1.2	100
	5	0.5	2.6	5.2	100
	6	0.5	2.0	4.0	100
	Total	2.5	8.0	3.2*	100
TS5-5	4	0.5	7.5	15	100

* Average purification efficiency

3.3. Discussion

The aim of this chapter was to obtain pure, fully characterised and active ADAMTS-4 and ADAMTS-5 proteinases in sufficient quantities (≥ 2.0 mg) for use in further studies. This process began with the expression and selection of TS4-4 and TS5-5 domain deletion mutants and continued with their purification through affinity chromatography. They were then characterised by western blotting as well as Coomassie and silver staining. The concentration of each proteinase was determined through active-site titration with N-TIMP-3, and further confirmed through active-site titration with the small molecule inhibitor CT-1746.

To begin with, FLAG-tagged ADAMTS-4 and ADAMTS-5 plasmids were separately cultured in mammalian HEK293 cells. As described in the introduction to this chapter, the combined catalytic and disintegrin domain of ADAMTS-4 (*i.e.* TS4-4) and ADAMTS-5 (*i.e.* TS5-5) were chosen as they had amongst the highest yields after purification. In addition, the molecular weight of TS5-5 (*i.e.* ~ 34 kDa) (17) is compatible with the PEGA beads that will form the solid support in the ADAMTS-5 combinatorial library in Chapter 4 (155). This

compatibility refers to the molecular weight limit of the proteinase that can diffuse into the PEGA bead to cleave the substrate on the solid support, and has been reported as ~120 kDa (156). A mammalian expression system was chosen instead of a bacterial one such as *E.coli*. This is because thiol-disulphide oxidoreductases such as disulphide-bond forming oxidase A have a propensity to oxidise cysteines successively, thereby leading to improper folding of proteins where non-consecutive disulphide bridges are necessary (188, 189). Looking at the available crystal structures of both proteinases, there are a total of ten overlapping disulphide bonds in the case of TS4-4 (2RJP, protein data bank), whereas there are five overlapping disulphide bonds in the case of TS5-5 (2RJQ, protein data bank). These factors make a mammalian expression system more suitable for expression of TS4-4 and TS5-5 proteinases compared to a bacterial one.

The antibiotic hygromycin B was chosen as the selection agent based on the pCEP4 vector used in the transfection of the HEK293 cells (16, 17). The higher the concentration of hygromycin B used, the greater the expression of the proteinase. However, if the concentration is too high—and once selection pressure has been removed—further culture of HEK293 cells will produce excessive amounts of proteinase, leading to detachment of cells from tissue culture plates and subsequent rounding-up (185). Literature concentrations of hygromycin B for TS4-4 (200 µg/ml (16) and TS5-5 (100 µg/ml (17) are thus at or below the lower end of typical working concentrations of 200–500 µg/ml (184). However, in the case of TS4-4, selection with 800 µg/ml hygromycin B was initially performed. Unfortunately, cells were only stable up to three days after conversion to serum-free medium before they rapidly dissociated from tissue culture plates. Notwithstanding, treatment with 100 µg/ml hygromycin B on newly cultured TS4-4 cells resolved this issue, suggesting that an unnecessarily high hygromycin B concentration was initially utilised. In addition, collection of

conditioned media was performed twice a week as opposed to once a week (16), which could reduce the total amount of degraded proteinase prior to collection. However, to identify the specific day at which degradation is occurring, western blotting would need to be performed on aliquots collected on a daily basis.

In contrast to TS4-4, selection of the TS5-5 proteinase with a hygromycin B concentration of 800 µg/ml did not result in detachment of cells after changing to serum-free medium. One possible explanation for this is that the TS5-5 proteinase is less prone to autocatalytic activity than the TS4-4 proteinase, as it is activated extracellularly by proprotein convertases (190). Nonetheless, to reduce further breakdown of either proteinase by cellular debris and undesirable metabolic by-products secreted by cells, the final conditioned medium was centrifuged and filtered. In addition, for long-term storage at -20 °C, sodium azide was added to the medium to prevent the growth of fungal spores and bacteria. These final post-collection steps are essential, as they ensure that the TS4-4/TS5-5 starting material is not contaminated or considerably degraded, which would otherwise reduce the amount of the desired proteinase prior to chromatographic purification.

To acquire pure TS4-4 and TS5-5 proteinases, affinity chromatography was performed. During the purification, the starting material was kept on ice in order to reduce breakdown of the proteinase at room temperature. The pH was decreased from 7.5 to 6.0 and then from 6.0 to 5.0 before elution of proteinase at pH 3.0 (as exemplified in Figure 3.1A for TS5-5). This stepwise drop in pH from 7.5 to 6.0 and 6.0 to 5.0 (shown for TS5-5 in Figure 3.1C) ensured the washing away of protein contaminants bound non-specifically to the stationary phase. As a result of these washing steps, fewer contaminated fractions of proteinase were collected at pH 3.0. Coomassie and silver stains revealed the final purity of the pooled TS4-4 and TS5-5 fractions (Figure 3.2B–C). The two bands at the top of the TS4-4

and TS5-5 silver stains in Figure 3.2C are likely to be keratin-1 (~66 kDa) and keratin-10 (~60 kDa) (191). Despite considerable effort to remove these bands, they continued to appear. They should be considered artifacts arising from atmospheric dust particles (191).

After confirming the activity of the purified proteinases, the amount of active proteinase was assessed through active-site titration using the reversible tight-binding inhibitor N-TIMP-3 (16, 192) (Figure 3.4). One advantage of measuring the active-site concentration of a proteinase is that it is a useful way to characterise some of its molecular properties (193). The determination of TS4-4 and TS5-5 concentrations through an A_{280} reading and the Beer-Lambert-Bouguer Law (176) fail to account for the presence of *inactive* proteinase, which contributes to the absorbance reading and thereby to the final concentration. In our case, TS4-4 and TS5-5 concentrations determined by active-site titration were consistently higher than those determined through an A_{280} reading. The lower concentrations determined through A_{280} are likely due to the impact that variations in buffer, pH, surfactant and ionic strength have on absorbance readings, as well as the use of inadequate blanks prior to sample measurements (194). In addition, the molar absorption coefficient was determined computationally (and not experimentally) using the ProtParam tool (www.expasy.org), which could also explain the discrepancy.

Another advantage of knowing whether *inactive* proteinase was present in each sample is associated with incubation of the final combinatorial library with TS5-5. The presence of *inactive* proteinase could result in reversible inhibition of the substrate due to competition for binders with the *active* form. The effectiveness of this inhibition would depend on the association and dissociation rate of the substrate for the inactive compared to the active form. The greater the overall binding to *inactive* proteinase, the worse this effect would become. This would result in the temporary loss of potential hits, as the *active*

proteinase would be unable to cleave the substrate if the latter was bound to the *inactive* proteinase. The presence of *inactive* proteinase in sample fractions would require adjusting the concentration of proteinase used for library incubation. Fortunately, 100% active fractions of the TS5-5 proteinase were determined through active-site titration (Table 3.4).

N-TIMP-3 has been frequently used for inhibition and active-site titration of ADAM and ADAMTS metalloproteinases (16, 17, 62, 195, 196). However, there are several limitations associated with using this protein as an active-site titrant. Firstly, N-TIMP-3 is expressed and refolded in bacteria, thus increasing the likelihood of misfolded protein being present. Secondly, its concentration may decrease during long-term storage, thereby affecting the accuracy of future titrations. Thirdly, the concentration of N-TIMP-3 was determined after two sequential active-site titrations and an initial A_{280} reading. Specifically, the concentration of N-TIMP-3 was determined by titrating against a known concentration of MMP-3, which had been expressed in bacteria and refolded. The concentration of MMP-3 was determined through active-site titration with a known concentration of TIMP-1, which had been expressed in mammalian cells but had not been refolded. The concentration of TIMP-1 was determined through an A_{280} reading.¹² The fact that each active-site titration would have introduced an error into the final concentration of N-TIMP-3 means that its final concentration would not be as accurate as it could be had it been weighed out. Given the accumulation of these errors, it was decided that a small molecule inhibitor should be used to confirm the concentration of TS4-4 and TS5-5. CT-1746 was chosen because it inhibited the matrix metalloproteinases MMP-2, MMP-3 and MMP-9 with low nanomolar or picomolar K_i values (197, 198). Its concentration could also be calculated accurately through weighing

¹² TIMP-1 was assumed to be 100% active as it did not have to be refolded. The concentration of TIMP-1 was confirmed to be the same as active-site titration of MMP-1/MMP-3 with a weighed amount of the small molecule inhibitor GM6001 (Ngee H. Lim, unpublished data).

and dissolution in the appropriate solvent. However, small molecule inhibitors also suffer from several issues. Firstly, they tend to have higher K_i values and thus a higher concentration of proteinase must be utilised for titration. This may lead to autocatalysis of the proteinase during the initial pre-incubation step with the inhibitor prior to addition of the substrate. Secondly, as they are frequently non-polar compounds, they must be dissolved in 100% DMSO, and must therefore be diluted to levels where the percentage of DMSO does not affect the activity of the proteinase. Despite these potential drawbacks, it is important to note that the use of an alternative class of inhibitor to confirm the concentration of active proteinase delivers a more thorough and accurate assessment of the active site concentration than either method in isolation.

The K_i values for CT-1746 (Table 3.1) showed that it inhibited both TS4-4 and TS5-5 similarly (16.06 ± 1.71 nM and 20.02 ± 1.21 nM, respectively). This is logical, given that the active-site within the crystal structures of the Cat-Dis domains of TS4-4 and TS5-5 have many conserved residues in common and adopt essentially identical conformations (179).

To maintain a stable long-term proteinase concentration, it is important to store it at the appropriate temperature. Many proteinases are therefore stored at either -20 °C or -80 °C. After active-site titration with N-TIMP-3, the initial concentration of an example TS4-4 proteinase batch was 8.65 μ M, while for TS5-5 it was 8.87 μ M. The concentrations of the same sample batches dropped to 4.81 μ M for TS4-4 and 2.52 μ M for TS5-5 after storage at -20 °C for 6 months and 5 months, respectively (Table 3.3, Appendix Figure 8.3). This reveals a 1.8-fold and 3.5-fold drop in TS4-4 and TS5-5 proteinase concentrations during storage at -20 °C. The decrease in concentration is approximately two-fold higher for TS5-5 than TS4-4 over approximately the same period of time and at a similar initial concentration. This could be because only a single batch of TS5-5 proteinase was purified (Table 3.4), so it was thawed

to room temperature and refrozen more frequently for use in kinetic studies than the five TS4-4 proteinase batches, leading to accelerated degradation. This is supported by Gendron *et al*, who noted that a degree of autocatalysis was detected for repeated cycles of thawing and re-freezing of ADAMTS-5 samples (17). Nevertheless, both TS4-4 and TS5-5 proteinases should only be stored at -20 °C for a short period of time. Otherwise, long-term storage at -80 °C is recommended.

The final amounts of TS4-4 and TS5-5 proteinases obtained were higher than those described in the literature (Table 3.4). Treatment with 100 µg/ml hygromycin B on newly cultured TS4-4 cells resulted in an average post-purification yield of 3.2 mg/L from five purification runs, which is approximately two-fold higher than the 1.8 mg/L described by Kashiwagi and co-workers (16). This improvement in yield is probably due to a combination of twice weekly changes of medium as well as a lower hygromycin B concentration selecting for cells with fewer plasmids, which then do not secrete as much TS4-4. This results in fewer cells detaching from tissue culture plates and an increase in the final yield of the proteinase. The yield of the TS5-5 proteinase after purification was 15 mg/L, compared to the 0.25 mg/L described by Gendron *et al* (17). It is important to interpret this number with caution, as it was based on purification from a single batch of conditioned medium, and later batches are likely to deliver a lower yield of TS5-5 as the HEK293 cells become continually destabilised under serum-free conditions.

In summary, the catalytic and disintegrin domain units of ADAMTS-4 (*i.e.* TS4-4) and ADAMTS-5 (*i.e.* TS5-5) were expressed, purified and characterised by western blotting, Coomassie staining and silver staining. The activity of TS4-4 and TS5-5 was confirmed through the cleavage of FRET substrates and their concentrations determined through active-site titration with N-TIMP-3 and CT-1746. A sufficient amount of each proteinase was obtained

for library screening; 8.0 mg for TS4-4 and 7.5 mg for TS5-5. The aim of the next chapter was to design, synthesise and screen an ADAMTS-5 combinatorial library to discover novel FRET peptide substrates.

Chapter 4. The design, synthesis and screening of an ADAMTS-5 combinatorial library

4.1. Introduction

In Chapter 3, the metalloproteinases ADAMTS-4 and ADAMTS-5 were obtained through mammalian tissue culture, purified by affinity chromatography and characterised by western blotting as well as Coomassie and silver staining. The level of activity of the purified proteinases was determined through cleavage of FRET peptide substrates, and active-site titration with N-TIMP-3 and CT-1746 delivered the concentration of each proteinase.

This chapter focuses on the development of a selective ADAMTS-5 FRET peptide substrate, a process which comprised several steps. Initially, the rate of cleavage of the current literature substrate by ADAMTS-5 was optimised by evaluating a range of surfactant concentrations. This was followed by assessing the level of selectivity of this substrate for ADAMTS-5 over a panel of metalloproteinases. For the purposes of comparison, these initial steps were also carried out for ADAMTS-4 and its respective literature substrate. The least selective substrate was found to be that for ADAMTS-5, thereby indicating a greater need for a more selective probe for ADAMTS-5 than ADAMTS-4, irrespective of the role of these proteinases in osteoarthritis. The current ADAMTS-5 literature substrate was thus modelled into the active site of ADAMTS-5 and derivatives with improved binding and selectivity designed through computational modelling. These derivatives were then synthesised by Fmoc-SPPS and tested for cleavage by ADAMTS-5. The most rapidly cleaved substrate formed the basis of a combinatorial library, which was designed by computational modelling, constructed by split-mix synthesis and screened against ADAMTS-5 to deliver novel hit FRET peptide substrates.

To optimise the rate of cleavage of the current ADAMTS-4 (*i.e.* 5-FAM-AE~LQGRPISIAK-TAMRA, compound **1**) and ADAMTS-5 (*i.e.* Abz-TESE~SRGAIY-Dpa-KK-NH₂, compound **2**) literature substrates (62, 63), different concentrations of Brij-35 surfactant

ranging from 0–5% (v/v) were utilised. Brij-35 was chosen as the preferred surfactant because it is frequently used in enzymatic assays of ADAMTS proteinases (17, 62, 199-201).

To determine the level of selectivity of **1** for ADAMTS-4 and **2** for ADAMTS-5 over a panel of available metalloproteinases, the rate of cleavage of each substrate by TS4-4/TS5-5 was assessed relative to MMP-1, -2, -3, -9, -13 and ADAMTS-4, -5.

The least selective ADAMTS-5 literature substrate was docked into the active site of its target TS5-5 proteinase using MOE software. To improve binding of the substrate and enhance selectivity, a series of derivatives were designed through the replacement of residues with natural and un-natural amino acids based on steric effects and electrostatic interactions. Each modification was followed by an energy minimisation step, with molecular dynamics simulations of the final structure used to obtain the optimal pose for each derivative.

The docked substrates were synthesised by manual Fmoc-SPPS and tested for cleavage on-resin by incubation with the target ADAMTS proteinase. To further improve binding and selectivity, the most rapidly cleaved substrate was used to form the foundation of a 16-mer combinatorial library, which was designed by computational modelling. By utilising conservative amino acid changes rather than more radical ones, a fine balance was struck between improving binding and selectivity but not abolishing cleavage entirely. The final library was constructed through split-mix synthesis and was limited to ~20,000 unique peptides to reduce the total number of possible sequences. This ensured that final hit substrates were rapidly and definitively identified by MALDI and tandem mass spectrometry following library screening.

Finally, the combinatorial library was screened against active TS5-5 proteinase to deliver novel hit substrates, which were removed manually.

To summarise, the purpose of this chapter was to develop a more selective ADAMTS-5 FRET peptide substrate through the design, synthesis and screening of an ADAMTS-5 combinatorial peptide library.

4.2. Results

4.2.1 Determination of the optimal surfactant concentration for FRET substrate cleavage by ADAMTS-4 and ADAMTS-5 proteinases

To determine the concentration of Brij-35 surfactant that would give an optimal rate of substrate cleavage by TS4-4 and TS5-5, a series of Brij-35 concentrations from 0 to 5% (v/v) were evaluated. These included 0.02% and 0.05% (v/v) Brij-35 as these values are commonly employed in ADAMTS-4 and ADAMTS-5 activity assays (17, 62, 199-201). In addition, 0%, 0.0005%, 0.005% and 5% (v/v) Brij-35 were also assessed. The rate of substrate **1** cleavage by TS4-4 and substrate **2** cleavage by TS5-5 was measured in RFU s⁻¹ and divided by the concentration of the proteinase used (nM). A calibration curve for the 5-FAM and Abz fluorophores was generated for each Brij-35 concentration (see Appendix Figure 8.5 and 8.6, respectively). These were used to convert each rate from RFU s⁻¹ nM⁻¹ to s⁻¹. Multiplication by the final concentration of each substrate delivered the rate, which was expressed as the mean ± SD in nM s⁻¹ (Figure 4.1).

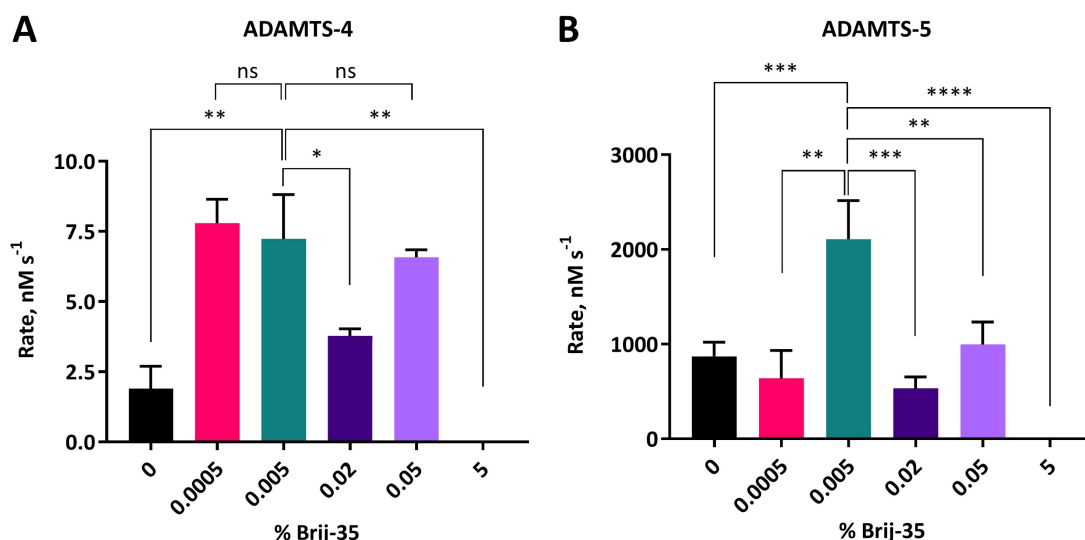


Figure 4.1. The effect of different surfactant concentrations on the rate of cleavage of FRET substrates by ADAMTS-4 and ADAMTS-5. (A) Rate data for cleavage of 5-FAM-AE~LQGRPISIAK-TAMRA (0.5 μ M) by TS4-4 at different concentrations of Brij-35; (B) rate data for cleavage of Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μ M) by TS5-5 at different concentrations of Brij-35. Rates are expressed in nM s⁻¹ as the mean $\pm$ SD for $n = 10$ -27 wells from 3-9 experiments (5-FAM-AE~LQGRPISIAK-TAMRA) or $n = 11$ -21 wells from 3-7 experiments (Abz-TESE~SRGAIY-Dpa-KK-NH₂) performed in triplicate or quadruplicate on different days at 37 °C. Statistical significance was assessed between 0.005% (v/v) Brij-35 and all other surfactant concentrations using a two-tailed parametric un-paired t-test with Welch's correction; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, n.s. not significant.

In the case of TS4-4 (Figure 4.1A), the highest rates of substrate cleavage were observed with 0.0005% and 0.005% (v/v) Brij-35 at 7.80 ± 0.85 nM s⁻¹ and 7.23 ± 1.58 nM s⁻¹, respectively. These values were statistically significant with respect to all other surfactant concentrations except for 0.05% (v/v) Brij-35, where the rate was 6.58 ± 0.27 nM s⁻¹. For TS5-5 (Figure 4.1B), a concentration of 0.005% (v/v) Brij-35 gave the highest rate of substrate cleavage at 2108 ± 409 nM s⁻¹, which was statistically significant relative to all other Brij-35 concentrations.

As a concentration of 0.005% (v/v) Brij-35 resulted in the second highest (*i.e.* with TS4-4) and highest (*i.e.* with TS5-5) rates of overall substrate cleavage, this concentration was chosen for all further assays described in this thesis.

4.2.2 Determination of the selectivity of FRET substrates for ADAMTS-4 and ADAMTS-5 over a panel of metalloproteinases

To determine the selectivity of each substrate for their respective proteinase, the rate of cleavage of **1** by TS4-4 and **2** by TS5-5 was compared to a panel of metalloproteinases during the linear time-window of substrate hydrolysis. This panel included collagenases (MMP-1, -13), gelatinases (MMP-2, -9), stromelysin-1 (MMP-3) and aggrecanases (ADAMTS-4, -5).

For each metalloproteinase, the rate of substrate cleavage was determined in RFU s⁻¹ (see Appendix Figure 8.7 and 8.8 for the raw rate data for each substrate). These values were converted to nM s⁻¹ in the same manner as in the previous section using the 5-FAM and Abz calibration curves from Appendix Figure 8.5C and 8.6C, respectively. Each rate was expressed as the percentage mean $\pm$ SEM as a function of TS4-4 cleavage of **1** or TS5-5 cleavage of **2** (Figure 4.2). The activity of each metalloproteinase was confirmed through activity assays corrected through a calibration curve (Appendix Figure 8.9A–B, respectively). In addition, the concentration of each MMP was corrected through active-site titration with N-TIMP-3 or CT-1746 (Appendix Figure 8.10).

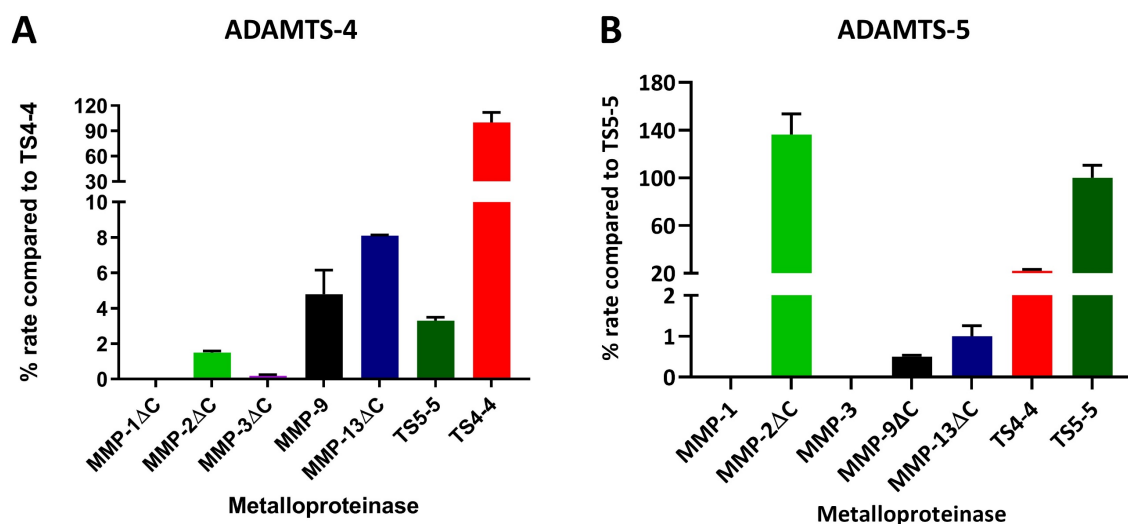


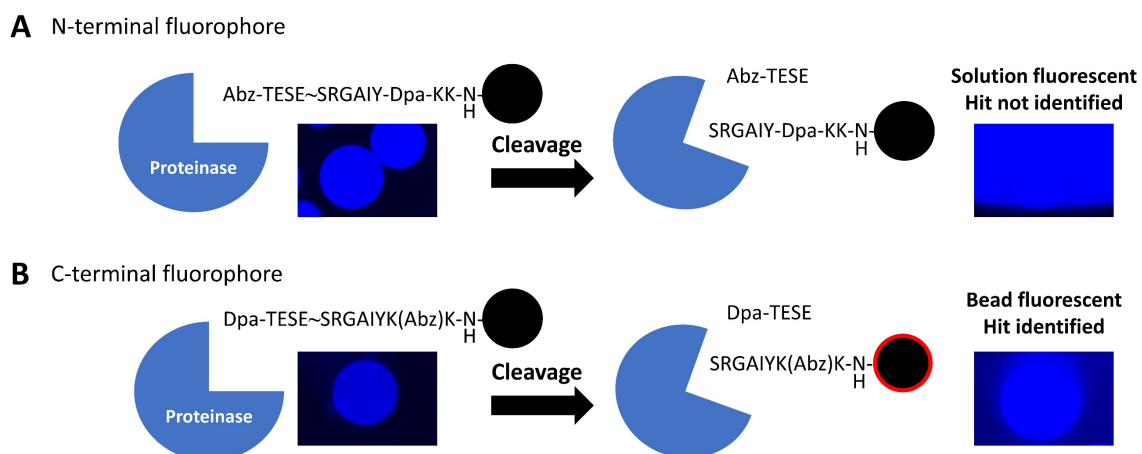
Figure 4.2. Selectivity studies of ADAMTS-4 and ADAMTS-5 FRET substrates with ADAMTS and MMP metalloproteinases. (A) Percentage rate data for cleavage of 5-FAM-AE~LQGRPISIAK-TAMRA (0.5 μ M) by different metalloproteinases. (B) Percentage rate data for cleavage of Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μ M) by different metalloproteinases. The rate of cleavage by each metalloproteinase is expressed as the percentage mean $\pm$ SEM as a function of TS4-4 and TS5-5 cleavage of 5-FAM-AE~LQGRPISIAK-TAMRA and Abz-TESE~SRGAIY-Dpa-KK-NH₂, respectively for $n = 6$ -24 wells from 2-8 independent experiments performed in triplicate at 37 $^{\circ}$ C. Please note that only one experiment with $n = 3$ wells was carried out for MMP-1, as no cleavage was observed up to 72 h at 92 nM. Δ C = catalytic domain, MMP-3 and MMP-9 = full length.

Relative to the rate of **1** cleavage by TS4-4, full length MMP-9 and the catalytic domain (Δ C) of MMP-13 showed cleavage rates $\geq 5\%$ (Figure 4.2A). Specifically, MMP-9 had a cleavage rate = $4.8 \pm 1.4\%$ and MMP-13 Δ C = $8.1 \pm 0.044\%$. All other cleavage rates were $< 5\%$. Substrate **1** thus shows a good level of selectivity for TS4-4, with the rate of cleavage by the tested metalloproteinases not exceeding 10% relative to TS4-4. In contrast, TS5-5 cleavage of **2** (Figure 4.2B) revealed $> 10\%$ cleavage by TS4-4 ($22 \pm 1\%$) and $> 100\%$ cleavage by MMP-2 Δ C ($136 \pm 17\%$). Therefore, substrate **2** showed poorer overall selectivity compared to **1**, with metalloproteinases cleaving at rates $> 10\%$ relative to the ADAMTS proteinase coupled with an extremely high rate of cleavage by MMP-2 Δ C. As a consequence of this, along with the key role of ADAMTS-5 in osteoarthritis, **2** (*i.e.* Abz-TESE~SRGAIY-Dpa-KK-NH₂), was taken forward as the starting point for the development of a more selective FRET

substrate for this proteinase. This stepwise process of: i) design and computational docking; ii) substrate synthesis and testing; and iii) combinatorial library synthesis and screening will be described in more detail in the following sub-sections.

4.2.3 General design and docking of ADAMTS-5 FRET peptide substrates

When designing a more selective ADAMTS-5 substrate, it is important to initially assess whether the structural design of **2** is compatible with combinatorial library screening, which is performed with substrates C-terminally bound to solid-phase PEGA resin beads. Scheme 4.1 shows two possible substrate library designs. In both cases—in the absence of proteinase—the substrate is intact and optically silent, as fluorescence from the Abz fluorophore is quenched by Dpa. However, in the presence of proteinase, cleavage of the substrate occurs, resulting in separation between Abz and Dpa and thereby an increase in fluorescence intensity. In the case of an N-terminal fluorophore design (Scheme 4.1A) some cleaved fragments containing Abz (*i.e.* Abz-TESE) would diffuse out of the bead, resulting in an increase in solution fluorescence. This would make it impossible to identify the hit bead. If a C-terminal fluorophore design is utilised (Scheme 4.1B), the increase in fluorescence would manifest itself in the form a ‘ring’ around the bead. This fluorescent ‘ring’ would gradually expand inwards as the proteinase diffused into the bead and continued to cleave the substrate, thus allowing the bead to be identified as a hit.



Scheme 4.1. Comparison between on-resin screening of combinatorial library designs containing a FRET peptide substrate with an N-terminal or C-terminal fluorophore. Schematic representation of Abz-TESE~SRGAIY-Dpa-KK-NH₂ (A) and Dpa-TESE~SRGAIYK(Abz)K-NH₂ (B) FRET substrates bound to a PEGA resin bead (black circle) *via* the C-terminus. Cleavage of either substrate by the proteinase results in separation of the Abz fluorophore and Dpa quencher, leading to an increase in fluorescence. Only for substrates containing a C-terminal fluorophore (B) can hits be identified, as an N-terminal fluorophore design (A) results in diffusion of the cleaved fragment from the bead into solution following substrate cleavage.

The C-terminal fluorophore design Dpa-TESE~SRGAIYK(Abz)K-NH₂ is thus compatible with combinatorial library screening and was chosen for computational docking of derivatives into the active site of TS5-5 (2RJQ, protein data bank). To inform the docking process, existing substrates for ADAMTS-5 were compiled from CutDB (cutdb.burnham.org), BRENDA (www.brenda-enzymes.org) and MEROPS (www.ebi.ac.uk/merops/) online databases. These substrates were aligned by cleavage site (Appendix Table 8.1) and the most commonly occurring amino acid was identified for the P4-P9' peptide positions (Figure 4.3).

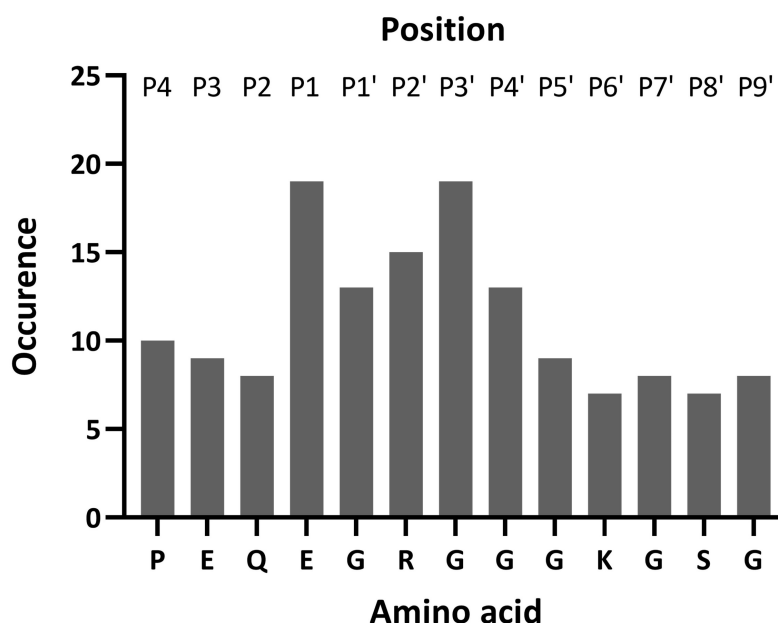


Figure 4.3. The most commonly occurring residues by peptide position for known substrates of **ADAMTS-5**. The most frequently occurring amino acid was determined for P4-P9' positions from the 57 peptide substrates of ADAMTS-5 shown in Appendix Table 8.1.

Detailed analysis of Table 8.1 and Figure 4.3 is beyond the scope of this thesis. However, Table 8.1 indicates the possible positions that could be used for a fluorophore (P3/P4) or quencher (P3' and P5' – P7'). Additionally, Figure 4.3 displays the preferred classes of amino acids that could be docked into the subsites of ADAMTS-5 (*e.g.* E in P1 and G in P3').

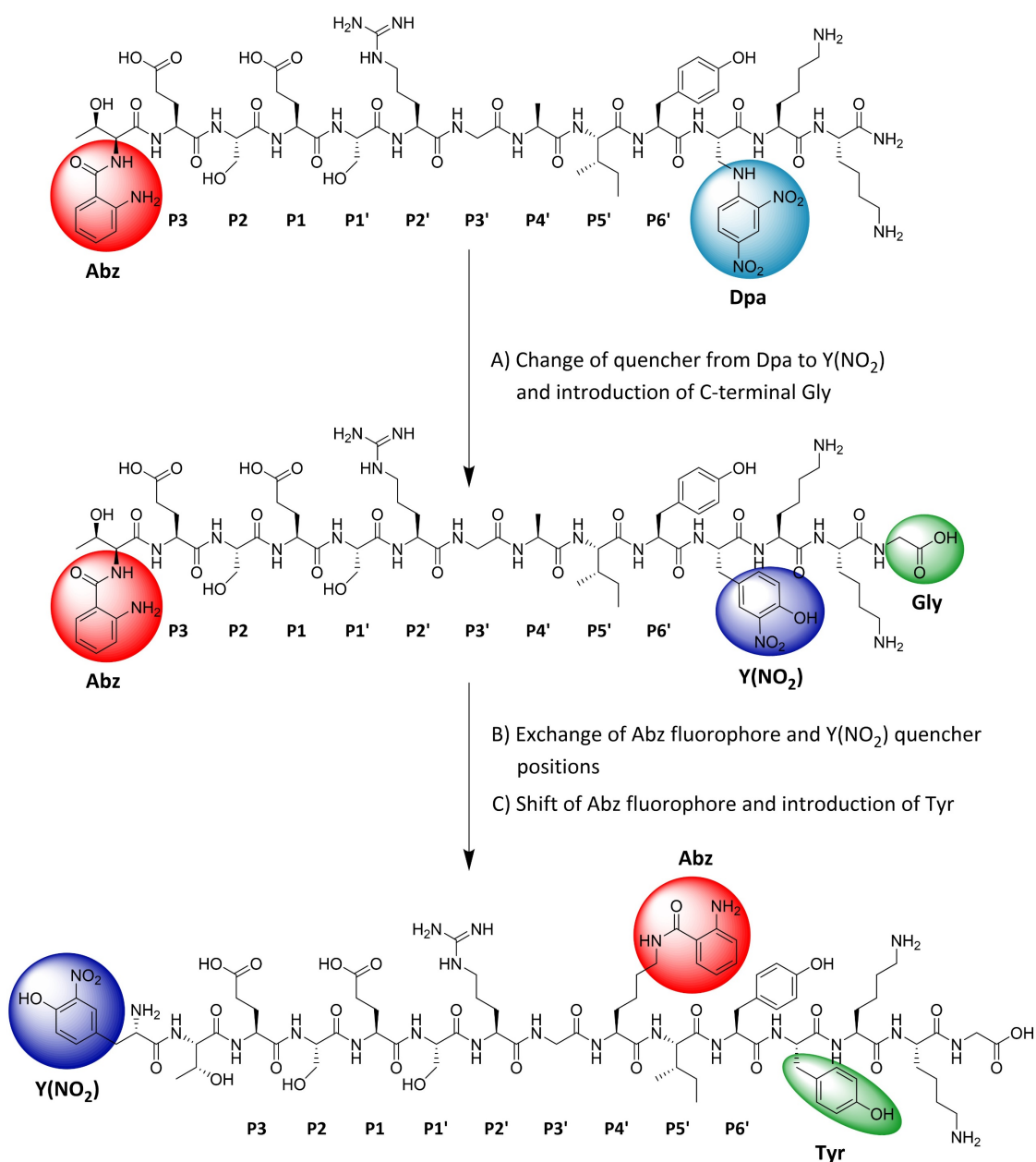
To design substrate derivatives with stronger binding and enhanced selectivity, single, conservative amino acid changes were made to the structure of each substrate. This was carried out to prevent potential disruption to cleavage by ADAMTS-5 through multiple non-conservative modifications. The docking process began with solvation of the substrate-ADAMTS-5 complex in water (see Chapter 2, section 2.2.1 for more details). This was followed by the introduction of amino acids into different P/P' sites based on several criteria: i) the substrate amino acid alignments in Table 8.1; ii) the preference for amino acids in specific positions in Figure 4.3; iii) the presence of favourable or un-favourable electrostatic interactions with the active site; and iv) steric hindrance between the amino acid side chains

and residues within the active site. In addition, un-natural amino acids were used to further explore binding pockets within the active site of the proteinase through non-covalent interactions such as hydrogen bonding or hydrophobic interactions. Energy minimisation of each substrate was performed after introduction of an amino acid. If the pose of the resultant substrate was considered unfavourable based on the criteria from steps iii) – iv), the modification was discarded. If it was considered favourable, a final molecular dynamics simulation was used to generate the optimal substrate pose for the active site of the proteinase.

4.2.4 Design and computational docking of a P4' fluorophore series of ADAMTS-5 FRET peptide substrates

Abz-TESE~SRGAIY-Dpa-KK-NH₂ (**2**) was initially docked into the active site of TS5-5 using MOE software and the initial series of derivatives developed according to Scheme 4.2. The Dpa quencher was exchanged for Y(NO₂) as this is a smaller, innocuous amino acid that is less likely to interfere with binding to the active site (step A). The C-terminal Gly was introduced to facilitate synthesis and fill out the limited space available at the C-terminus. Next, the fluorophore and quencher were exchanged (step B) for reasons described in Scheme 4.1 and the fluorophore shifted to the P4' position onto the side chain of lysine (step C). This position was selected as there is sufficient space available to introduce the fluorophore onto the lysine side chain, with favourable electrostatic interactions between the aromatic amine and the negatively charged pocket of the proteinase. A Tyr was added to the P7' position, as this closely resembles the Dpa quencher in terms of structure and also represents a conservative change. Figure 4.4 shows the current substrate, *i.e.* Abz-TESE~SRGAIY-Dpa-KK-NH₂ (**2**, Figure 4.4A) and the initial P4' substrate, *i.e.* Y(NO₂)TESESRGK(Abz)IYYKKG (**3**, Figure 4.4B) docked into the active site of TS5-5. The position of the Dpa quencher in the initial substrate (Figure 4.4A, green inset) appears to be

unfavourable as a consequence of the electrostatic repulsion between the two aromatic nitro groups and the negatively charged electrostatic molecular energy surface of the active site. This validated the introduction of a Tyr at the P7' position in **3**. A summary of the different FRET peptide substrates containing a P4' fluorophore modelled into TS5-5 are shown in Table 4.1, with a snapshot of each amino acid modification from MOE indicated by a red arrow. The quality of the fit for each docked substrate was assessed through application of the criteria in steps iii) – iv) at the end of sub-section 4.2.3. The quality of the fit cannot be easily quantified through calculations of potential energy, as these include water molecules of the solvation shell and their interactions with the substrate-ADAMTS-5 complex. The substrates in Table 4.1 (and all others in Tables 4.2–4.3 thereafter) are thus listed in the order that the modifications were carried out.



Scheme 4.2. Multi-step design process for the development of an $P4'$ fluorophore series of ADAMTS-5 FRET peptide substrates informed by computational docking. The current substrate Abz-TESE~SRGAIY-Dpa-KK-NH₂ was used as the starting point for the design process. (A) The Dpa quencher at $P7'$ (teal) was changed for $Y(NO_2)$ (dark blue), with a Gly carboxylic acid (green) added to the C-terminus; (B) The $Y(NO_2)$ quencher (dark blue) was swapped with the N-terminal Abz fluorophore (red); (C) The Abz fluorophore (red) was moved to $P4'$, with Tyr (green) added to $P7'$ in its stead.

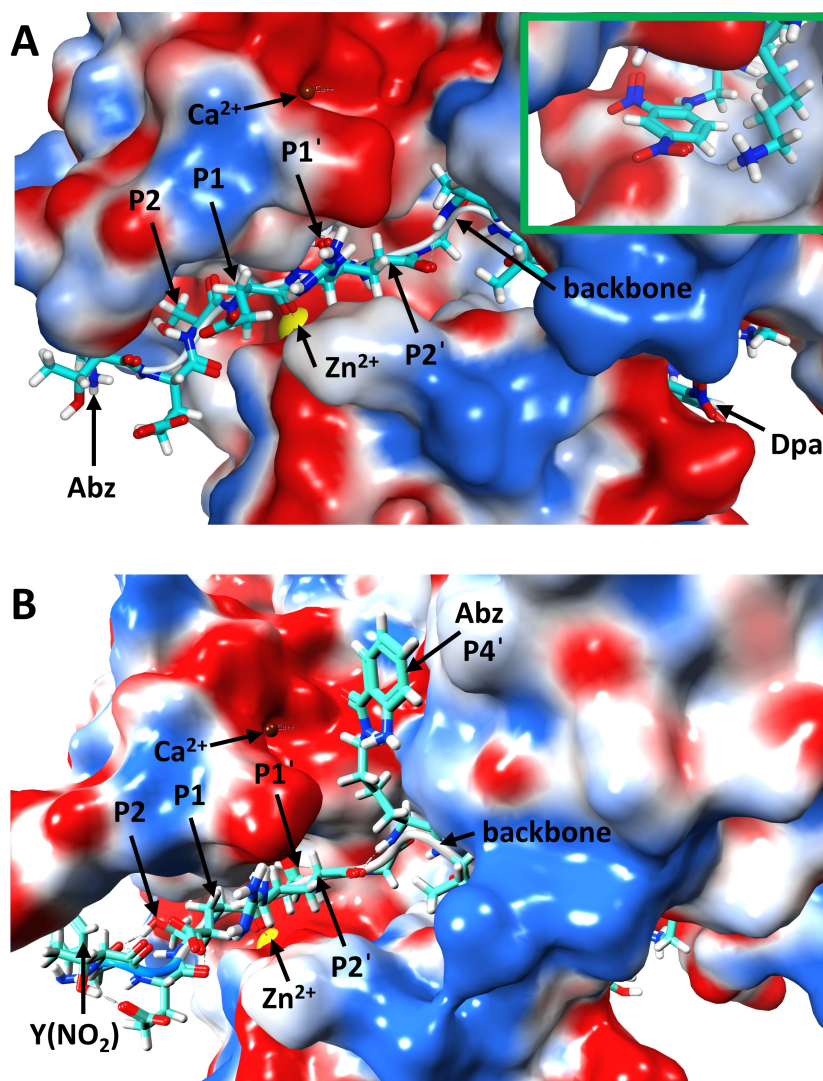
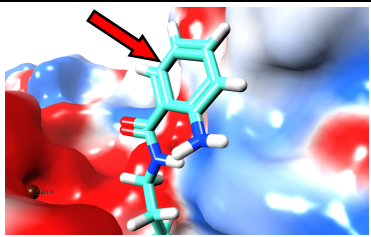
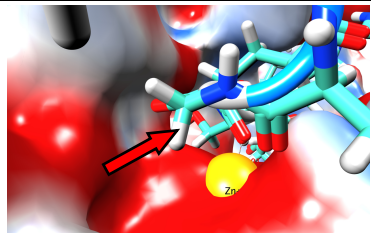
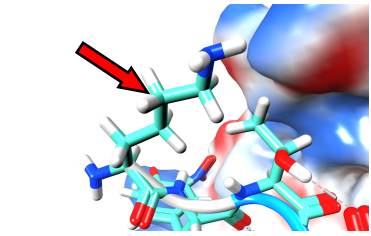
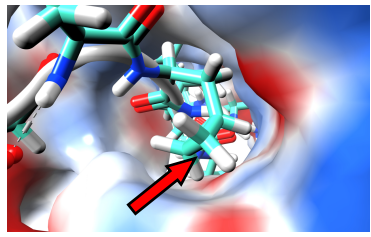
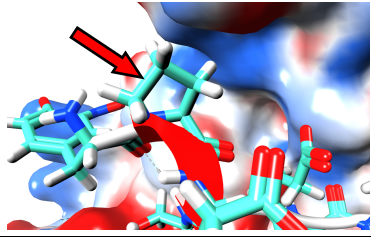
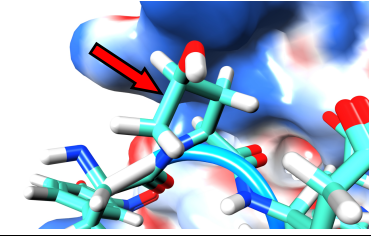
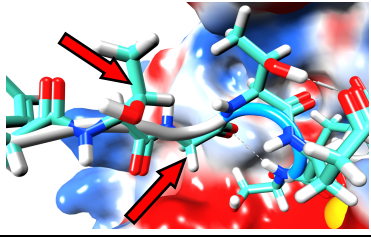


Figure 4.4. Representative images of FRET peptide substrates docked into the active site of the crystal structure of ADAMTS-5 using computational modelling. The substrates Abz-TESESARGAIY-Dpa-KK-NH₂ (A) and Y(NO₂)TESESRGK(Abz)IYYKKG (B) were docked into the active site of the crystal structure of ADAMTS5-5 (2RJQ, protein data bank) using MOE software. The green inset in (A) represents a magnification of the Dpa quencher and its surroundings. The negatively and positively charged electrostatic energy surfaces are colour-coded red and blue, respectively. The zinc (Zn²⁺) and calcium (Ca²⁺) ions are depicted by yellow and dark red balls, respectively. The backbone of the peptide is represented by white 'piping', with non-covalent interactions designated by dotted lines. Carbon atoms are indicated in cyan, hydrogen atoms in white, nitrogen atoms in blue and oxygen atoms in red.

Table 4.1. Summary of FRET peptide substrates containing a P4' fluorophore docked into the active site of the crystal structure of ADAMTS-5.

Compound	Peptide sequence	Snapshot of modification (MOE)
3	Y (NO ₂) TESES R G K (Abz) IYYKKG	
4	Y (NO ₂) TE G ESRGK (Abz) IYYKKG	
5	K Y (NO ₂) TESESRGK (Abz) IYYKKG	
6	Y (NO ₂) TESESRGK (Abz) J YYKKG <i>J = β-cyclopropyl-alanine</i>	
7	Y (NO ₂) P ESESRGK (Abz) IYYKKG	
8	Y (NO ₂) X ESESRGK (Abz) IYYKKG <i>X = 4-hydroxyproline</i>	
9	Y (NO ₂) TG TESESRGK (Abz) IYYKKG	

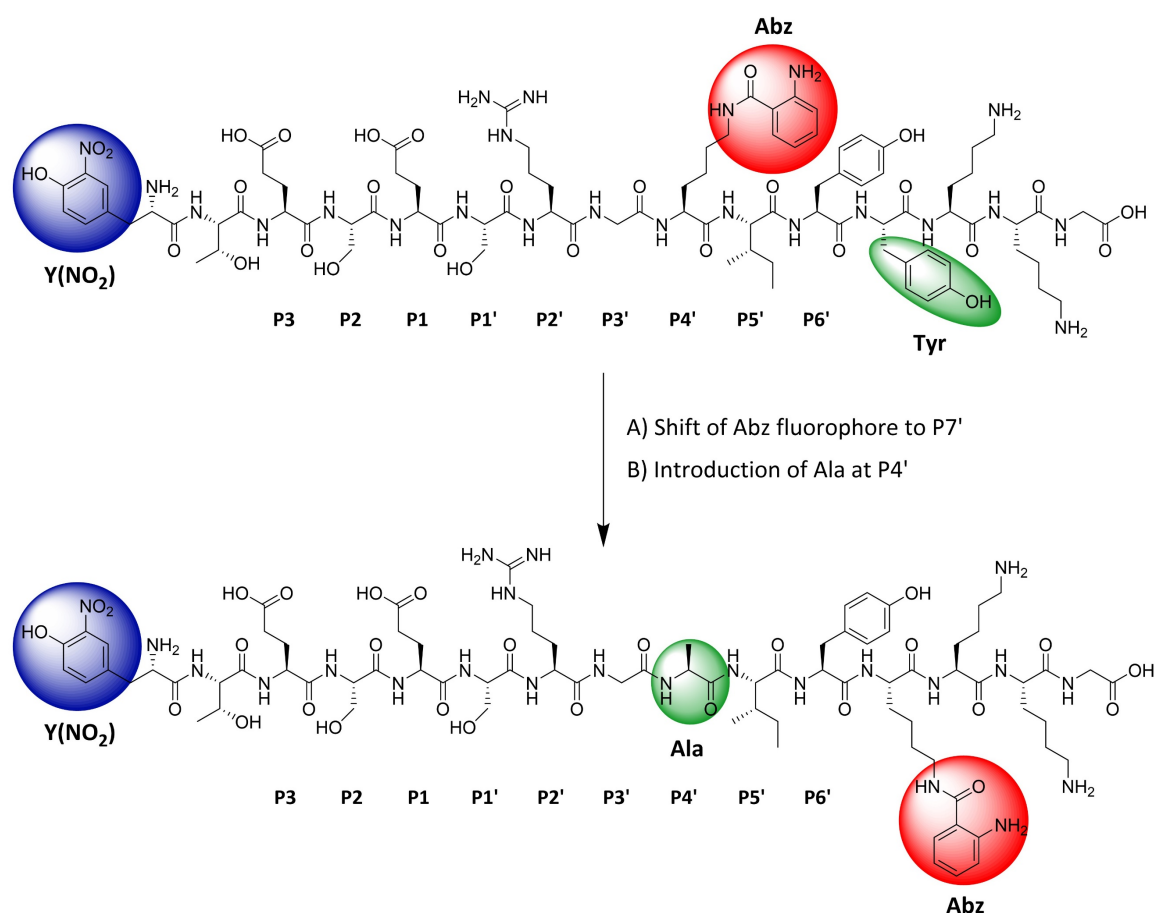
The initial substrate Y(NO₂)TESESRGK(Abz)IYYKKG (compound **3**, Table 4.1) underwent conservative modification using single amino acid changes. The first change involved introduction of a Gly residue in P2, as this residue is found in the main aggrecan cleavage site (202) (compound **4**, Table 4.1). The introduction of a Lys residue at the N-terminus was carried out to increase the solubility of the overall substrate (compound **5**). The un-natural amino acid β -cyclopropyl-alanine (J) was introduced instead of Ile at P5' (compound **6**) to improve selectivity at that position whilst ensuring a conservative change.

Multiple substrates from Figure 4.3 (18%) contained a proline residue at P4. As a consequence of this, it was introduced at this position along with a 4-hydroxyproline (X) (compounds **7-8**). Finally, the peptide sequence was extended with Thr and Gly residues (compound **9**) to account for any possibility that a longer peptide sequence may be required for successful substrate cleavage. This is because many of the natural substrates for ADAMTS-5 (*e.g.* aggrecan) are longer than the 15 residues found in **3**.

4.2.5 Design and computational docking of a P7' fluorophore series of ADAMTS-5 FRET peptide substrates

Three substrates in Table 8.1 contained a quencher in P7', including the existing ADAMTS-5 literature substrate **2**. In addition, one of the substrates incorporated a 2,4-dinitrophenyl (Dnp) quencher into the side chain of Lys, namely C₆-(GPXPXG)₂-GPXGTK(Mca)GE~L~E~GR~GTK(Dnp)-GIS(GPXPXG)₂-GPX-NH₂ (where C₆ is hexanoic acid and X is hydroxyproline) (203, 204). As a result of these observations, the K(Abz) fluorophore was moved to P7' (Scheme 4.3, step A). An alanine was used to replace it at P4' so that the initial substrate within the P7' series, *i.e.* Y(NO₂)TESESRGAIYK(Abz)KKG (compound **10**), was consistent with the existing literature substrate Abz-TESE~SRGAIY-Dpa-KK-NH₂ (**2**, step B).

The initial substrate of the P7' series is shown docked into the active site of TS5-5 in Figure 4.5, with a snapshot of the K(Abz) modification shown in Table 4.2 (compound **10**).



Scheme 4.3. Two-step design process for the development of a P7' fluorophore series of ADAMTS-5 FRET peptide substrates informed by computational docking. The docked substrate Y(NO₂)TESESRGK(Abz)IYYKKG was used as the starting point for the design process. (A) The Abz fluorophore (red) was shifted to P7' to replace Tyr (green); (B) An Ala residue (green) was added to P4' to replace the shifted Abz fluorophore.

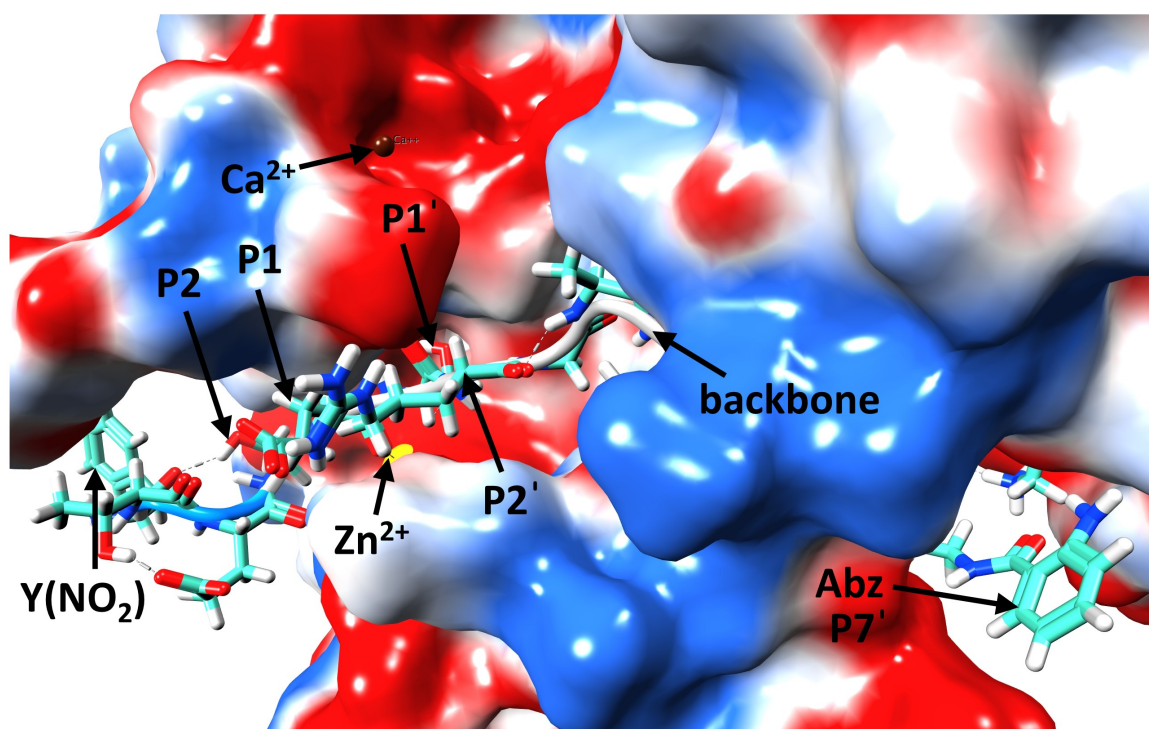
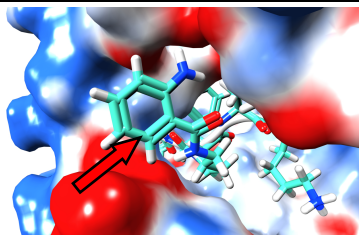
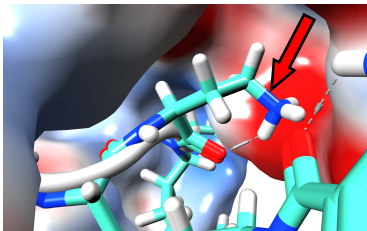
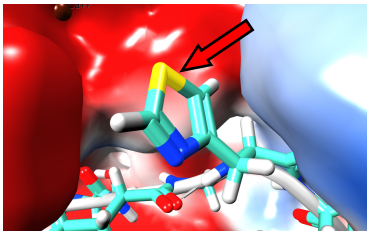
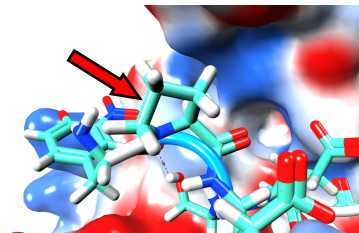
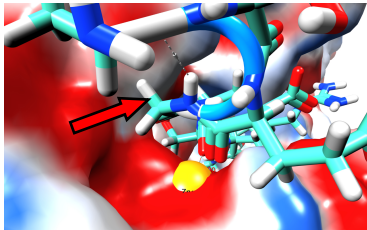


Figure 4.5. Representative image of a FRET peptide substrate with a P7' fluorophore docked into the active site of the crystal structure of ADAMTS-5 using computational modelling. The substrate Y(NO₂)TESESRGAIYK(Abz)KKG was docked into the active site of the crystal structure of ADAMTS5-5 (2RJQ, protein data bank) using MOE software.

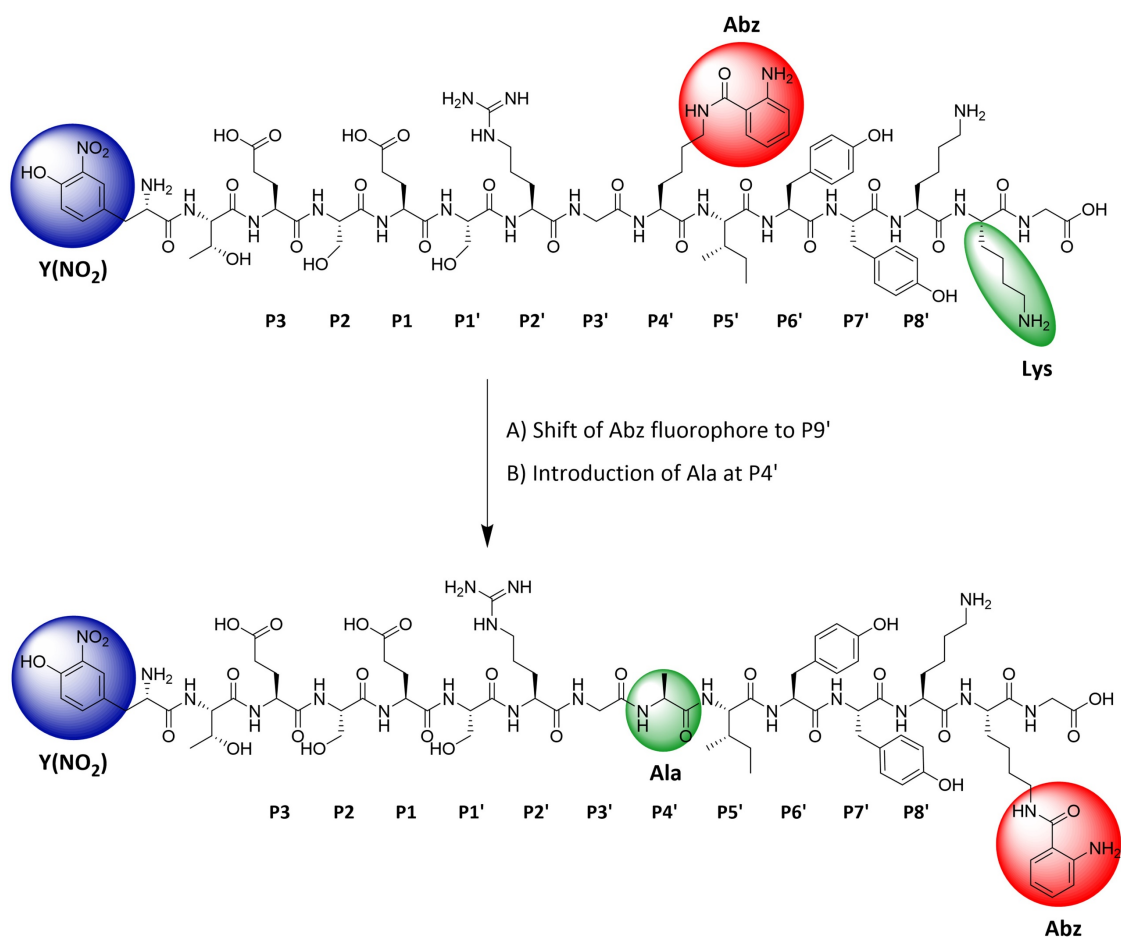
Looking at Table 4.2, the initial change to **10** involved the replacement of Lys at P8' with a 2,4-diaminobutyric acid residue (B, compound **11**). This is because the interaction of the longer lysine side chain with the “roof” of the binding pocket was considered sterically unfavourable during modelling. A 4-thiazolyl-alanine (Z) was introduced instead of the Ala at P4', as the sulphur atom of the thiazole would have a favourable non-covalent interaction with the positively charged molecular energy surface on the right-hand side (compound **12**). The insertion of a Pro at P4 (compound **13**) and a Gly at P2 (compound **14**) were carried out for the same reasons as mentioned in the P4' fluorophore series.

Table 4.2. Summary of FRET peptide substrates containing a P7' fluorophore docked into the active site of the crystal structure of ADAMTS-5.

Compound	Peptide sequence	Snapshot of change (MOE)
10	Y (NO ₂) TESES ^R GAIY K (Abz) KKG	
11	Y (NO ₂) TESES ^R GAIYK (Abz) B KG <i>B = 2,4-Diaminobutyric acid</i>	
12	Y (NO ₂) TESES ^R G Z IYK (Abz) KKG <i>Z = 4-Thiazolyl-alanine</i>	
13	Y (NO ₂) P ESES ^R GAIYK (Abz) KKG	
14	Y (NO ₂) TE G ES ^R GAIYK (Abz) KKG	

4.2.6 Design and computational docking of a P9' series of ADAMTS-5 FRET peptide substrates

The final series of substrates to be docked into TS5-5 involved a shift of K(Abz) to the P9' position with the concomitant introduction of Ala at P4' (Scheme 4.4). Ala was introduced at P4' to closely resemble the literature substrate **2**. The K(Abz) was moved to P9' as there were Lys residues already present and sufficient space was available to accommodate this fluorophore. The docked structure of the initial peptide in this series—namely Y(NO₂)TESESRGAIYYKK(Abz)G (compound **15**)—can be seen in Figure 4.6 and the K(Abz) change in Table 4.3. Further modifications to this structure followed a similar pattern to the P4' and P7' series; a Pro at P4 (compound **16**) and a Gly at P2 (compound **19**). A Leu was introduced as a conservative change at P7' (compound **17**) and the un-natural amino acid β -cyclopropyl-alanine (J) was also used at that position (compound **18**).



Scheme 4.4. Two-step design process for the development of a P9' fluorophore series of ADAMTS-5 FRET peptide substrates informed by computational docking. The docked substrate Y(NO₂)TESESRGK(Abz)IYYKKG was used as the starting point for the design process. (A) The Abz fluorophore (red) was shifted to P9' to replace Lys (green); (B) An Ala residue (green) was added to P4' to replace the shifted Abz fluorophore.

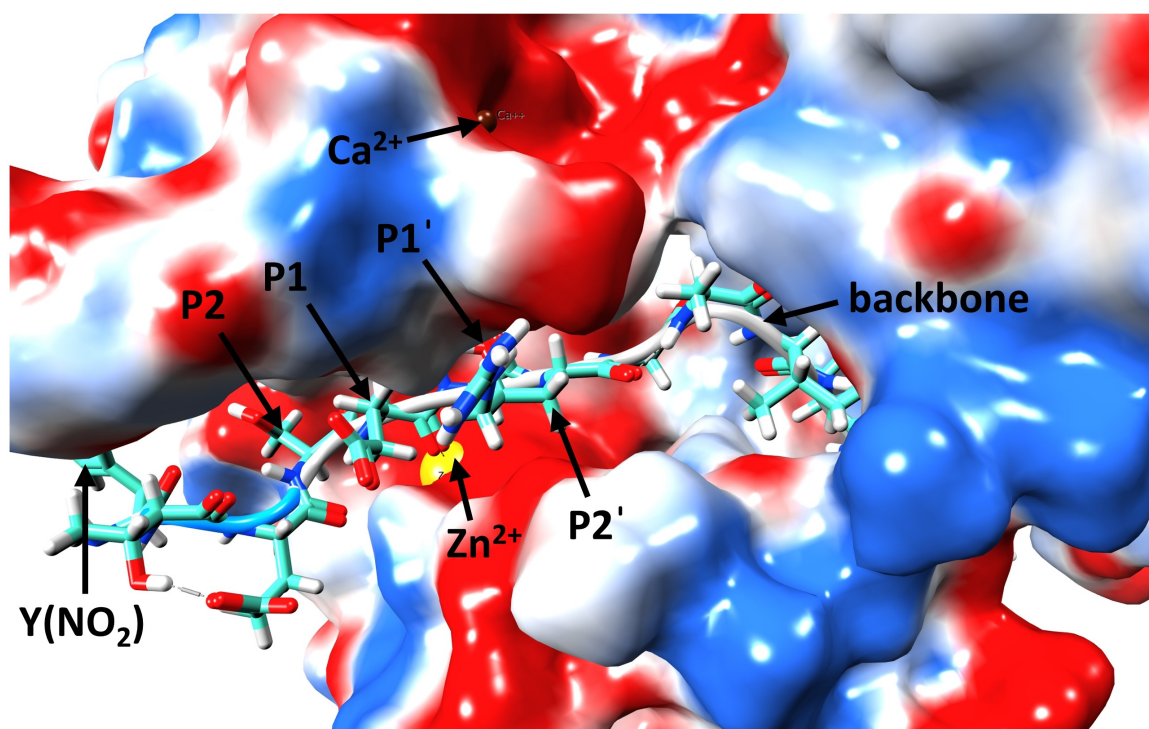


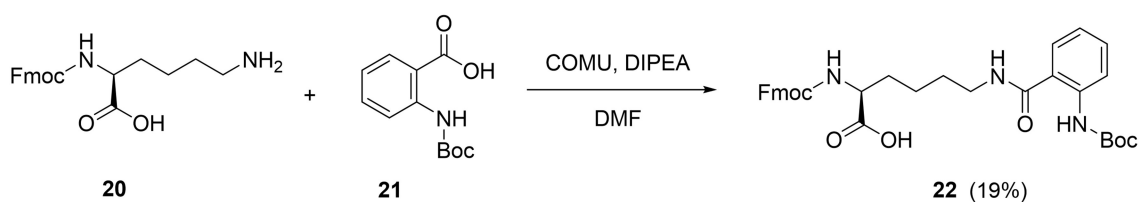
Figure 4.6. Representative image of a FRET peptide substrate with a P9' fluorophore docked into the active site of the crystal structure of ADAMTS-5 using computational modelling. The substrate Y(NO₂)TESES_{RG}AIYYKK(Abz)G was docked into the active site of the crystal structure of ADAMTS5-5 (2RJQ, protein data bank) using MOE software.

Table 4.3. Summary of FRET peptide substrates containing a P9' fluorophore docked into the active site of the crystal structure of ADAMTS-5.

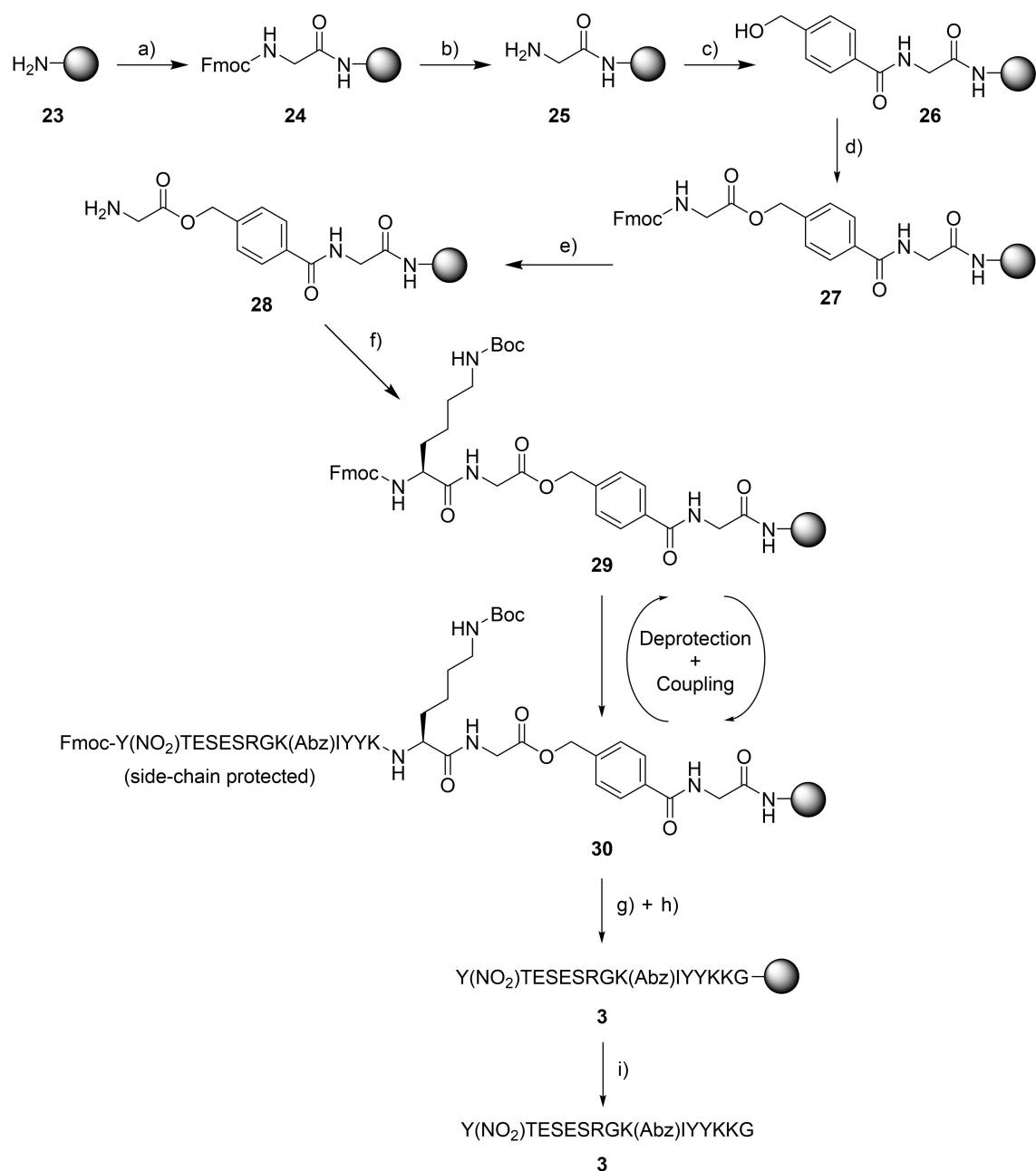
Compound	Peptide sequence	Snapshot of change (MOE)
15	Y (NO ₂) TESES ^R GAIYYK K (Abz) G	
16	Y (NO ₂) P ES ^R ESRGAIYYKK (Abz) G	
17	Y (NO ₂) TESES ^R GAIY L KK (Abz) G	
18	Y (NO ₂) TESES ^R GAIY J KK (Abz) G <i>J</i> = β -cyclopropyl-alanine	
19	Y (NO ₂) TE G ESRGAIYYKK (Abz) G	

4.2.7 Fmoc-based solid-phase peptide synthesis of modelled ADAMTS-5 FRET peptide substrates

Having modelled a series of peptide substrates with the Abz fluorophore in P4', P7' and P9' positions, they were synthesised using Fmoc-SPPS to rapidly test for cleavage on-resin by TS5-5. The Abz fluorophore was introduced into the peptide sequence through Fmoc-Lys(*retro*-Abz-Boc)-OH. As this amino acid is expensive to purchase and a considerable amount would be required for later combinatorial library synthesis, it was also synthesised according to Scheme 4.5 (205). Fmoc-L-Lys-OH (**20**) was coupled to Boc-Abz-OH (**21**) using COMU to deliver Fmoc-Lys(*retro*-Abz-Boc)-OH (**22**) in 19% yield. The HPLC chromatograms and ¹H-NMR spectrum of the purified compound can be found in Appendix Figures 8.11 and 8.12, respectively. The low yield from this synthesis was due to difficulties in chromatographic separation of the product from by-products and unreacted starting material. As a result of this, purification had to be performed multiple times to gradually separate out **22**, which was subsequently used in the synthesis of the FRET peptide substrates by manual Fmoc-SPPS. The general synthetic procedure for this is shown in Scheme 4.6 and is exemplified by the synthesis of the P4' fluorophore peptide Y(NO₂)TESESRGK(Abz)IYYKKG (**3**).



Scheme 4.5. Synthesis of the small molecule fluorophore building block Fmoc-Lys(*retro*-Abz-Boc)-OH used in Fmoc-SPPS of ADAMTS-5 FRET peptide substrates. Reaction conditions: i) Boc-Abz-OH (**21**, 1.2 equiv.), COMU (1.2 equiv.), DIPEA (2.45 equiv.), DMF, RT, 30 min; ii) Fmoc-L-Lysine-OH (**20**, 1.0 equiv.), RT, 21 h.



Scheme 4.6. Synthetic route to ADAMTS-5 FRET peptide substrates through manual Fmoc-SPPS using PEGA resin. Synthetic pathway to peptide substrates illustrated for the peptide Y(NO₂)TESESRGK(Abz)IYYKKG (**3**). Reaction conditions: **a**) PEGA₁₉₀₀ resin (**23**, 1.0 equiv.), Fmoc-Gly-OH (4.0 equiv.), DMF, NEM (3.93 equiv.), TBTU (4.0 equiv.), RT, 1.5 h; **b**) 20% piperidine/DMF, RT, 20 min; **c**) HMBA (4.0 equiv.), DMF, NEM (3.93 equiv.), TBTU (4.0 equiv.), RT, 2 h; **d**) Fmoc-Gly-OH (4.0 equiv.), DCM, 1-Melm (4.6 equiv.), MsNT (4.0 equiv.), RT, 1 h x 2; **e**) 20% piperidine/DMF, RT, 5 min; **f**) Fmoc-Lys(Boc)-OH (4.0 equiv.), DMF, NEM (3.93 equiv.), TBTU (4.0 equiv.), RT, 1 h; **g**) 20% piperidine/DMF, 1 x 2 min, 1 x 20 min; **h**) TFA : TIPS : H₂O (95% : 2.5% : 2.5%), RT, 1 x 10 min, 1 x 3 h; **i**) 0.1 M NaOH, RT, 1 h.

Fmoc-Gly-OH was coupled to PEGA₁₉₀₀ resin (**23**) to give **24**. This was followed by deprotection of the Fmoc group and coupling of the base labile HMBA linker to deliver **26**. The coupling of the initial Gly of the peptide sequence was performed twice with MsNT activation and furnished **27**. Subsequent deprotection followed by coupling of Fmoc-Lys(Boc)-OH with TBTU to give **29** was repeated for each amino acid as it was added to the growing peptide chain. A Kaiser test (Chapter 2, section 2.3.4) was used to test for successful couplings. Once the desired peptide **30** had been acquired, side chain protecting groups were removed using a cleavage cocktail consisting of TFA and scavengers. The resulting resin beads containing bound peptide substrate **3** were taken forward to assess cleavage by the TS5-5 proteinase (sub-section 4.2.8).

To confirm that **3** had been successfully synthesised, the resin beads that were not used to assess cleavage by TS5-5 were treated with 0.1 M NaOH to cleave the HMBA linker and release the peptide from the resin. This afforded crude peptide **3**. A total of 17 peptide sequences were synthesised with crude purities ranging from 64-91% by UHPLC. Their characterisation data is shown in Table 4.4 and their on-resin cleavage was evaluated according to Scheme 4.7.

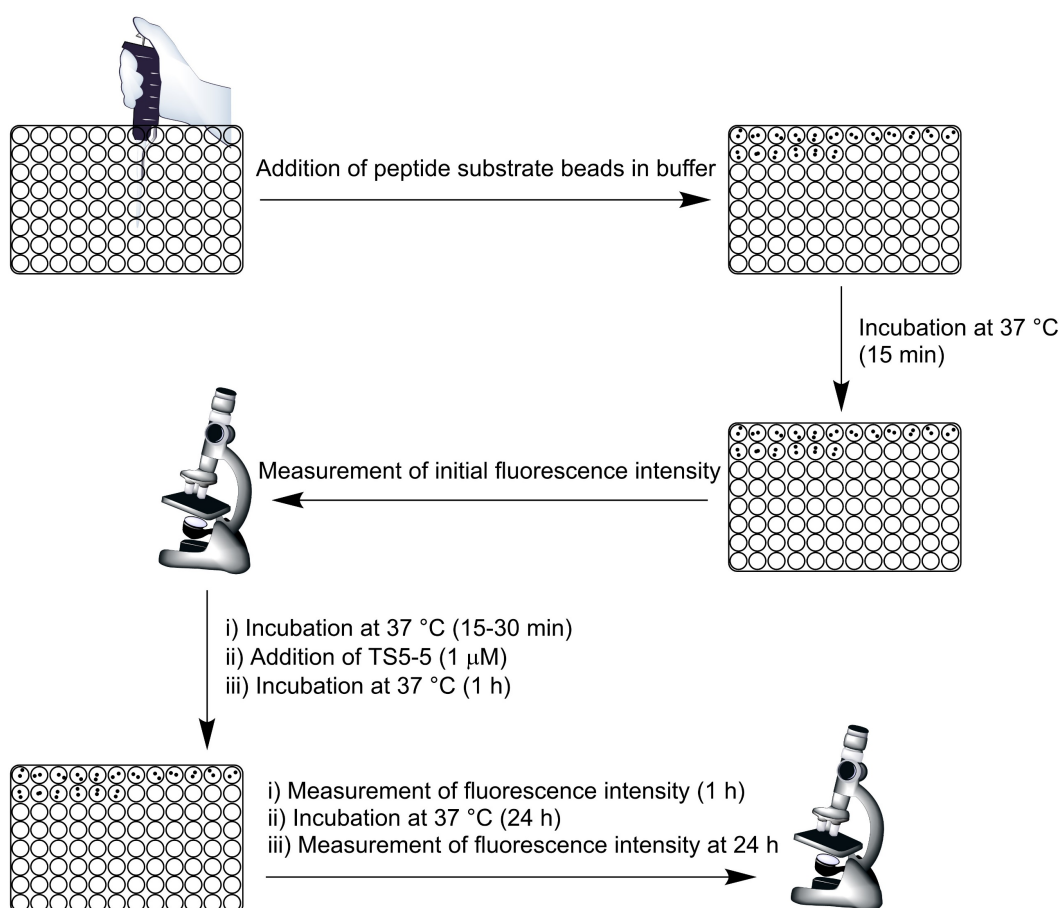
Table 4.4. Summary of characterisation data for ADAMTS-5 FRET peptide substrates synthesised by manual Fmoc-SPPS using PEGA resin.

Compound	Peptide sequence	Crude % Purity (by UHPLC)	HRMS (calcd) [M+H] ⁺	HRMS (found) [M+H] ⁺
P4'	3 Y(NO ₂)TESESRGK(Abz)IYYKKG	≥ 81	1972.9405	1972.9428
	4 Y(NO ₂)TE G ESRGK(Abz)IYYKKG	≥ 71	1942.9299	1942.9314
	5 K Y(NO ₂)TESESRGK(Abz)IYYKKG	≥ 76	2101.0355	2101.0378
	6 Y(NO ₂)TESESRGK(Abz) J YYKKG	≥ 77	1970.9248	1970.9262
	7 Y(NO ₂) P ESERGGK(Abz)IYYKKG	≥ 64	1968.9456	1968.9464
	8 Y(NO ₂) X ESERGGK(Abz)IYYKKG	≥ 65	1984.9405	1984.9382
	9 Y(NO ₂) TG ESERGGK(Abz)IYYKKG	≥ 72	2131.0096	2131.0080
P7'	10 Y(NO ₂)TESESRGAIY K (Abz)KKG	≥ 78	1880.9143	1880.9150
	11 Y(NO ₂)TESESRGAIYK(Abz) B KG	≥ 80	1852.8830	1852.8813
	12 Y(NO ₂)TESESRG Z IYK(Abz)KKG	≥ 75	1963.8972	1963.8980
	13 Y(NO ₂) P ESERGAIYK(Abz)KKG	≥ 72	1876.9194	1876.9202
	14 Y(NO ₂)TE G ESRGAIYK(Abz)KKG	≥ 76	1850.9037	1850.9057
P9'	15 Y(NO ₂)TESESRGAIYY K (Abz) G	≥ 78	1915.8826	1915.8854
	16 Y(NO ₂) P ESERGAIYYKK(Abz)G	≥ 66	1911.8877	1911.8853
	17 Y(NO ₂)TESESRGAIY L KK(Abz)G	≥ 91	1865.9034	1865.9061
	18 Y(NO ₂)TESESRGAIY J KK(Abz)G	≥ 88	1863.8877	1863.8887
	19 Y(NO ₂)TE G ESRGAIYYKK(Abz)G	≥ 82	1885.8721	1885.8747

Key: J = β-cyclopropyl-alanine; X = 4-Hydroxyproline; B = 2,4-Diaminobutyric acid; Z = 4-Thiazolyl-alanine. Residues altered within each peptide series are bolded in red.

4.2.8 Evaluation of modelled FRET peptide substrates against the ADAMTS-5 proteinase

The resin-bound FRET peptide substrates were tested for cleavage by the TS5-5 proteinase. Scheme 4.7 illustrates the general screening process (see Chapter 2, section 2.3.7 for the procedure).



Scheme 4.7. Determination of the extent of cleavage of synthesised FRET peptide substrates on-resin after incubation with ADAMTS-5. Workflow for measuring the change in fluorescence intensity of substrate beads after incubation with TS5-5. PEGA₁₉₀₀ beads (10–30, black) containing FRET peptide substrates were incubated in TNC buffer with 0.005% (v/v) Brij-35 in a 96-well plate at 37 °C (15 min). After measurement of the initial fluorescence intensity, substrates were incubated at 37 °C (15–30 min) before the addition of TS5-5 (1 μM). The fluorescence intensity of the substrate beads was then measured 1 h and 24 h after incubation at 37 °C.

Briefly, resin beads containing each substrate were added to wells containing buffer and incubated at 37 °C for 15 minutes. After the initial fluorescence intensity had been recorded, TS5-5 proteinase was added and the change in fluorescence intensity (ΔFI) measured after 1 h and 24 h. In addition, to confirm that the changes in fluorescence intensity were a result of FRET substrate cleavage by TS5-5 and not poor substrate stability, substrates were incubated in buffer only and ΔFI measured after 1 h and 24 h. The raw data for all of these experiments is shown in Appendix Figure 8.13.

To determine the change in fluorescence intensity after 1 h and 24 h while accounting for substrate stability, the ΔFI for each substrate in the presence of TS5-5 (ΔFI_{TS5-5}) was subtracted from the ΔFI for each substrate in the absence of TS5-5 ($\Delta FI_{control}$). The processed data is shown in Figure 4.7 for each individual peptide substrate after 1 h (A) and 24 h (C) and for substrates summed by P4', P7' and P9' fluorophore positions after 1 h (B) and 24 h (D).

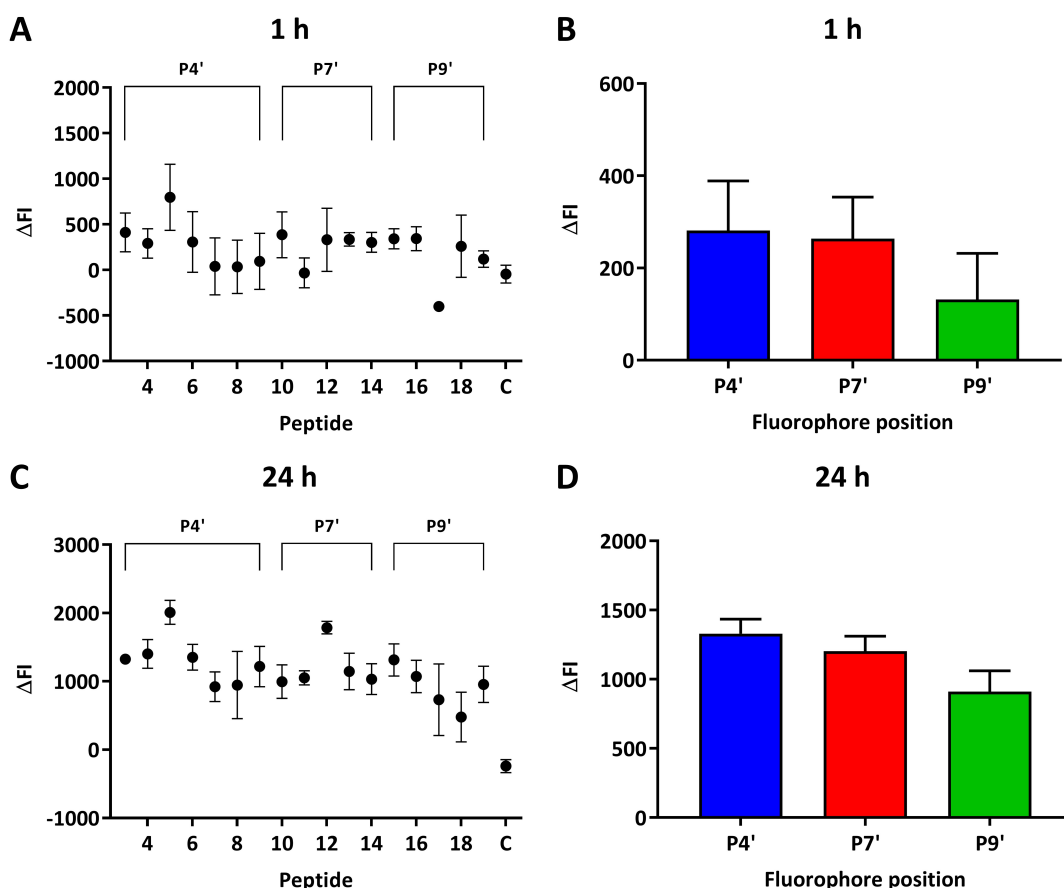


Figure 4.7. Change in fluorescence intensity at different time points after incubation of on-resin FRET peptide substrates with ADAMTS-5. (A, C) Individual changes in the mean fluorescence intensity (ΔFI) $\pm$ SEM for peptide substrates on resin in the presence of TS5-5 (ΔFI_{TS5-5}) subtracted from the fluorescence intensity in the absence of TS5-5 ($\Delta FI_{control}$): $\Delta FI = \Delta FI_{TS5-5}$ (background-subtracted FI at time t – background-subtracted FI at time 0) – $\Delta FI_{control}$ (background-subtracted FI at time t – FI at time 0) at 1 h (A) and 24 h (C) for $n = 3$ wells from three independent experiments performed in single replicate at 37 °C. (B, D) Changes in mean fluorescence intensity (ΔFI) $\pm$ SEM summed by fluorophore position (P4', P7' and P9') after 1 h (B) and 24 h (D) for data in (A) and (C), respectively. In all cases, cleavage of peptide substrates was monitored at an $\lambda_{ex} = 350$ nm and $\lambda_{em} = 438$ nm in TNC buffer with 0.005% (v/v) Brij-35. C = control, un-modified PEGA₁₉₀₀ resin with no substrate.

Incubation of P4', P7' and P9' fluorophore substrates with TS5-5 for 1 h revealed no substantial increase (*i.e.* < 1000 fluorescent units) in mean ΔFI , with values between 0 – 500 fluorescence units (Figure 4.7A). The only exception was peptide **5**, which showed an increase of 796 ± 364 , which was approximately two-fold higher than the next closest substrate **3** (411 ± 213). The drop in fluorescence intensity for peptide **17** (-402 ± 11) likely reflects a small increase in fluorescence intensity in the presence of TS5-5 ($\Delta FI_{TS5-5} = 128 \pm 11$, Appendix Figure 8.13A) coupled with an unusually high control fluorescence intensity ($\Delta FI_{control} = 530 \pm 152$, Appendix Figure 8.13B), the latter of which may have resulted from variations in fluorescence intensity between individual substrate beads. When ΔFI was summed by fluorophore position (Figure 4.7B), a higher increase in ΔFI was found for the P4' and P7' fluorophore positions compared to the P9' fluorophore position (*cf.* 281 ± 108 and 264 ± 90 to 132 ± 100). As the individual increases in fluorescence intensity for all peptide substrates were not substantial (*i.e.* < 1000 fluorescence units), it was difficult to assess which substrates were cleaved most readily by TS5-5. Despite the fact that the ΔFI was computed for substrates by fluorophore position, there was no clear preference for substrate cleavage with a fluorophore in either of the P4', P7' and P9' positions after 1 h.

In contrast, incubation with TS5-5 for 24 h resulted in a much larger increase in ΔFI (*i.e.* > 1000 units) for almost all peptide substrates, with the exception of peptides **7**, **8**, **10** and **17–19** (Figure 4.7C). Values above 1500 fluorescent units were observed for peptides **5** (2010 ± 174) and **12** (1786 ± 192). When changes were summed by position (Figure 4.7D), a stepwise decrease in mean $\Delta FI \pm SEM$ was observed when the fluorophore was shifted from P4' (1329 ± 105) -> P7' (1202 ± 110) -> P9' (911 ± 151). In both 1 h and 24 h experiments (Figure 4.7A and C, respectively), control PEGA₁₉₀₀ beads showed no increase in ΔFI . The slight decrease in ΔFI for PEGA₁₉₀₀ beads at 1h (-46 ± 98) and 24 h (-240 ± 95) time points is a

consequence of differences in fluorescence intensity between individual beads during measurement. As the P4' fluorophore series of FRET peptide substrates had the highest ΔFI after 24 h (1329 ± 105), the P4' position was selected as the optimal position for fluorophore placement. Within this series, peptide **5** had the highest ΔFI after 24 h (2010 ± 174) and was thus chosen as the initial lead candidate for combinatorial library synthesis.

To illustrate the change in fluorescence intensity for an individual bead containing a FRET peptide substrate, images of the same bead from peptide **5** were captured at 0 h, 1h and 24 h time points (Figure 4.8). The 'ring' of fluorescence intensity described earlier can be seen to gradually expand inwards over time (Figure 4.8A -> C). This increase in fluorescence intensity can be seen quantitatively in Figure 4.8D, although it appears to plateau prior to the 24 h timepoint.

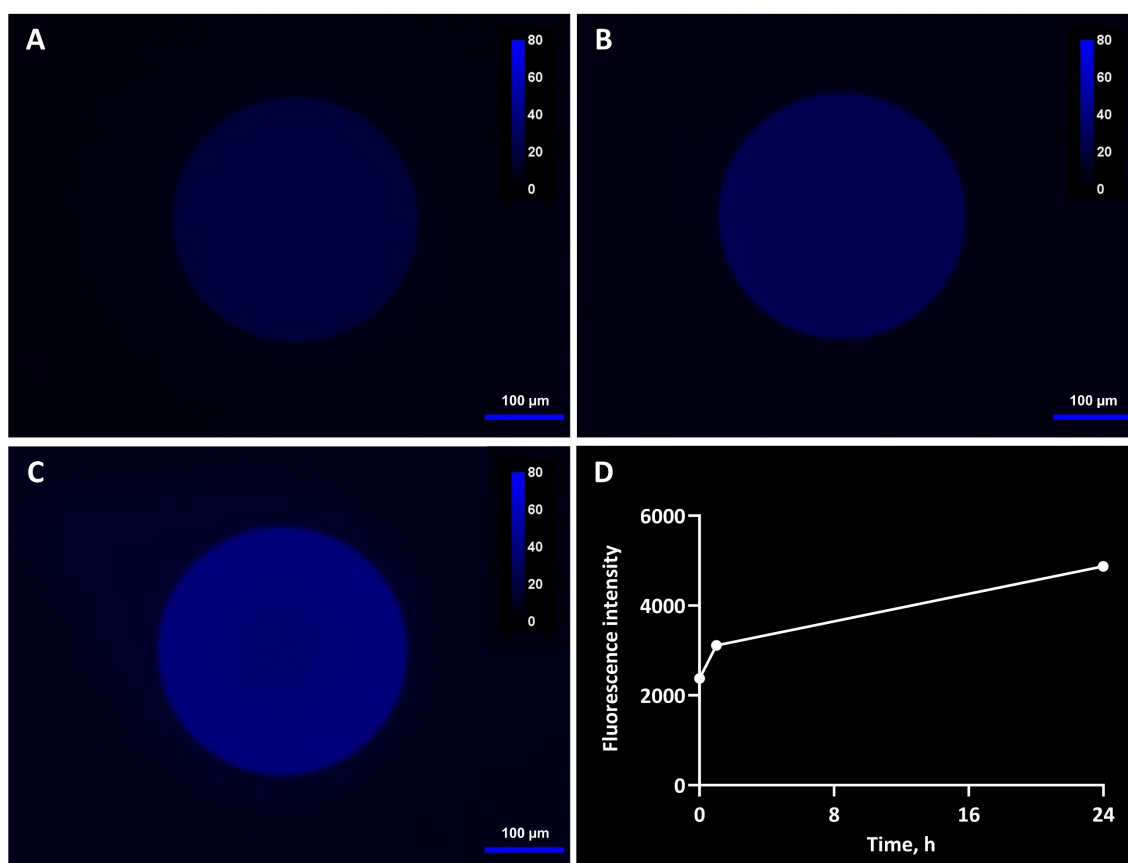


Figure 4.8. Qualitative and quantitative assessment of the change in fluorescence intensity of a single resin bead containing a FRET peptide substrate before and after incubation with ADAMTS-5. (A-C) Fluorescent images of the same PEGA₁₉₀₀ bead containing the peptide substrate KY(NO₂)TESESRGK(Abz)IYYKKG (**5**) at time = 0 h (A) and 1 h (B) and 24 h (C) after incubation with TS5-5 (500 nM). The bead was incubated in TNC buffer with 0.005% (v/v) Brij-35 and imaged using a fluorescent microscope (λ_{ex} = 350 nm and λ_{em} = 438 nm). These images were processed and false coloured using Fiji software. (D) Change in total fluorescence intensity over time for the same PEGA₁₉₀₀ bead containing the peptide substrate KY(NO₂)TESESRGK(Abz)IYYKKG after the addition of TS5-5 (1 μM). The total fluorescence intensity of the bead was measured using a three-point circle (r = 50 μm) and subtracted from the same radius of background solution in cellSens Dimension software.

The combinatorial library was thus designed using the structure of peptide **5** as the starting point (Figure 4.9). In the same fashion as the initial docking performed for each P' fluorophore series, amino acids were selected and substrates docked through the four-step process described at the end of sub-section 4.2.3.

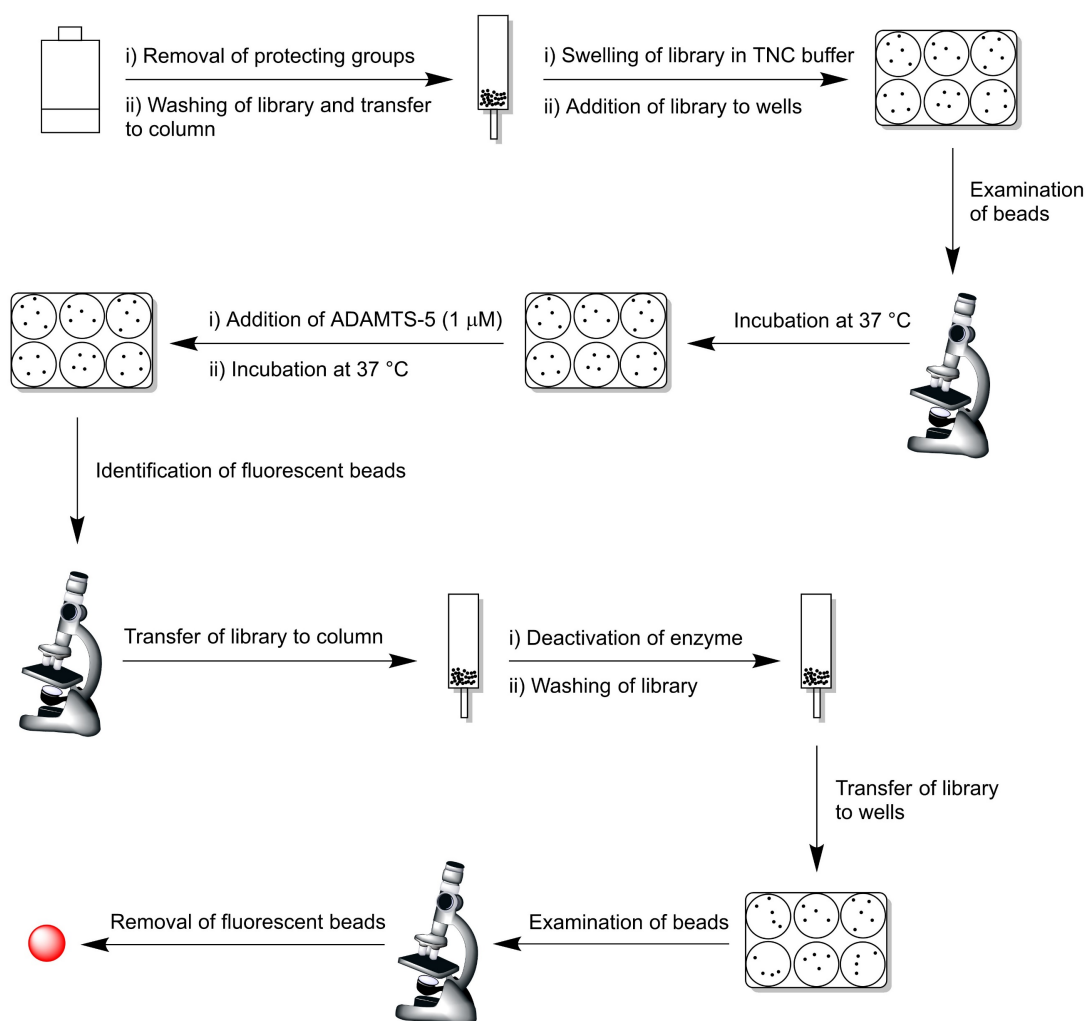
P6	P5	P4	P3	P2	P1		P1'	P2'	P3'	P4'	P5'	P6'	P7'	P8'	P9'	P10'
K	Y(NO ₂)	T	E	S	E	~	S	R	G	K(Abz)	I	Y	Y	K	K	G
		S	D	T				K	A		A	H	F	U		
		N		G				O			V	Z				
		Q		N				B			J					

Figure 4.9. Design of a combinatorial library of FRET peptide substrates for screening against ADAMTS-5. A combinatorial library was designed using the peptide substrate KY(NO₂)TESESRGK(Abz)IYYKKG (**5**) as the parent compound, with variable positions indicated below the grey area. The quencher Y(NO₂) is coloured in blue, and the fluorophore K(Abz) in red. Natural amino acids are indicated using their single letter code with un-natural amino acids coloured in purple and abbreviated as follows: O = ornithine, B = 2,4-Diaminobutyric acid, J = β -cyclopropyl-alanine, Z = 4-Thiazolyl-alanine, U = 2,3-Diaminopropionic acid.

However, more specific changes were made for the combinatorial library. Firstly, positions that were regarded as sufficiently optimised through docking and amino acid alignments were fixed. These included Lys in P6 (to improve solubility), Glu in P1 and Ser in P1' (these residues occur frequently in known ADAMTS-5 substrates, see Figure 4.3 and Appendix Table 8.1); along with Lys in P9' (to improve solubility) and Gly in P10' (to facilitate synthesis and reduce steric hindrance at the C-terminus).

Secondly, to facilitate hit identification by MALDI and tandem mass spectrometry, the total number of unique sequences was kept below 50,000. As a result of this, each unfixed position was limited to an average of three amino acid modifications. Conservative amino acid modifications were once again employed to ensure that cleavage of library substrates would not be disrupted by radical changes to side chain functional groups. For example, the P3' position was additionally modified with Ala, as this does not differ substantially from Gly. Un-natural amino acids were introduced once more to improve selectivity at specific positions. For example, derivatives of Lys were used through the systematic shortening of the side chain *via* O and B at P2'.

Combinatorial library synthesis was performed using the “split-mix” method (145) (described extensively in Chapters 1 and 2) with uniform PEGA₁₉₀₀ resin beads (182). The final library had a theoretical number of 3⁹ (19,683) unique peptide sequences represented on ~100,000 beads. The library incubation process was performed in accordance with Scheme 4.8. Initially, peptide side chain protecting groups were removed using a cleavage cocktail consisting of TFA and scavengers. Resin beads were then washed extensively, transferred to a column, swelled in buffer and transferred to the wells in tissue culture plates. Beads were carefully examined under a fluorescent microscope to check for any false positives before incubation at 37 °C for 30 minutes. A representative example of library beads prior to the addition of proteinase is shown in Figure 4.10A. After the addition of TS5-5, beads were incubated at 37 °C and examined for any signs of fluorescence 30 minutes later. A faint ‘ring’ of fluorescence was seen to be forming around beads as TS5-5 diffused inside and cleaved copies of each substrate (Figure 4.10B). Incubation was continued for a further 30 minutes to verify this cleavage. Figure 4.10C revealed that cleavage of substrate beads was in fact occurring. As a result of this, the library was transferred to a column and the TS5-5 proteinase deactivated with TFA. This deactivation was performed in order to enable hit beads to be identified without fear of continued substrate cleavage occurring on all other beads. After washing of the library, it was transferred back to the wells in the tissue culture plates and hit beads were identified and manually removed under a fluorescent microscope. A representative image of a hit bead is shown in Figure 4.10D. A total of 20 beads were isolated. They were cleaved from the PEGA₁₉₀₀ resin using an aqueous solution of TEA prior to sample preparation for MALDI and tandem mass spectrometry.



Scheme 4.8. Incubation of a combinatorial library of FRET peptide substrates with ADAMTS-5, deactivation of the proteinase and isolation of hits. A 16-mer combinatorial library was synthesised using the “split-mix” method. Initially, acid-labile protecting groups were cleaved from library peptides using 95% TFA: 5% scavengers. The library was then washed with DCM, DMF, 5% (v/v) DIPEA/DMF, water and TNC buffer with 0.02% (w/v) NaN_3 and transferred to a column. After swelling in buffer, the library was added to 6-well plates and examined under a fluorescent microscope. The library was then incubated at 37 °C (30 min) and TS5-5 (1 μ M) added. After 1 hour, the library was added to a column and proteinase deactivated with 2% (v/v) TFA and washed with 2% (w/v) NaHCO_3 and water. The library was then transferred to multiple 6-well plates and fluorescent beads were removed manually after analysis under a fluorescent microscope.

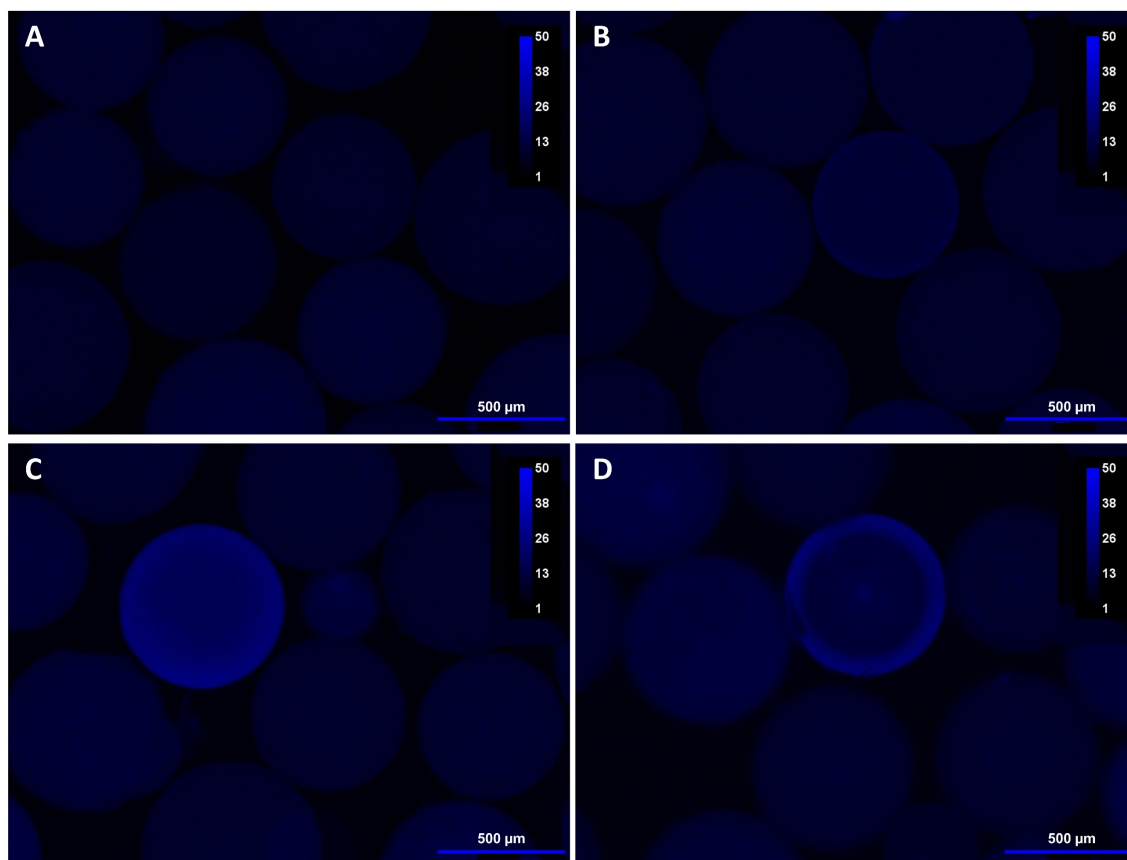


Figure 4.10. Qualitative time-course assessment of the change in on-resin fluorescence intensity of a combinatorial library of FRET peptide substrates before and after incubation with ADAMTS-5. (A-D) Fluorescent images of PEGA₁₉₀₀ beads from the combinatorial library containing different peptide substrates at time = 0 h (A) and 30 min (B), 1 h (C) and 24 h (D) after incubation with TS5-5 (500 nM). The library was incubated in TNC buffer with 0.005% (v/v) Brij-35 and imaged using a fluorescent microscope ($\lambda_{\text{ex}} = 350 \text{ nm}$ and $\lambda_{\text{em}} = 438 \text{ nm}$). These images were processed and false coloured using Fiji software.

4.3. Discussion

In this chapter, the ADAMTS-5 proteinase was chosen as a target for the development of a more selective FRET peptide substrate. This process began with the design and computational docking of peptide derivatives of Abz-TESE~SRGAIY-Dpa-KK-NH₂ (**2**) followed by their synthesis through Fmoc-SPPS and screening against active TS5-5. This led to the identification of the initial lead peptide KY(NO₂)TESESRGK(Abz)IYYKKG (**5**), which was used as a foundation for the development of a combinatorial peptide library. This library was

constructed by split-mix synthesis and screened against active TS5-5 to identify novel hit peptide substrates.

The decision to select ADAMTS-5 over ADAMTS-4 as the preferred target for substrate development was supported by a comparison between the selectivity studies of their respective substrates (*i.e.* **2** and **1**) against a panel of different metalloproteinases. A wide range of different metalloproteinases were included in this panel so that selectivity could be assessed across different families (*e.g.* ADAMTSs vs MMPs) and sub-families (*e.g.* collagenases vs gelatinases). It was clear that **2** was less selective for TS5-5 compared to **1** for TS4-4, given the high rate of cleavage by MMP-2 Δ C ($136 \pm 17\%$) and TS4-4 ($22 \pm 1\%$). In addition, substrate **1** was identified through screening of a large phage-display library (10^{12} phage) (63). This would make it more difficult to improve upon, as this not only involved five rounds of biopanning, but also delivered a $K_m = 13.5 \mu\text{M}$ and $k_{\text{cat}}/K_m = 1.41 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ for cleavage by TS4-4 (63). This low micromolar K_m and high catalytic efficiency indicate both strong substrate binding and rapid substrate cleavage, respectively (206). In contrast, substrate **2** was derived from part of the sequence of rat brevican where ADAMTS-4 cleavage occurs (62), making it a better overall target for improvement, as it was not identified through any formal biological or chemical screening process and had a considerably higher reported K_m of $76 \mu\text{M}$ (62).

All of the metalloproteinases included in the selectivity studies were obtained from laboratory stocks and the MMPs were initially tested for activity against Mca-PLG~L-Dpa-AR-NH₂. To ensure that accurate concentrations were available for selectivity studies following long term storage at -20°C or -80°C , each metalloproteinase was titrated with N-TIMP-3. The only exceptions were MMP-13 Δ C (as this was purchased commercially), and MMP-1 Δ C, which had already been titrated by a trusted member of the laboratory.

It is well known that surfactants modulate enzyme activity through non-covalent interactions (207). A study by Park and co-workers using different types of ionic and non-ionic surfactants revealed different rates of Mca-PLG~L-Dpa-AR-NH₂ or Mca-PKPLA~L-Dpa-AR-NH₂ FRET substrate cleavage by MMP-7, MMP-9, MMP-14 and MMP-26 (208). In addition, different concentrations of Brij-35 were evaluated in terms of Mca-PKPLA~L-Dpa-AR-NH₂ cleavage by MMP-7, MMP-9 and MMP-14. The highest rates of cleavage were found to be between approximately 0.007–0.02% (v/v) Brij-35 (208), which was close to the critical micelle concentration of this surfactant (90 µM). To the best of our knowledge, the effect of different surfactant concentrations on substrate cleavage by ADAMTS-4 and ADAMTS-5 proteinases has not been assessed, despite optimisation of other components of the assay buffer (17). This is important because the final concentration of Brij-35 is crucial in the context of combinatorial library screening. A concentration of Brij-35 within the optimal range of Brij-35 concentrations would yield a higher rate of substrate cleavage and would thus maximise signal detection on beads, enabling more facile identification of substrate hits than if an un-optimised concentration were used. This is also important because the rate of cleavage of substrates on PEGA resin beads by proteinases are generally 5-fold lower than in solution (160).¹³ As a result of these factors, the rate of ADAMTS-4 and ADAMTS-5 substrate cleavage was optimised using a range of different Brij-35 concentrations (Figure 4.1). As mentioned in the results section, the most common concentrations of Brij-35 utilised in assays of ADAMTS-4 and ADAMTS-5 proteinases are 0.02% (v/v) and 0.05% (v/v) Brij-35. As can be seen from Figure 4.1, there is an increase in the rate of substrate cleavage when 0.05% (v/v) Brij-35 is used compared to 0.02% (v/v). This is highly statistically significant for

¹³ This value was given through a personal communication with Professor Morten P. Meldal. The publication states that 10–50-fold higher concentrations of enzyme are required for cleavage of substrates on-resin compared to in solution.

ADAMTS-4 ($p = 0.0004$) but not statistically significant for ADAMTS-5 ($p = 0.0632$). Nevertheless, these concentrations do not give the maximal amount of substrate cleavage, which probably lies between 0.0005–0.005% (v/v) Brij-35 for ADAMTS-4 and around 0.005% (v/v) Brij-35 for ADAMTS-5. As 0.005% (v/v) Brij-35 gave a high rate of substrate cleavage for both proteinases, it was adopted in all further studies.

Screening of the docked peptide substrates with TS5-5 after 24 h led to a stepwise decrease in fluorescence intensity when the fluorophore was shifted from P4' to P7' to P9'. However, this decrease was only found to be substantial between P4' and P9' fluorophore positions. This difference could be due to an increase in steric hindrance and unfavourable stereoelectronic interactions between the Abz fluorophore on the lysine side chain and the molecular energy surface of the proteinase at the P9' position compared to the P4' position. Additional positions for fluorophore placement such as P5', P6', P8' and P10' were excluded due to unfavourable steric effects that would arise from introduction of a long alkyl chain and bulky aromatic fluorophore in the restricted space available within the active site of the proteinase.

The availability of sufficient space to avoid unfavourable steric effects is especially important for the development of an *in vivo* FRET peptide substrate. Fluorophores such as Cy5.5 have a much higher molecular weight compared to Abz, with large biphenyl rings that would sterically clash with the residues in the active site at P7' and P9'. This makes the P4' position the most suitable, as it provides the largest amount of space available to accommodate a Cy5.5 fluorophore, in addition to a favourable negatively charged electrostatic energy surface. Alternatively, the stepwise decrease in fluorescence intensity could simply be due to the specific set of peptides synthesised within each series, with an alternative series delivering a different result. Nonetheless, given the preferable cleavage of

peptide **5** by TS5-5 and the availability of space within the P4' position for potential introduction of an *in vivo* fluorophore, this position was used as a foundation for the combinatorial library.

Although the final hit peptides were cleaved by TS5-5, a high rate of non-selective cleavage by MMP-2 and TS4-4 remains a possibility. This is because the literature substrate **2** whence the peptide substrates were derived was hydrolysed most rapidly by MMP-2 and TS4-4 after TS5-5 (Figure 4.2). Therefore, to reduce non-selective cleavage of FRET substrate hits by MMP-2 and TS4-4 (and thereby improve their final selectivity), a pre-screening step prior to incubation with TS5-5 could have been employed. Briefly, the combinatorial library would be incubated with MMP-2, the proteinase deactivated, fluorescent hits removed, and the library washed. This step would then be repeated for TS4-4. By removing non-selective hits at each pre-screening step, the remaining substrates would be more selective for TS5-5 by virtue of this, as cleavage by TS4-4 and MMP-2 would theoretically be worse. Practically however, the most important considerations for pre-screening would be the concentration of proteinase and the final incubation time. For example, given that hit beads were identified within 30 minutes of library incubation with TS5-5, this time window could also be optimal for removing hits that are favourably cleaved by MMP-2/TS4-4. In order to attempt to achieve a similar rate of substrate cleavage by 500 nM TS5-5 within the 30-minute time window, MMP-2 could be used at 700 nM and TS4-4 at 2.3 μ M. These numbers are based on the assumption that the relative cleavage rates of most library substrates by MMP-2 and TS4-4 (*i.e.* the selectivity) would be similar to the current literature substrate, where $\sim$ 1.4-fold higher and $\sim$ 4.5-fold lower cleavage rates were reported with respect to TS5-5 (Figure 4.2). After removal of these hits, screening against TS5-5 would thus deliver fluorescent hits possessing an improved level of selectivity for this proteinase compared to those obtained

from the single screen. In addition, it is important that effective pre-screening as opposed to excessive pre-screening is carried out. Excessive pre-screening with multiple proteinases could result in very few library beads available for incubation with TS5-5, limiting the total number of potential hits. The trade-off between effective and excessive pre-screening would have to be determined experimentally.

The question of whether to utilise a large combinatorial library or a small combinatorial library for screening requires careful consideration. Whilst combinatorial libraries do not possess the same size as phage display libraries, the overall size of the library is still important. In principle, the development of a large library (> 100,000 unique sequences) is advantageous, as it increases the possible number of sequences, which leads to a wider coverage of chemical space and thus a greater chance that rapidly cleaved substrates would be obtained following screening. Unfortunately, the difficulties in identifying peptide sequences from PEGA beads using MALDI and tandem mass spectrometry after library screening limit the final size of the library to < 50,000 unique sequences. For example, if a large library is used, there are often too many possibilities for a particular sequence, making the resultant MALDI spectrum difficult to interpret. A peptide sequence that is an effective substrate is biologically unusable if the sequence cannot be identified. On the other hand, a smaller library enables more facile identification of peptide sequences as there are fewer possibilities. Once these hit peptides have been identified, a second library can be synthesised based on the lead sequence and the process can be repeated. This approach leads to successive rounds of substrate optimisation leading to a final “super” substrate.

An alternative method to identifying hit peptide substrates is to utilise a technique known as substrate phage display (209). Based on phage display (210), large libraries of

peptide substrates are encoded on the surfaces of bacteriophages and are screened against a specific target proteinase (209, 211). This technique has been used to map the substrate specificity of ADAMTS-13 based on mutations to residues of a von Willebrand factor fragment (212). Very recently, a novel approach utilising two substrate phage display libraries has been coupled with next generation sequencing (SPD-NGS) to find novel substrates and cleavage sites for caspase-2, -3, -6, -7, -8 (213). The first library resulted in detailed mapping of the subsite amino acid specificity for each proteinase and the identification 10^4 to 10^5 unique cleavage sites, with additional ones identified in the second library (213). This approach was also applied to ADAM10 and ADAM17. The biggest advantage of SPD-NGS over combinatorial chemistry is that it employs larger libraries with a high level of diversity ($\sim 10^9$ unique sequences), potentially leading to more hit substrates if applied to ADAMTS-5 (213).

To conclude, 20 peptide hits were isolated after screening of a combinatorial peptide substrate library with ADAMTS-5. This library was designed after a combination of computational, synthetic, and biochemical approaches: *in silico* substrate modelling, manual Fmoc-SPPS and *in vitro* substrate testing with TS5-5. The identification of the hit peptide sequences from library screening, as well as their kinetic parameters and selectivity for ADAMTS-5 over the major cleaving metalloproteinases in Figure 4.2 (*i.e.* MMP-2 Δ C and TS4-4) will be covered in the next chapter.

Chapter 5. The identification and characterisation of peptide hits from ADAMTS-5 combinatorial library screening

5.1. Introduction

Chapter 4 described the computational docking of FRET peptide substrates, their synthesis through Fmoc-SPPS and subsequent testing for on-resin cleavage by ADAMTS-5. The most rapidly cleaved peptide KY(NO₂)TESESRGK(Abz)IYYKKG formed the basis of a solid-phase combinatorial library, which was synthesised by split-mix synthesis. After incubation of the library with ADAMTS-5 for one hour, 20 fluorescent hit beads were removed manually under a fluorescent microscope. The peptides were cleaved from the resin beads and sequencing performed by MALDI and tandem mass spectrometry.

This chapter focuses on the identification of peptide sequences from hit beads using manual *de novo* peptide sequencing, which involved a combination of MALDI and tandem mass spectrometry. This was followed by their synthesis and purification prior to biochemical characterisation, which involved: i) confirming solution-phase cleavage by ADAMTS-5; ii) determining kinetic parameters for cleavage by ADAMTS-5; iii) determining selectivity for ADAMTS-5 over ADAMTS-4, MMP-2 and MMP-9; and iv) comparing the sensitivity of the most rapidly cleaved and selective peptide substrate(s) to the existing literature substrate for detection of ADAMTS-5 activity.

In the initial step, *de novo* peptide sequencing was used to identify peptide hits from mass spectra through two pieces of computational software: Bruker Compass DataAnalysis and the custom-made LibMSCalc (183). The former software determined the peptide sequence from the mass spectra, whereas the latter software established how well the data matched the determined peptide sequence. Peptides hits were then synthesised by Fmoc-SPPS and purified by HPLC.

To confirm that peptide hits were cleaved by ADAMTS-5 in solution, they were incubated with TS5-5 and the rate of cleavage measured over time. Assuming substrate cleavage followed Michaelis-Menten kinetics (214), a more detailed analysis was performed on the parameters that contributed to this cleavage rate, namely k_{cat} and K_{m} . Briefly, the k_{cat} value describes the rate of conversion of substrate to product whereas the K_{m} value describes the strength of peptide binding to the active site of the proteinase (206). Given that conservative amino acid changes were employed in the design of the combinatorial library, significant improvements in k_{cat} and K_{m} values (*i.e.* ≥ 10 -fold) were not expected. Taken together however, these kinetic parameters would provide a more detailed explanation for any differences between the overall cleavage rates of the hit peptides by ADAMTS-5.

To assess the level of selectivity that each hit peptide had for ADAMTS-5, cleavage was also determined by TS4-4, MMP-2 Δ C and MMP-9 Δ C. This is because the former two proteinases were found to cleave the existing ADAMTS-5 literature substrate most rapidly (Figure 4.2), making any reduction in cleavage of hit sequences by these proteinases a priority. In addition, MMP-9 Δ C was chosen as it is part of the same gelatinase family as MMP-2 Δ C.

Finally, the most selective and rapidly cleaved substrate(s) were evaluated for their ability to detect ADAMTS-5 activity at low nanomolar concentrations using a low micromolar concentration of substrate.

To sum up, the objectives of this chapter were to: i) identify hit peptide substrates through mass spectrometry; ii) synthesise these peptides by Fmoc-SPPS; iii) purify and characterise these peptides by HPLC-MS; iv) determine the rate of substrate cleavage by ADAMTS-5 in solution; v) assess the selectivity of these substrates for ADAMTS-5; and vi) determine the detection sensitivity of the best hits for ADAMTS-5 activity.

5.2. Results

5.2.1. Identification of peptide hits from MALDI and tandem mass spectra using computational software

The general structure of peptide hits was KY(NO₂)XXXESXXK(Abz)XXXXKG, where “fixed” residues were indicated by their one-letter code and “X” represented one of several possible amino acids from the combinatorial library. To obtain the mass of the quasi-molecular ion for each peptide hit, an initial MALDI spectrum was acquired. Figure 5.1A shows a representative MALDI spectrum of a peptide hit displayed in Bruker Compass DataAnalysis, with the quasi-molecular ion indicated by [M + H]⁺ (2038.99954 u).

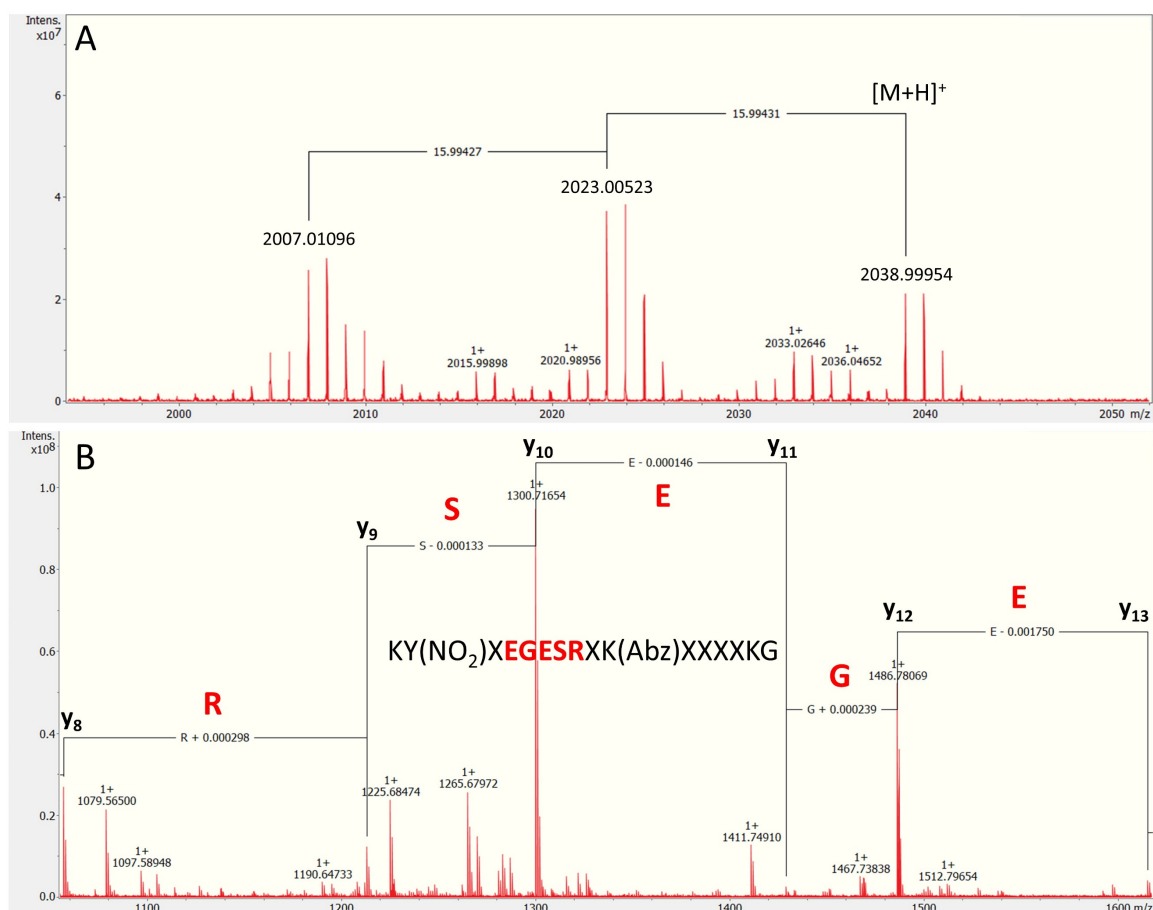


Figure 5.1. MALDI and tandem mass spectra of a peptide hit from ADAMTS-5 library screening. (A) Representative MALDI spectrum of a peptide hit showing the sequential loss of oxygen atoms (-16 u) from the quasi-molecular ion $[M + H]^+$ (2038.99954 u). (B) A representative tandem mass spectrum following fragmentation at 2023.00523 u, where the quasi-molecular ion has lost an oxygen atom from the nitrotyrosine moiety, *i.e.* $[M - O + H]^+$. Please note that only a portion of the spectrum is shown between ~1050-1600 m/z for clarity, with the residues lost from fragment ion peaks indicated in red single letter code. Each fragment ion peak that results from the loss of a residue forms part of the y-ion series and is assigned both a letter and number according to the criteria shown in Figure 1.13 (Chapter 1). Analysis of both spectra was performed in Bruker Compass DataAnalysis.

The quasi-molecular ion peak was identified by the characteristic loss of oxygen atoms from the nitrotyrosine moiety. The fragment ion that resulted from the loss of a single oxygen (*i.e.* $[M - O + H]^+$, 2023.00523 u) was fragmented further through collision-induced dissociation to give a tandem mass spectrum containing a y-ion series. Figure 5.1B shows part of the fragmentation pattern of the $[M - O + H]^+$ peak between ~1050-1600 m/z. Starting from the right-hand side of the spectrum, the loss of E from the y_{13} fragment ion at

1615.82154 u resulted in the y_{12} ion with mass 1486.78069 u; the loss of G gave the y_{11} ion (1429.75899 u); the loss of E the y_{10} ion (1300.71654 u), and so on until the y_8 fragment ion had been reached. By piecing together the amino acids that had been lost, the sequence that formed the core of the peptide was determined as EGESR. By continuing to subtract residues from different fragment ion peaks, the final peptide sequence was elucidated.

However, the signal intensity of the quasi-molecular ion in the initial MALDI spectrum determined whether further fragmentation by tandem mass spectrometry was possible. A sufficiently strong signal intensity was assessed on the basis of being not more than 10-fold lower than the most intense background ion. Therefore, of the 20 peptide hits that were analysed by MALDI after cleavage from resin beads, only 17 had sufficiently strong signals for further fragmentation. Of these 17 peptides, two were too convoluted to analyse in that it was not possible to establish a y -ion or b -ion series from the spectrum. Additionally, six peptides either had insufficient fragment ion peaks to determine a sequence or there was more than one possible sequence that could be determined from the fragmentation pattern. This left a total of 9 peptides, whose fragment ion mass lists were imported into LibMScalc and scored to conclusively confirm the peptide sequences. Table 5.1 shows the results from the mass spectrometry analysis and scoring as well as the parent compound used in combinatorial library synthesis. Appendix Figures 8.14–8.31 display the MALDI and tandem mass spectra for each of the peptide hits.

Table 5.1. Peptide hits identified through MALDI and tandem mass spectrometry

Compound	Sequence [†]	HRMS (calc.) [M+H] ⁺	HRMS (found) [M+H] ⁺	Mass difference	Score
5*	KY (NO ₂) TESESRGK (Abz) IYYKKG	N/A	N/A	N/A	N/A
31	KY (NO ₂) SE GESRGK (Abz) JYF KKG	2038.99867	2038.99954	± 0.00088	10,080
32	KY (NO ₂) TE N ESRGK (Abz) IYYKKG	2128.04634	2128.04349	± 0.00285	4189
33	KY (NO ₂) QD SE SBAK (Abz) JHYU KG	2001.94188	2001.94370	± 0.00182	5432
34	KY (NO ₂) TE G ESRGK (Abz) JZ YKKG	2059.96599	2059.96713	± 0.00114	7025
35	KY (NO ₂) TE T ES OAK (Abz) IY F KKG	2071.05003	2071.04741	± 0.00262	4382
36	KY (NO ₂) S ESESRGK (Abz) IYYKKG	2087.019795	2087.01731	± 0.00249	9772
37	KY (NO ₂) NDT ES OAK (Abz) AHFU KG	1959.93132	1959.93098	± 0.00034	5669
38	KY (NO ₂) NE SE S KAK (Abz) VH YKKG	2060.02013	2060.01896	± 0.00117	4948
39	KY (NO ₂) SE N ESRGK (Abz) IYYKKG	2114.03061	2114.02730	± 0.00331	3258

* Parent compound for combinatorial library synthesis. Key: O = ornithine, B = 2,4-Diaminobutyric acid, J = β-cyclopropyl-alanine, Z = 4-Thiazolyl-alanine, U = 2,3-Diaminopropionic acid. [†]Residues that differ from the parent compound in that position are bolded in red.

The differences between the theoretical and experimental masses for each peptide sequence were well within ± 0.02 u (< 10 ppm), which is considered an acceptable error for assigning both the quasi-molecular ion for each peptide hit as well as the difference in mass between the fragment ion peaks (*i.e.* the peak-to-peak relationship) in the tandem mass spectra (215).

All peptide hits in Table 5.1 were identified using a combination of Bruker Compass DataAnalysis and LibMSCalc. The only exception was peptide **39**. Here, the sequence identified through the fragmentation pattern (Appendix Figure 8.31) gave a peptide of mass 2114.02730 u, which corresponds to compound **39** in Table 5.1. However, the relationship between this mass (2114.02730 u) and the mass of the ion that underwent fragmentation to deliver the tandem mass spectrum for peptide **39** (*i.e.* 2124.04884 u) cannot be explained. Nevertheless, given that the mass of 2114.02730 u matches the fragmentation pattern, is present in the initial MALDI spectrum as a quasi-molecular ion (Appendix Figure 8.30) and is also a possible peptide from the combinatorial library, peptide hit **39** was included in Table 5.1. In addition, peptide **39** also completes a series of sequential single amino acid modifications at the N-terminus (*i.e.* -SENE-) based on peptide **5** (-TESE-), **32** (-TENE-) and **36** (-SESE-). This series is important as it enables structure-activity relationships to be probed should substrate cleavage by ADAMTS-5 be detected.

5.2.2. Synthesis of peptide hits by Fmoc-SPPS

As a consequence of time constraints resulting from the significant amount of time spent manually determining the peptide sequences in the previous section, the current literature substrate, the library parent compound **5** and hit peptides **31–39** were outsourced to GL Biochem which synthesised and purified them. The HPLC

chromatograms and mass spectra for these peptides were kindly provided and can be found in the Appendix (section 8.3.2). Each peptide was found to have a purity $\geq 95\%$ by HPLC and the mass was confirmed through HPLC-MS.

5.2.3. Evaluation of purified peptide hits for cleavage by ADAMTS-5

The current literature substrate Abz-TESE~SRGAIY-Dpa-KK-NH₂ (**2**), the library parent compound **5** and peptide hits **31–39** were tested for cleavage by the TS5-5 proteinase in solution. Although cleavage of **2** was described in Chapter 3, it was purchased from a different company (Severn Biotech). To minimise the possibility of discrepancies between the cleavage rates of the same substrate from different companies, **2** from GL Biochem was tested directly alongside peptides **5**, **31–39** to ensure consistency of data and results. In all cases, rates were determined in the same way as described in Figure 4.1 of Chapter 4 by accounting for the concentration of TS5-5 and expressing them in $\mu\text{M s}^{-1}$ using the calibration curve from Figure 8.6C. Representative rate curves for the cleavage of each substrate can be found in the Appendix (Figure 8.32). Final rates were expressed as the mean $\pm$ SEM from $n = 3\text{--}4$ individual experiments and are shown in Figure 5.2.

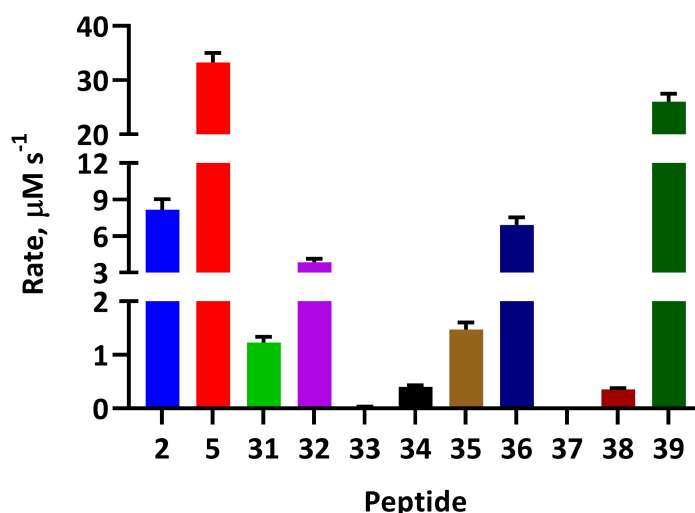


Figure 5.2. Cleavage of the current substrate, parent compound and peptide hits by the ADAMTS-5 proteinase. Rate data for cleavage of the current literature substrate 2, library parent compound 5 and hit peptides 31–39 by TS5-5. The rate of cleavage of each peptide substrate (20 μM) by TS5-5 is expressed as the mean ± SEM for $n = 9$ –12 wells from 3–4 independent experiments performed in triplicate at 37 °C. Peptides are numbered and labelled according to Table 5.1.

Compared to the current literature substrate, the library parent compound **5** showed the highest improvement in overall cleavage by TS5-5 (*cf.* $33.29 \pm 1.74 \mu\text{M s}^{-1}$ to $8.18 \pm 0.86 \mu\text{M s}^{-1}$, ~4-fold vs current). This indicated that the computational modelling of **5** into the active site of ADAMTS-5 described in Chapter 4 had been successful. Amongst the hit peptides identified by combinatorial library screening, the highest rate of substrate cleavage was found for **39** ($26.05 \pm 1.45 \mu\text{M s}^{-1}$), which showed an ~3-fold improvement in cleavage by TS5-5 compared to the current substrate. This was followed by **36** ($6.91 \pm 0.63 \mu\text{M s}^{-1}$) and **32** ($3.86 \pm 0.29 \mu\text{M s}^{-1}$). Compared to **5** (-TESE~) the latter two peptides differed by a single amino acid residue, *i.e.* for **36**, Thr → Ser in P4 (-SESE~) and for **32** Ser → Asn in P2 (-TENE~), whereas **39** was a combination of both of these changes (-SENE~).

When Thr in the P4 position of **5** was replaced with Ser to give **36**, or Ser in P2 was replaced with Asn to give **32**, ~5-fold and ~9-fold decreases in the cleavage rate were observed, respectively. In contrast, when both of these changes were combined to give **39**, a smaller ~1.3-fold decrease in overall cleavage rate by TS5-5 was seen. Peptide hits that varied by 4–5 amino acid residues from the parent compound **5** had far lower rates of cleavage by TS5-5: $1.48 \pm 0.13 \mu\text{M s}^{-1}$ for **35** and $0.36 \pm 0.024 \mu\text{M s}^{-1}$ for **38** for example. Where more than one un-natural amino acid was introduced into the peptide sequence, as in the case of **37** (O and U) and **33** (B, J and U), the lowest cleavage rates were observed: $2.81 \times 10^{-3} \pm 3.78 \times 10^{-4} \mu\text{M s}^{-1}$ and $2.44 \times 10^{-2} \pm 6.23 \times 10^{-3} \mu\text{M s}^{-1}$, respectively.

5.2.4. Determination of kinetic parameters for cleavage of peptide hits by ADAMTS-5

To analyse the results in Figure 5.2 in more detail, the cleavage of each peptide by TS5-5 was assumed to follow Michaelis-Menten kinetics (214). The parameters that describe this model of enzyme kinetics—namely the k_{cat} and K_{m} values—were thus determined. Initially however, the most up-to-date concentration of TS5-5 was determined by active-site titration with CT-1746 using the procedure described in Chapter 2 (section 2.1.2.16) to give curves similar to those shown in Chapter 3. This was to account for any change in the concentration of the TS5-5 proteinase following long-term storage (> 2 years) at - 80 °C.

To determine the k_{cat} and K_{m} for each peptide, the rate of cleavage was determined at a fixed concentration of TS5-5 but at a series of different substrate concentrations. The current literature substrate, the library parent compound **5** and hit peptides **32**, **36** and **39** were analysed. Hit peptides **31**, **33**, **34**, **35**, **37** and **38** were not, owing to low rates of cleavage by TS5-5 that would necessitate the use of large quantities

of substrate to determine these parameters. For each peptide that was examined, the rates at the different substrate concentrations were fitted to the Michaelis-Menten equation. A representative curve for the current substrate is shown in Figure 5.3.

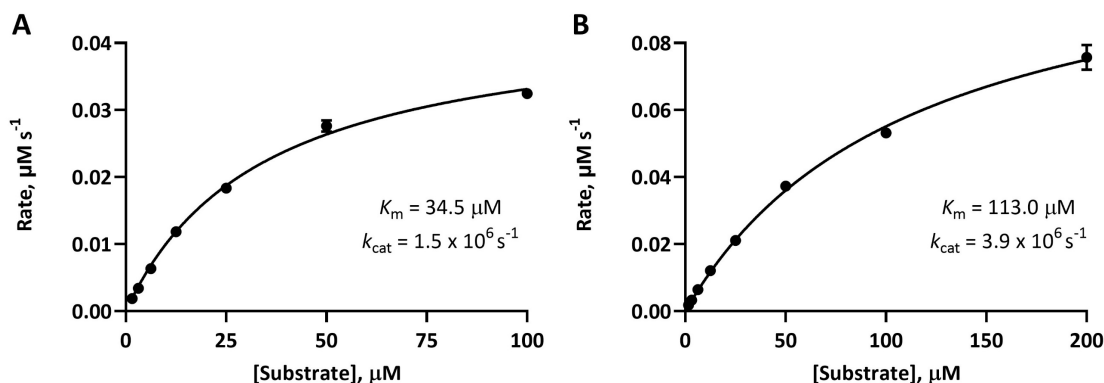


Figure 5.3. Kinetics of substrate cleavage by the ADAMTS-5 proteinase. Representative Michaelis-Menten saturation curves showing cleavage of Abz-TESE~SRGAIY-Dpa-KK-NH₂ at different concentrations by a fixed concentration of TS5-5 (30 nM). (A) Rate data prior to inner-filter correction; (B) rate data following inner-filter correction. The rate at each concentration is expressed in $\mu\text{M s}^{-1} \pm \text{SEM}$ for $n = 3$ wells from a single experiment performed in triplicate at 37 °C.

Fitting the rate data at different substrate concentrations to the Michaelis-Menten equation and utilising the concentration of TS5-5 delivered a $K_m = 34.5 \mu\text{M}$ and $k_{\text{cat}} = 1.5 \times 10^6 \text{ s}^{-1}$, respectively (Figure 5.3A). However, when determining accurate kinetic parameters, the inner-filter effect must be considered. This effect is most pronounced at substrate concentrations $> 20 \mu\text{M}$ (216). At these concentrations, upon cleavage of the FRET substrate and separation of the fluorophore and quencher, the fragment containing the “free” fluorophore is substantially quenched by un-cleaved intact peptides or cleaved quencher fragments, reducing the amount of detected fluorescence. This decrease in detected fluorescence affects the linearity of the cleavage rate, which in turn results in less accurate kinetic parameters after fitting to the Michaelis-Menten equation (178, 200). Correction for the inner-filter effect can be carried out in two ways: a) dilution of

the sample until the fluorophore and quencher fragments are too far apart for quenching to occur; or b) through the use of an empirical method that involves preparing different substrate concentrations with and without a fixed concentration of free fluorophore. The latter method was used in line with a literature protocol (178) and is described in more detail in section 2.1.2.20 of Chapter 2. Following correction for the inner-filter effect, values of $K_m = 113.0 \mu\text{M}$ and $k_{\text{cat}} = 3.9 \times 10^6 \text{ s}^{-1}$ were obtained (Figure 5.3B). These higher values highlight the importance of inner-filter correction for determining accurate kinetic parameters for proteinase substrates.

A summary of corrected kinetic parameters for individual experiments of the current substrate **2**, library parent compound **5** and hit peptides **32**, **36** and **39** are displayed in Table 5.2 with the mean $\pm$ SEM for each dataset. Individual saturation curves for each experiment both before and after inner-filter correction can be found in the Appendix (Figures 8.33–8.37).

Table 5.2. Kinetic parameters for cleavage of the current substrate, library parent compound and selected hit peptides by the ADAMTS-5 proteinase.

Entry	Peptide	$k_{\text{cat}}, \times 10^6 \text{ s}^{-1}$	$K_{\text{m}}, \mu\text{M}$	$k_{\text{cat}}/K_{\text{m}}, \times 10^3 \mu\text{M}^{-1} \text{ s}^{-1}$
2	Abz-TESE~SRGAIY-Dpa-KK-NH ₂	3.9	113.0	34.6
		4.1	132.1	30.7
		2.9	130.6	21.8
		3.6 ± 0.3	125.2 ± 5.0	29.1 ± 3.1
5	KY(NO ₂)TESESRGK(Abz)IYYKKG	1.8	29.9	59.6
		2.2	32.8	67.3
		1.6	25.6	62.6
		1.9 ± 0.1	29.5 ± 1.7	63.2 ± 1.8
32	KY(NO ₂)TENESRGK(Abz)IYYKKG	0.51	36.6	14.0
		0.57	30.4	18.8
		0.53	29.7	17.7
		0.50	25.5	19.6
		0.53 ± 0.014	30.5 ± 2.0	17.5 ± 1.1
36	KY(NO ₂)SESESRGK(Abz)IYYKKG	1.2	61.7	18.7
		0.9	75.6	11.7
		1.4	56.6	24.2
		1.1 ± 0.1	64.6 ± 4.6	18.2 ± 2.9
39	KY(NO ₂)SENESESRGK(Abz)IYYKKG	4.0	66.2	60.7
		2.7	72.2	37.9
		3.2	92.3	34.7
		3.3 ± 0.3	76.9 ± 6.4	44.4 ± 6.7

N.B. All bolded values represent mean ± SEM

Compared to the current literature substrate, the library parent compound **5** and hit peptides **32** and **36** had ~2-fold, ~7-fold and ~3-fold lower catalytic rate constants, respectively (*cf.* $3.6 \pm 0.3 \times 10^6 \text{ s}^{-1}$ to $1.9 \pm 0.1 \times 10^6 \text{ s}^{-1}$ for **5**; $0.53 \pm 0.014 \times 10^6 \text{ s}^{-1}$ for **32** and $1.1 \pm 0.1 \times 10^6 \text{ s}^{-1}$ for **36**). In contrast however, the k_{cat} value for **39** was similar to the

current literature substrate at $3.3 \pm 0.3 \times 10^6 \text{ s}^{-1}$. Nevertheless, neither **5** nor the peptide hits **32**, **36**, and **39** had better catalytic rate constants than the current substrate. This shows that upon formation of the ADAMTS-5-substrate complex, the conversion of substrate to product remained slower when compared to **2**.

On the other hand, the K_m value was found to decrease for the library parent compound and all hit peptide substrates. The current substrate was found to have a K_m value of $125.2 \pm 5.0 \text{ }\mu\text{M}$, which is higher than the $76 \text{ }\mu\text{M}$ value reported in the literature by Troeberg *et al* (62). For compound **5** and **32**, this value fell by ~ 4 -fold to $29.5 \pm 1.7 \text{ }\mu\text{M}$ and $30.5 \pm 2.0 \text{ }\mu\text{M}$, respectively. However, for **36** and **39** a smaller ~ 2 -fold decrease was observed, with values of $64.6 \pm 4.6 \text{ }\mu\text{M}$ and $76.9 \pm 6.4 \text{ }\mu\text{M}$ determined, respectively. The lower K_m values for **5** and **32** indicate stronger binding of the substrate to the ADAMTS-5 proteinase and greater stability of the ADAMTS-5-substrate complex.

Finally, the highest catalytic efficiencies (k_{cat}/K_m values) were recorded for peptides **5** ($63.2 \pm 1.8 \times 10^3 \text{ }\mu\text{M}^{-1} \text{ s}^{-1}$) and **39** ($44.4 \pm 6.7 \times 10^3 \text{ }\mu\text{M}^{-1} \text{ s}^{-1}$). This revealed that TS5-5 cleaved these substrates more effectively than the current literature substrate ($k_{\text{cat}}/K_m = 29.1 \pm 3.1 \times 10^3 \text{ }\mu\text{M}^{-1} \text{ s}^{-1}$) and verified the success of both computational modelling and combinatorial library screening for identifying novel substrates of ADAMTS-5 with improved cleavage rates and kinetic parameters.

However, to assess the selectivity of the hit peptides for ADAMTS-5, their rate of cleavage by TS5-5 was compared to a panel of metalloproteinases. This panel consisted of TS4-4 and MMP-2 Δ C, as they cleaved the current literature substrate most rapidly (see Figure 4.2, Chapter 4). MMP-9 Δ C was also included due to its close relation to MMP-2 Δ C. MMP-9 Δ C was readily available, and TS4-4 had been purified and characterised in

Chapter 3. However, because MMP-2ΔC was not readily available, it had to be expressed and purified first.

5.2.5. Expression and selection of proMMP-2ΔC for purification

HEK293 cells were used to obtain human MMP-2ΔC for selectivity studies. The expression and selection process were similar to that described for ADAMTS-4 and ADAMTS-5 in Chapter 3 (Scheme 3.1) and will be covered only briefly here. Cryopreserved HEK293 cells stably transfected with the pCEP4 expression vector encoding untagged proMMP-2ΔC (62) were thawed and cultured in DMEM under a range of hygromycin B concentrations (50 μg/ml – 800 μg/ml). A final concentration of 800 μg/ml hygromycin B was used, as this was successful with TS5-5 and is close to the reported literature value of 1000 μg/ml (62). After passaging to larger plates, medium was replaced with serum-free DMEM supplemented with lactalbumin enzymatic hydrolysate. The conditioned medium was collected and changed every 3–5 days. Centrifugation was used to remove the cell debris, from which the supernatant was decanted and filtered. This medium was frozen at -20 °C after the addition of 0.02% (w/v) sodium azide.

5.2.6. Purification of proMMP-2ΔC by gelatin-sepharose and native sephacryl affinity chromatography

ProMMP-2ΔC was purified at 4 °C from conditioned medium using a Gelatin-Sepharose 4B column on an ÄKTA prime FPLC. Figure 5.4 shows a representative zymogram that enables the detection of active MMP-2 through the degradation of gelatin, which is indicated by a loss in Coomassie staining. This figure also shows the outcome of the stepwise purification process for proMMP-2ΔC. Initially, the column was washed with 5% (v/v) DMSO TNC buffer, equilibrated with TNC buffer and the combined

flow-through collected (FT1, lane 3). The conditioned medium containing proMMP-2 Δ C was diluted 1:1 with dilution buffer consisting of 100 mM tris-HCl, pH 8.0, with 0.02% (w/v) NaN₃ and loaded onto the column (LD, lane 2) and flow-through collected (FT2, lane 4). Contaminants binding to the column were then removed by washing with 1% (v/v) DMSO TNC (W, lane 5). ProMMP-2 Δ C was finally eluted with 5% (v/v) DMSO TNC and the collected fractions immediately frozen at -20 °C to prevent degradation of the proteinase. Small aliquots of these fractions were collected and can be seen in lanes 10–15 (numbered 1, 5, 10, 15, 20 and 25). The gelatin zymogram reveals that both proMMP-2 Δ C and active MMP-2 Δ C successfully eluted from the column at ~44 kDa and ~34 kDa, respectively when compared to the positive controls of active MMP-2 Δ C at ~34 kDa (lane 8) and MMP-2 containing active proteinase at ~63 kDa and proMMP-2 Δ C at ~44 kDa (lane 6).

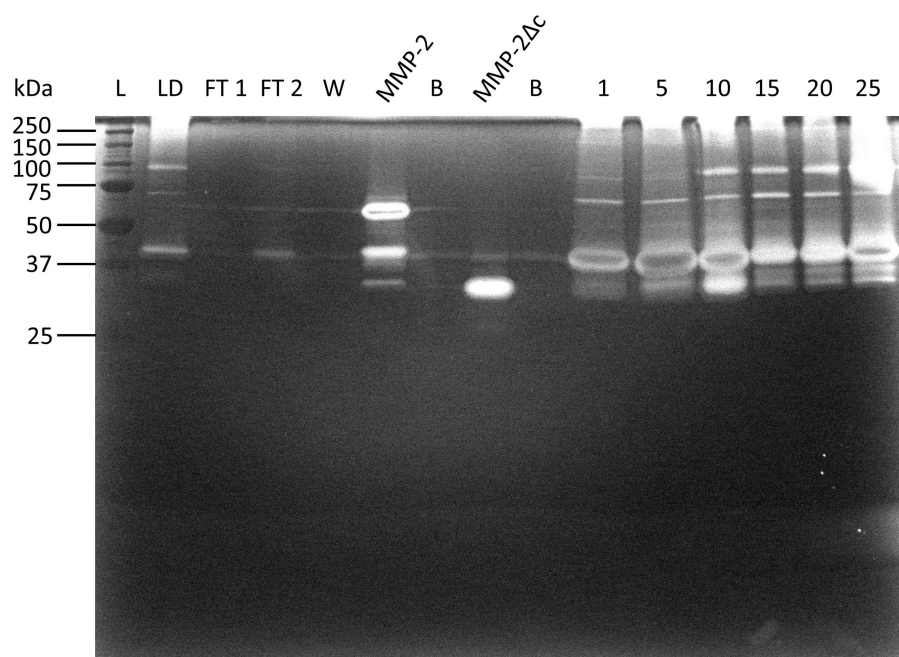


Figure 5.4. Representative affinity purification of proMMP-2ΔC through stepwise elution. Gelatin zymogram of eluted proMMP-2ΔC fractions visualised by Coomassie staining. L = ladder, LD = load, FT = flow-through, W = wash, MMP-2 = full-length, ΔC = catalytic domain and B = loading buffer. MMP-2 and MMP-2ΔC control fractions were loaded at 1.41 μg and 2.36 μg, respectively.

To remove DMSO from the sample fractions, a second purification was performed on an ÄKTA FPLC (UPC-900 monitor and P-920 pump) using a native sephacryl S-200 column. Initially, the collected proteinase fractions were thawed, pooled, concentrated to < 5 ml in a protein spin concentrator and an A_{280} reading taken. The native Sephacryl S-200 column was equilibrated with TNC buffer, the concentrated fractions loaded onto the column and the purification followed at a wavelength of 280 nm using Unicorn 5.2 software. Proteinase fractions that had eluted from the column were collected and frozen at -20 °C. Small aliquots were also collected and gelatin zymography performed to assess the activity of these samples and to decide which should be pooled (Figure 5.5A). Fractions C7–C9 were pooled as they contained only active MMP-2ΔC. They were concentrated in a protein spin concentrator and frozen

down for long-term storage at -80 °C in TNC buffer with 0.05% (v/v) Brij-35 and 10% (v/v) glycerol. Coomassie staining of the pooled fractions was performed to determine the purity of the proteinase. Figure 5.5B confirms that pure MMP-2 Δ C had been acquired at ~34 kDa.

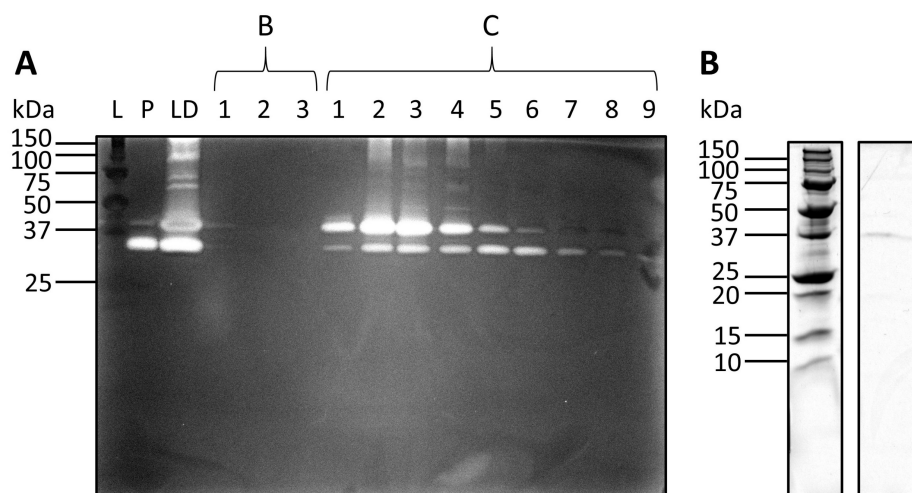


Figure 5.5. Representative purification of MMP-2 Δ C fractions using Sephacryl affinity chromatography. (A) Representative gelatin zymogram of eluted MMP-2 Δ C fractions visualised by Coomassie staining after removal of DMSO. L = ladder, P = MMP-2 Δ C, LD = load, Δ C = catalytic domain. Control MMP-2 Δ C was loaded at 2.36 μ g. *Please note that both “B” and “C” fraction groups were arbitrarily assigned by the purification software.* (B) Representative Coomassie stain of a purified MMP-2 Δ C fraction after S-200 Sephacryl purification. This pooled fraction was loaded at 0.45 μ g.

5.2.7. Validation of MMP-2 Δ C activity and determination of concentration through active-site titration with N-TIMP-3

To further validate the activity of the purified MMP-2 Δ C and determine the concentration of active proteinase, active-site titration was performed as previously described for both ADAMTS-4 and ADAMTS-5 proteinases (Figure 5.6).

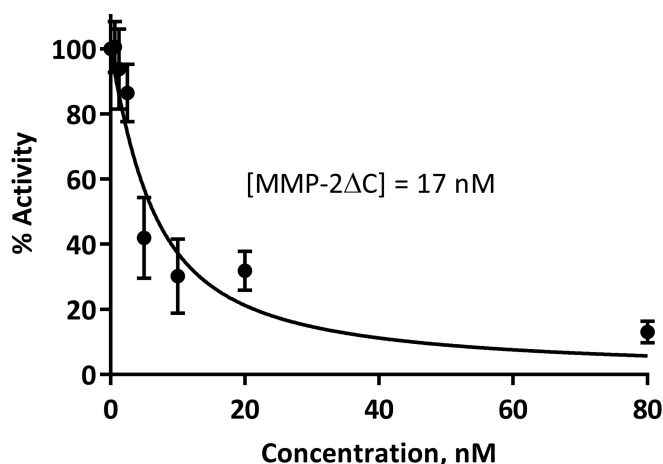


Figure 5.6. Determination of MMP-2 Δ C concentration through active-site titration with an endogenous TIMP-3 inhibitor. Representative curve for active-site titration of MMP-2 Δ C with N-TIMP-3. Rates are expressed as a function of percentage activity without inhibitor. Values are given as mean percentage activity $\pm$ SEM, with $n = 3$ wells used for each N-TIMP-3 concentration at 37 °C.

Extrapolation of the linear portion of the curve to the x-axis gave the concentration of active proteinase. Having established that pure and active MMP-2 Δ C had been obtained, this proteinase was used alongside TS4-4 and MMP-9 Δ C for the selectivity studies of the library parent compound and hit peptide substrates.

5.2.8. Determination of selectivity of peptide hits for ADAMTS-4, MMP-2 and MMP-9

To assess the selectivity of the hit peptides for ADAMTS-5, they were tested against TS4-4, MMP-2 Δ C and MMP-9 Δ C and rate data determined as described for Figure 5.2. The results of the selectivity studies are shown in Figure 5.7, with the individual rate curves shown in the Appendix (Figures 8.38–8.40). In addition, Figure 5.7A displays the selectivity data for TS5-5 again purely for the purpose of comparing it to the other proteinases.

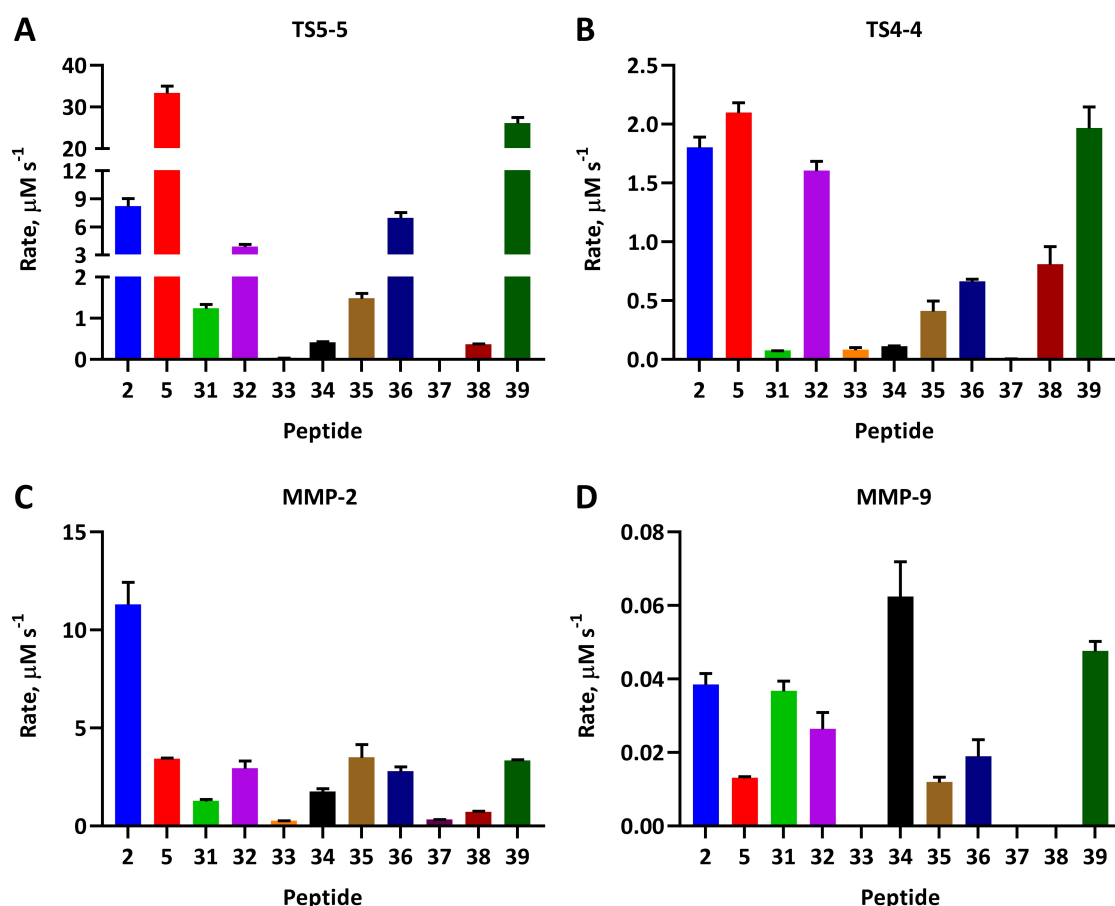


Figure 5.7. Selectivity studies of the current substrate, parent compound and hit peptides against ADAMTS and MMP proteinases. Rate data for cleavage of the current substrate **2**, library parent compound **5** and hit peptides **31–39** (all 20 μM) by (A) TS5-5, (B) TS4-4, (C) MMP-2 ΔC and (D) MMP-9 ΔC . The rate of cleavage of each metalloproteinase is expressed as the mean $\pm$ SEM for $n = 9\text{--}12$ wells from 3–4 independent experiments performed in triplicate at 37 °C. ΔC = catalytic domain. The selectivity of each substrate for ADAMTS-5 can be seen in Figure 5.2 and is included for comparison purposes.

In the case of TS4-4 selectivity (Figure 5.7B), both the library parent compound **5** and hit peptide **39** showed essentially identical rates of cleavage to that of the current substrate **2** (*cf.* $2.1 \pm 0.1 \mu\text{M s}^{-1}$ and $2.0 \pm 0.2 \mu\text{M s}^{-1}$ to $1.8 \pm 0.1 \mu\text{M s}^{-1}$) thereby revealing no reduction in cleavage by TS4-4. Furthermore, hit peptide **32** also showed a similar cleavage rate by TS4-4 to the aforementioned peptides ($1.6 \pm 0.1 \mu\text{M s}^{-1}$). All other peptides (*i.e.* **31**, **33**, **34** and **35–38**) showed lower rates of cleavage by TS4-4 compared to the current substrate.

Looking at MMP-2ΔC selectivity, the rate of substrate cleavage by this proteinase was not only reduced for library parent compound **5**, but also for all hit peptides **31–39**. More specifically, the rate decreased from $11.3 \pm 1.1 \mu\text{M s}^{-1}$ for **2** to between $3.4 \pm 6.5 \times 10^{-2} \mu\text{M s}^{-1}$ and $0.2 \pm 1.8 \times 10^{-2} \mu\text{M s}^{-1}$ when taking into account the highest and lowest rates of cleavage, respectively.

The selectivity study with MMP-9ΔC is shown in Figure 5.7D. This data is difficult to interpret as the cleavage rates are $< 0.1 \mu\text{M s}^{-1}$ for all substrates. In addition, the raw cleavage data in Appendix Figure 8.40 shows that in several cases, the measured fluorescence intensity is not above background fluorescence (peptides **5** and **35**). The data indicates that no substrate is cleaved well by MMP-9ΔC, with hit peptides **33**, **37** and **38** showing no signs of cleavage after 48 h at a final concentration of 30 nM.

Overall, the most selective substrates for ADAMTS-5 were found to be the library parent compound **5** and the hit peptide **39**. When comparing their rate of cleavage by TS5-5 to the other proteinases in Figure 5.7, peptide **5** showed a ~16-fold preference for cleavage by TS5-5 over TS4-4, an ~10-fold preference over MMP-2ΔC and an ~2561-fold preference over MMP-9ΔC. Similarly, peptide **39** revealed selectivity preferences of ~13-fold, ~8-fold and ~548-fold over TS4-4, MMP-2ΔC and MMP-9ΔC, respectively.

5.2.9. Determination of peptide 5 and peptide 39 sensitivity for ADAMTS-5

Having discovered that the library parent compound **5** and hit peptide **39** were the most selective substrates for ADAMTS-5, the sensitivity of these probes for detection of ADAMTS-5 activity was evaluated. Initially, a series of substrate concentrations ranging from 20 μM to 0.63 μM were prepared and the rate of substrate cleavage determined over the course of one hour at a fixed concentration of TS5-5. The aim of

this experiment was to determine the lowest concentration of substrate that would be comfortably cleaved by TS5-5 above the level of the background fluorescence intensity of the substrate in the absence of TS5-5 (*i.e.* $\Delta\text{RFU} > 300$). The results are shown in Figure 5.8A and 5.8B for **5** and **39**, respectively.

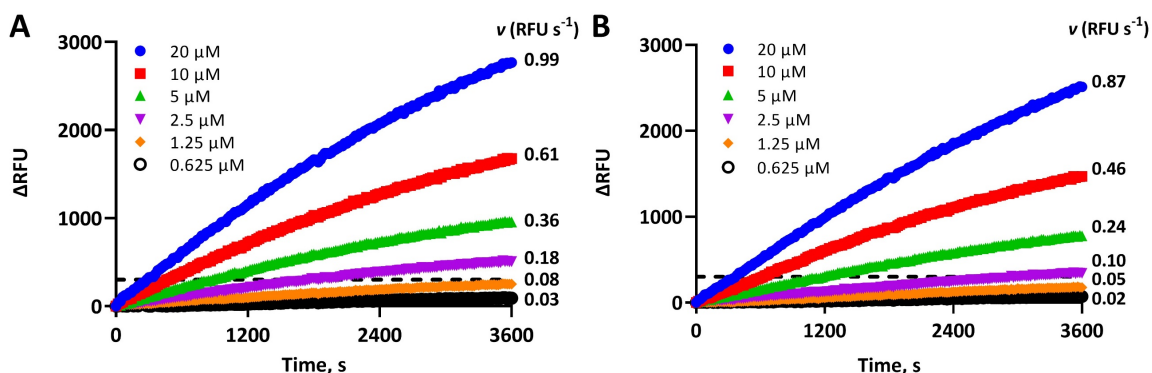


Figure 5.8. Comparison between the cleavage of library parent compound **5 and hit peptide **39** at different concentrations of substrate.** Rate data for cleavage of parent compound **5** (A) and hit peptide **39** (B) at a range of substrate concentrations (20 μM – 0.63 μM) but at a fixed concentration of TS5-5 (57 nM). The rate (v) is given in RFU s^{-1} and exhibits a linear increase with each rise in substrate concentration. Each time point is expressed as ΔRFU ($\text{RFU at time } t - \text{RFU at time } 0$) $\pm$ SEM. For each substrate concentration, a single experiment with $n = 3$ wells was performed at 37 °C. The dotted line indicates $\Delta\text{RFU} = 300$.

Figure 5.8 shows that for cleavage of both **5** and **39** at 57 nM TS5-5, substrate concentrations of 5 μM and 2.5 μM can be used to detect cleavage rates with $\Delta\text{RFU} \geq 300$ after 1 hour. However, in order to further probe how sensitive each of these FRET substrates was at detecting TS5-5 activity, low nanomolar concentrations of TS5-5 (1 nM, 0.5 nM, 0.25 nM and 0.125 nM) were combined with a 5 μM concentration of **5** and **39**. In addition, the current literature substrate **2** was also analysed at 5 μM with these TS5-5 concentrations to establish whether detection sensitivity had been improved (Figure 5.9).

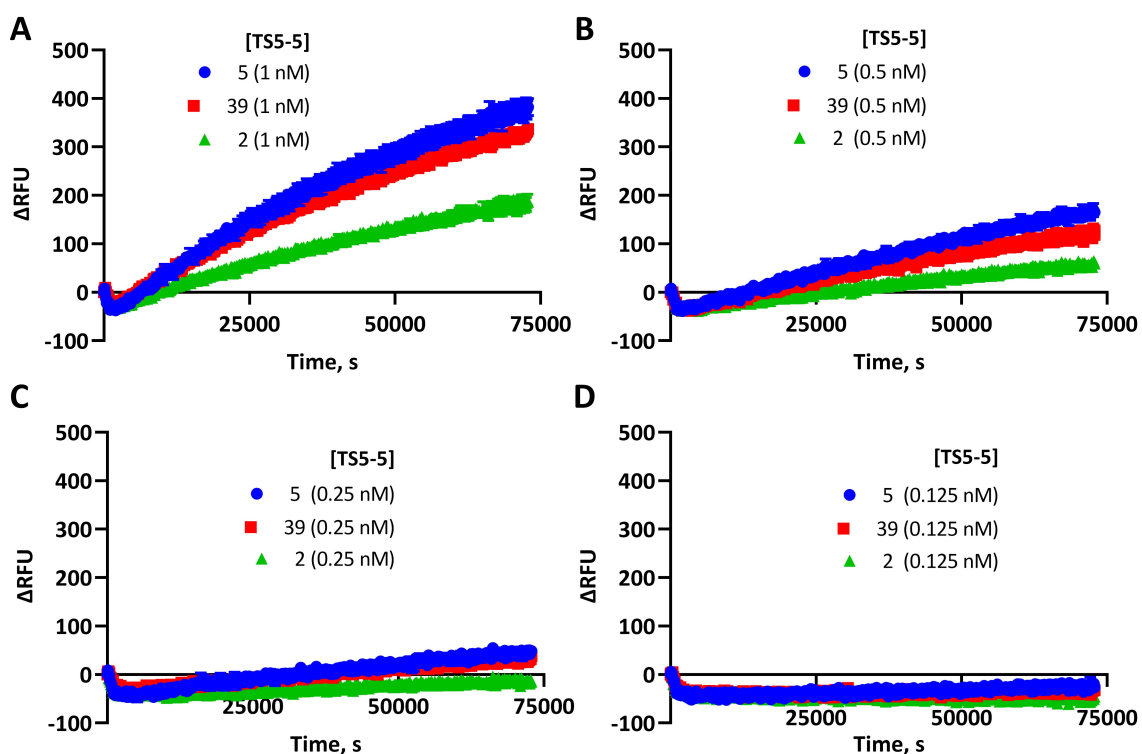


Figure 5.9. Probing the lowest concentration of substrate that can detect ADAMTS-5 activity. Rate data for cleavage of library parent compound **5**, hit peptide **39** and the current literature substrate **2** at a fixed concentration of substrate (5 μ M) with (A) 1 nM (B) 0.5 nM (C) 0.25 nM and (D) 0.125 nM concentrations of TS5-5 over 20 h. Each time point is expressed as Δ RFU (RFU at time t – RFU at time 0) $\pm$ SEM. For each concentration of TS5-5, a single experiment with $n = 3$ wells was performed at 37 $^{\circ}$ C.

As can be seen in Figure 5.9A, peptides **5**, **39** and **2** are able to detect ADAMTS-5 activity above the level of background fluorescence at a 1 nM concentration of TS5-5 after 20 h (Δ RFU > 100). Yet when the concentration is decreased to 0.5 nM TS5-5, only peptides **5** and **39** remain able to maintain a Δ RFU > 100 after 20 h, as Δ RFU = 59 for **2**. The poor sensitivity of **2** becomes apparent when 0.25 nM TS5-5 is used (Figure 5.9C), as it is no longer able to detect TS5-5 activity above background fluorescence up to 20 h when compared to **5** and **39**, where Δ RFU > 30. Neither of the probes **5**, **39** and **2** are able to detect TS5-5 activity at 0.125 nM above background fluorescence up to 20 h. However, this may become possible for peptide **5** after 20 h as substrate cleavage appears to continue increasing in a linear fashion.

A final experiment using a substrate concentration 2.5 μM for **5**, **39** and **2** at TS5-5 concentrations of 1 nM, 0.5 nM, 0.25 nM and 0.125 nM was performed (Figure 5.10). Only peptides **5** and **39** were able to detect TS5-5 activity above background ($\Delta\text{RFU} > 50$) at a concentration of 1 nM for the duration of the experiment (20 h). Although activity was detectable above background for **5** using 0.5 nM TS5-5 (Figure 5.10B), it is unlikely that a longer experimental duration would improve the ΔRFU as the cleavage rate can be seen to be plateauing on either side of the x-axis when $y = 0$. No cleavage was observed at 0.25 nM or 0.125 nM concentrations of TS5-5 with either of the three substrates evaluated here (Figure 5.10C and 5.10D).

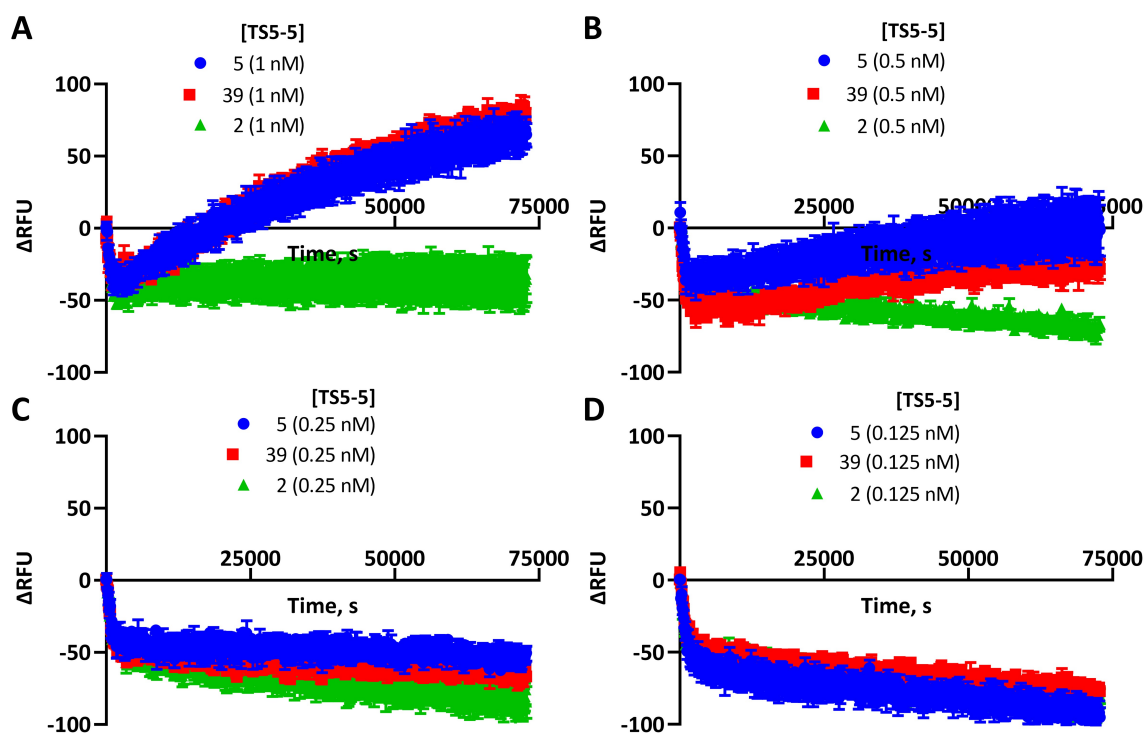


Figure 5.10. Further investigation into the lowest concentration of substrate that can detect ADAMTS-5 activity. Rate data for cleavage of library parent compound **5**, hit peptide **39** and the current literature substrate **2** at a fixed concentration of substrate (2.5 μM) with (A) 1 nM (B) 0.5 nM (C) 0.25 nM and (D) 0.125 nM concentrations of TS5-5 over 20 h. Each time point is expressed as ΔRFU ($\text{RFU at time } t - \text{RFU at time } 0$) $\pm$ SEM. For each concentration of TS5-5, a single experiment with $n = 3$ wells was performed at 37 $^{\circ}\text{C}$.

In summary, the library parent compound **5** and hit peptide **39** are able to detect ADAMTS-5 activity at concentrations as low as 0.25 nM using 5 μ M substrate, or 1 nM using 2.5 μ M substrate for experiments performed for ~20 h. These substrates thus display improved detection sensitivity over the current literature substrate **2** in addition to an enhanced cleavage rate and better overall selectivity for ADAMTS-5.

5.3. Discussion

This chapter described the identification of hit peptides **31–39** from fluorescent library beads through a combination of MALDI and tandem mass spectrometry. Along with the library parent compound **5**, these hit peptides were synthesised and characterised by a chemical company and their cleavage by TS5-5 evaluated in solution. More detailed kinetic parameters for cleavage by TS5-5—including the catalytic rate constant and Michaelis constant—were also calculated. Selectivity studies with TS4-4, MMP-2 Δ C and MMP-9 Δ C were used to determine the level of selectivity of substrates **5** and **31–39** for ADAMTS-5. Finally, the lead substrates **5** and **39** were evaluated for the detection of ADAMTS-5 activity at low nanomolar concentrations of proteinase.

The identification of hit peptide sequences from spectra acquired through MALDI and tandem mass spectrometry was performed using the commercially available Bruker Compass DataAnalysis software along with the in-house software LibMSCalc (183). The former software was used to identify the peptide sequence through the fragmentation pattern of b-ions and y-ions found in the tandem mass spectrum of each hit peptide. Unfortunately, this software has several limitations. Firstly, it does not gauge how probable the identified sequence is when constructed from the fragmentation pattern. For cases where a complete y-ion or b-ion series is available this does not present a

problem. However, complete fragmentation was only possible for one sequence (*i.e.* hit peptide **31**); whereas in several cases multiple fragmentations were observed where more than one possible series of amino acids could be present within that sequence. A case in point was hit peptide **39**, where both SENE and TDNE N-terminal fragment ion series were observed, making it difficult to determine which of these series belonged to the mass of the quasi-molecular ion. This was where LibMSCalc proved invaluable. By inputting the mass of the quasi-molecular ion into a pre-defined library and importing an experimental mass list of fragment ions, the software was able to match each theoretical peptide sequence and mass list of fragment ions to the experimental mass list. A score was then given for how well the experimental mass list fit the theoretical mass list and thus the peptide sequence of the inputted mass of the quasi-molecular ion. The higher the score, the more likely the listed peptide sequence(s) matched the experimental mass list of fragment ions. For example, a score of 10,000 indicated an excellent fit. In the case of hit peptide **39**, the N-terminal SENE sequence was given a higher score over TDNE, as the fragment ion peaks from the experimental mass list had a higher intensity than those from TDNE. Consequently, the SENE sequence was taken as the N-terminal portion of hit peptide **39**. In the absence of this software, generating a preference for either the SENE or TDNE sequence would only be possible by analysing the intensity of each of the different experimental fragment ion peaks, which would be time-consuming and prone to errors. The combination of these tools was thus critical to the analysis of the mass spectrometry data and the ability to definitively identify the final peptide sequences.

A structure-activity relationship was found for the hit peptide sequences **32**, **36** and **39**, which are related to the library parent compound **5** through single amino acid changes at one of the four N-terminal amino acids following KY(NO₂). Specifically, the

amino acids involved are -TESE~ (**5**), -TENE~ (**32**), -SESE~ (**36**) and -SENE~ (**39**). An ~9-fold decrease in cleavage rate by TS5-5 was observed in going from **5** to **32** through the replacement of Ser with Asn in the P2 position (*cf.* $33.29 \pm 1.74 \mu\text{M s}^{-1}$ to $3.86 \pm 0.29 \mu\text{M s}^{-1}$). In contrast, an ~5-fold decrease in cleavage rate was found when moving from **5** to **36**, where Ser was retained in P2 while Thr was exchanged with Ser in the P4 position (*cf.* $33.29 \pm 1.74 \mu\text{M s}^{-1}$ to $6.91 \pm 0.63 \mu\text{M s}^{-1}$). The larger decrease in rate observed for **32** could be due to the weakening of an important hydrogen bonding interaction in the P2 position. The replacement of the hydroxyl group in Ser with an amide group in Asn resulted in a weaker hydrogen bond donor, as the lone pair on the nitrogen atom of the amide is delocalised into the carbonyl group, thereby reducing its electronegativity. In contrast, for **36**, the hydroxyl functional group was maintained as Ser was kept in the P2 position and the replacement of Thr with Ser in P4 was a conservative change, likely explaining the smaller decrease in cleavage rate when **5** was converted to **36** compared to when it was converted to **32**. When the changes to **5** that generated both **32** and **36** were combined to give **39** (Ser → Asn in P2 and Thr → Ser in P4), a small ~1.3-fold decrease in cleavage rate was found (*cf.* $33.29 \pm 1.74 \mu\text{M s}^{-1}$ to $26.05 \pm 1.45 \mu\text{M s}^{-1}$), suggesting a possible synergistic effect between Ser and Asn that prevented the cleavage rate from dropping to that observed with **32** or **36**.

Looking at the cleavage of peptides **5**, **32**, **36** and **39** by TS5-5 along with the current substrate **2** in more detail revealed somewhat of a pattern with regard to the catalytic rate constant and Michaelis constant. Compared to **2**, the catalytic rate constant is shown to decrease ~2-fold and ~7-fold for peptides **5** and **32**, respectively (*cf.* $3.6 \pm 0.3 \times 10^6 \text{ s}^{-1}$ to $1.9 \pm 0.1 \times 10^6 \text{ s}^{-1}$ and $0.53 \pm 0.014 \times 10^6 \text{ s}^{-1}$). In contrast, the Michaelis constant was found to also decrease ~4-fold for both of these peptides (*cf.* $125.2 \pm 5.0 \mu\text{M}$ to 29.5

$\pm 1.7 \mu\text{M}$ and $30.5 \pm 2.0 \mu\text{M}$ for **5** and **32**, respectively). The catalytic efficiencies ($k_{\text{cat}}/K_{\text{m}}$'s) of **5** and **32** were $63.2 \pm 1.8 \times 10^3 \mu\text{M}^{-1} \text{s}^{-1}$ and $17.5 \pm 1.1 \times 10^3 \mu\text{M}^{-1} \text{s}^{-1}$, with **5** being ~ 2 -fold higher than the current substrate ($29.1 \pm 3.1 \times 10^3 \mu\text{M}^{-1} \text{s}^{-1}$). This shows that although slower conversion of substrate to product is occurring compared to the current substrate (*i.e.* lower k_{cat}), stronger binding of **5** and **32** to the active site of ADAMTS-5 (*i.e.* lower K_{m}) is driving substrate cleavage; and in the case of **5** improves the overall cleavage rate by TS5-5.

On the other hand, peptides **36** and **39** have only ~ 2 -fold improvements in K_{m} compared to **2** (*cf.* $64.6 \pm 4.6 \mu\text{M}$ and $76.9 \pm 6.4 \mu\text{M}$ to $125.2 \pm 5.0 \mu\text{M}$), with peptide **39** having a comparative k_{cat} to the current substrate (*cf.* $3.3 \pm 0.3 \times 10^6 \text{s}^{-1}$ to $3.6 \pm 0.3 \times 10^6 \text{s}^{-1}$). Perhaps in the case of these peptides, the k_{cat} value drives substrate cleavage by TS5-5 more than the K_{m} , given that peptide **39** displays the second highest $k_{\text{cat}}/K_{\text{m}}$ value, which is ~ 1.5 -fold higher than the current substrate (*cf.* $44.4 \pm 6.7 \times 10^3 \mu\text{M}^{-1} \text{s}^{-1}$ to $29.1 \pm 3.1 \times 10^3 \mu\text{M}^{-1} \text{s}^{-1}$).

The difference in the K_{m} value of the current substrate ($125 \mu\text{M}$) relative to that described in the literature ($76 \mu\text{M}$) (62) could be a consequence of peptide purity or differences in the measurements of cleavage rates from different microplate readers. In addition, the value reported in the literature is from a single experiment, whereas here it is from the mean of three experiments performed in triplicate. As a result of a more rigorous determination of K_{m} , the $125 \mu\text{M}$ value is likely to be more accurate.

The use of the MALDI method to identify peptide sequences from hit beads suffers from a major caveat. Even if a small combinatorial is used for library screening (*e.g.* $< 10,000$ unique peptide sequences), there is typically a 50% attrition rate for hits.

In other words, only 50% of sequences from hit beads are identified in a study.¹⁴ In the case of this study, this value was 45% (9 out of 20). The issue arises from the inherent difficulty in obtaining peptide sequence data from a single bead for MALDI analysis. This can arise due to the peptide “sticking” to the bead during the extraction process or the final signal of the quasi-molecular ion in the MALDI spectrum being too weak to perform tandem mass spectrometry to determine the peptide sequence. Having identified the final peptide sequence however, there is also no guarantee that it will be cleaved by the proteinase in solution. This is because cleavage on-resin does not consistently translate to cleavage in solution, the reason for which is not currently known. One possibility may be the unique microenvironment within a PEGA bead that may induce more favourable cleavage. This could be due to the presence of basic residues on a substrate within each bead that favour hydration by an aqueous buffer, and thus facilitates cleavage by the ADAMTS-5 proteinase (161).

To improve the success rate in determining peptide sequences from hit beads, the MPM encoding approach—which utilises a barcode consisting of set of 3D vectors—could be employed (see Chapter I, sub-section 1.5.3 for more details) (171, 172). This is because MPM encoding has several advantages over the MALDI method. Firstly, deactivation of the proteinase after the library incubation step would not be necessary, as identification of the peptide sequence does not require the presence of an intact (uncleaved) peptide on the bead, but rather a “reading” of the bead’s “barcode.” Secondly, it would simplify the library pre-screening step described earlier. The danger of pre-screening with multiple proteinases prior to incubation with TS5-5 is that cleavage of

¹⁴ Personal communication with Professor Morten Meldal.

non-selective substrates generates fragments that can diffuse into other beads and remain there through non-covalent interactions. If these beads are then used for incubation with TS5-5 and are later found to be fluorescent hits, the conclusive identification of the peptide sequence by mass spectrometry would be hampered by the presence of those peptide fragments. Finally, the identification of hit beads relies on barcode recognition between the images of the fluorescent hit substrates to those taken during each step of the synthetic pathway. After a single round of comparisons, recognition between two batches of images was ~80–90%, which increased to ~95–98% after a second round of comparisons, thereby ensuring a high likelihood of identifying a hit (171). This technique has been recently used to screen a small tripeptide library against a mutated version of the Peptidyl-prolyl cis-trans isomerase FKBP-12 enzyme and identified ligands with micromolar binding affinities (217). Unfortunately, MPM encoding was not adopted for identifying hit beads from library screening, as five weeks were spent attempting to get the CCD cameras on the rotating disk to work. This was likely due to the software not registering them correctly. As re-writing the software for a newer version of Windows would require too much time, the MALDI method was used instead.

To conclude, 9 peptide sequences were identified through mass spectrometry and the kinetics of their cleavage by ADAMTS-5 probed in detail alongside the library parent compound **5**. The most rapidly cleaved lead substrates **5** and **39** showed a high level of selectivity for ADAMTS-5 over ADAMTS-4, MMP-2 and MMP-9. Both of these substrates were able to sensitively detect low nanomolar concentrations of ADAMTS-5 in solution. The peptide sequences **5** and **39** represent a new generation of rapidly cleaved, highly selective and sensitive FRET peptide substrates for monitoring ADAMTS-

5 activity *in vitro*, *ex vivo* and potentially *in vivo*. Examples of possible applications in these settings will be discussed in the next chapter.

Chapter 6. Conclusions and future work

6.1. Summary of this work

The aim of this thesis was to develop a selective FRET substrate for monitoring the activity of ADAMTS-5. This could be applied to imaging early proteolytic events in osteoarthritis, based on the hypothesis that the degradation of aggrecan precedes that of type II collagen during the progression of this disease. Towards this goal, the ADAMTS-5 proteinase was purified and characterised (Chapter 3), and a combinatorial library of FRET peptide substrates synthesised and screened against it (Chapter 4). The subsequent hits were identified by mass spectrometry and assessed for their cleavage by ADAMTS-5. The selectivity of each hit for ADAMTS-5 was then compared to ADAMTS-4, MMP-2, and MMP-9. Finally, the sensitivity of the lead substrates for detecting ADAMTS-5 activity was briefly explored (Chapter 5).

The results from this approach were two substrates, which (to the best of our knowledge) represent the most selective *in vitro* FRET probes for monitoring ADAMTS-5 activity to date. The first of these, namely KY(NO₂)TESESRGK(Abz)IYYKKG (**5**), was obtained through computational design and initial screening against ADAMTS-5 on-resin. This peptide was then taken forward as the foundation of the combinatorial library, where renewed screening against ADAMTS-5 delivered the second substrate KY(NO₂)SENE SRGK(Abz)IYYKKG (**39**). When compared to the current ADAMTS-5 literature substrate Abz-TESE~SRGAIY-Dpa-KK-NH₂ (**2**), the novel substrates collectively showed the following improvements: i) better overall cleavage by ADAMTS-5 (~3-fold/~4-fold faster cleavage rate); ii) stronger binding to ADAMTS-5 (~2-fold/~4-fold lower K_m); iii) better overall catalytic efficiency (~1.6-fold/~2-fold higher k_{cat}/K_m); and iv) higher overall selectivity for ADAMTS-5 over ADAMTS-4 (~13-fold/~16-fold), MMP-2 (~8-fold/~10-fold) and MMP-9 (~548-fold/~2561-fold). In addition, both **5** and **39** also

exhibited greater detection sensitivity than the current substrate, as they could be used in enzymatic assays at an ~4-fold lower final concentration (*cf.* 5 μM to 20 μM) and detect down to 0.25 nM proteinase concentrations.

6.2. The importance of selectivity in the design of FRET substrates

6.2.1. The difference between selectivity and specificity

While developing FRET substrates for target proteases, the terms “selectivity” and “specificity” are frequently used interchangeably, yet it is important to draw a distinction between the two. Within this thesis, the “selectivity” of a substrate for a target protease has been defined as the preferential cleavage of the substrate by the target protease over other non-target proteases that cleave it. The level of selectivity of the substrate for the target protease is thus determined by a comparison between the rate of substrate cleavage by the target protease compared to other proteases. The substrate is thus expressed as being “x-fold more selective” for the target protease (where “x” is a non-negative integer). For example, using the same concentration of proteinase, Lim and Meinjohanns *et al* determined the rate of cleavage of the FRET substrate Y(NO₂)GPLG~LEEAK(Abz)G as 1.1 nanomoles s⁻¹ by MMP-12 but only 0.089 nanomoles s⁻¹ by MMP-13 (138). This substrate is therefore ~12-fold more selective for MMP-12 than MMP-13 (138).

On the other hand, “specificity” could be defined as cleavage of a substrate exclusively by a single protease. In other words, the substrate shows an “absolute level” of selectivity for the target protease. Given that there are currently seven main families of known proteases classified by catalytic type (1), with at least 600 individual members in humans (1), achieving specificity with a synthetic FRET substrate for one protease is

practically impossible. This is because it would not only necessitate extensive testing for cleavage against every known protease but would also assume that no further proteases would be discovered in the future. A more practical approach to improving specificity would therefore be to screen against key off-target proteases that might interfere with signal detection, even if this would limit the overall “breadth” of substrate coverage. Should specificity be preferred as a term to selectivity, a more useful working definition of it would be one where there was a 100–1000-fold preference for cleavage of one substrate over another (*i.e.* 100–1000-fold more “selective” means “specific”). This is because the difference in concentrations between different proteases *in vivo* may be in the 10-fold range but are less likely to be in the 100–1000-fold range.

Regardless, given the difficulty in obtaining such high levels of selectivity, it is more practical to utilise the term “selectivity” to describe the preference of substrates for their target proteases rather than “specificity” in the context of this thesis. The term “selectivity” has been used throughout this thesis and will continue be used in the remainder of this chapter.

6.2.2. The importance of selectivity for *in vivo* FRET peptide substrates

It is vital to consider the selectivity of FRET substrates for their target proteases when designing them for *in vivo* applications. The ability to monitor the activity of a target protease selectively would provide information on its mechanism of action and its involvement at different stages of a disease. Prior to administration of a FRET substrate into a living system (*e.g.* through intravenous injection), it is important to assess the number and type of non-target proteases that may be present and could conceivably cleave the substrate. The selectivity of the substrate would then be tuned towards the

target protease and away from the non-target proteases. This tuning process must also consider the relative amounts of non-target proteases in the tissues, as well as the period of time that they are present for during the course of the imaging experiment, should such information be available.

6.3. The importance of kinetic parameters in the design of FRET substrates

6.3.1. Modulating the Michaelis constant and the catalytic rate constant *in vitro*

While selectivity of the substrate for the target protease is essential, the kinetic parameters describing the interaction between these entities are also important. For example, the Michaelis constant (K_m) needs to be low for the FRET substrate to bind to the active site of the protease sufficiently strongly to form the protease-substrate complex and thereby trigger cleavage. Once the subsequent complex has formed, the catalytic rate constant (k_{cat}) must be high enough that substrate cleavage occurs rapidly and generates a detectable fluorescent signal, which is amplified by continuous cleavage of each molecule of FRET substrate (52). If both K_m and k_{cat} values are too low, then the substrate will bind too strongly to the protease, resulting in limited cleavage and thereby poor signal detection as the substrate struggles to be released from the active site. If the opposite is the case (*i.e.* K_m and k_{cat} are too high), only a small amount of substrate will bind to the protease due to its poor affinity, also leading to reduced overall cleavage and signal detection—regardless of the value of k_{cat} . Therefore, successful *in vitro* (and *in vivo*) imaging of protease activity using a FRET substrate depends on a high value for the k_{cat}/K_m parameter (*i.e.* low K_m + high k_{cat}). The higher this value, the closer the protease is to achieving “kinetic perfection,” a state characterised by a rate of cleavage limited

only by the speed at which it encounters substrate in solution (206). These kinetic parameters can be modulated through structure-activity relationship studies based on computational modelling combined with chemical synthesis and testing *in vitro*.

6.3.2. Considering pharmacokinetic parameters for *in vivo* FRET peptide substrates

Successful imaging of protease activity *in vivo* is dependent upon key pharmacokinetic parameters such as distribution, metabolism, excretion and toxicity. Once the FRET substrate has been administered into a human or animal either intravenously or intra-articularly (the latter being more difficult in practise), it would typically be distributed into a number of different tissues, including the liver and kidneys (206). Ideally, an imaging probe such as a FRET substrate needs to be metabolically stable in plasma to limit breakdown prior to arrival at the target location (*e.g.* cartilage). Once arrived, it needs to be rapidly absorbed by the ECM and cleaved by the protease of interest (*e.g.* ADAMTS-5), so that activity can be detected prior to clearance of the probe (41). Metabolic stability of the probe could be attained through minor structural changes to the peptide, such as the use of an amide as opposed to a carboxylic acid at the C-terminus. Rapid excretion is also important to avoid unwanted toxicity through extended residence times in tissue such as cartilage. Such toxicity could result from non-specific off-target binding of the FRET substrate, which is thought to be proportional to the overall lipophilicity of the imaging probe (41). Including polar and/or charged residues in the peptide sequence should be able to reduce this effect (41).

6.3.3. Considering the target protease concentration *in vivo*

Another important consideration for using FRET substrates *in vivo* is the overall concentration of the protease in the body and at the target site. Attempts to assess the

concentration of proteases *in vivo* are limited, as they: i) determine expression levels, which do not provide information on the amount of protease that has been translated or activated; ii) cannot account for internalisation and recycling of proteases following extracellular excretion; and iii) cannot accurately measure protease activity, especially if this is at a very low level. In addition, the target protease should be predominantly expressed in the tissue being studied, although in reality this is often not the case as proteases are typically expressed at different levels in multiple tissues throughout the body. This is exemplified by ADAMTS-5, whose protein expression (in addition to that in articular cartilage) is scattered throughout multiple organs in the body. According to the human protein atlas, the expression of ADAMTS-5 is medium-to-high in the thyroid gland, lungs, oesophagus, gastrointestinal tract organs, gallbladder, liver, pancreas, kidneys and male- and female reproductive organs; amongst others (218, 219). These protein expression values should only be used as a guide, as they are estimates based on a knowledge-based annotation made up of data from immunohistochemistry with both monoclonal and polyclonal anti-ADAMTS-5 antibodies, RNA-sequencing and protein/gene characterisation (219). Nonetheless, they provide some insight into where target protease activity may be localised. It is therefore important that the FRET substrate be administered locally if possible. For example, in the case of imaging ADAMTS-5 activity in osteoarthritic joints, injection would be feasible for larger joints such as those of the shoulders, knees and hips. Unfortunately, it would be much less realistic for joints in the neck, fingers and base of the thumb. This is because the former is located in the sensitive area of the spinal cord and the latter two have a small injection area. Generally speaking however, local administration coupled with a high $k_{\text{cat}}/K_{\text{m}}$ value

and good selectivity would ensure the best chance of successfully measuring protease activity at the desired target location.

6.3.4. Summary

The development of a selective FRET substrate must thus satisfy a considerable number of diverse criteria: obtaining sufficient selectivity for the target protease, ensuring non-target protease cleavage is minimised and ensuring a high value of $k_{\text{cat}}/K_{\text{m}}$. For *in vivo* applications, this would also include providing favourable pharmacokinetic parameters along with an understanding of the expression levels and amounts of the protease at the target site. Other than pharmacokinetic parameters, the path taken for the development of an *in vitro* ADAMTS-5 FRET substrate in this thesis has attempted to address these concerns, with one alternative approach for the key step of each Chapter explicated in the discussion sub-section. This next section will discuss the potential applications for the novel FRET substrates discovered in this thesis.

6.4. Potential applications for selective ADAMTS-5 FRET substrates

6.4.1. General *in vitro* applications

The most common uses of FRET substrates *in vitro* involve assessing the activity of proteinases, determining their concentration through active-site titration with an endogenous or small molecule inhibitor and determining inhibition constants for novel inhibitors. These applications have been described for TS4-4 and TS5-5 in Chapter 3. However, FRET substrates utilised in these types of experiments do not need to be selective. Provided substrate cleavage can be measured for the duration of the experiment, any general substrate may be employed. For example, Mca-PLG~L-Dpa-AR-NH₂ is frequently used for activity assays and active-site titrations of MMPs. However,

the selective ADAMTS-5 FRET substrates **5** and **39** identified in Chapter 5 could potentially have more specific *in vitro* and *in vivo* applications in diseases where ADAMTS-5 activity plays a key role.

6.4.2. Application to tissue culture models of osteoarthritis

As has been mentioned several times, a number of *in vitro* experiments that simulate osteoarthritis through the treatment of cartilage explants with IL-1 have shown that degradation of the ECM occurs through initial breakdown of aggrecan followed by the loss of type II collagen (78-80). Aggrecan degradation is mediated by ADAMTS-4 and ADAMTS-5 (120, 121), whereas degeneration of type II collagen is primarily governed by MMP-13 (91, 92). By collecting an aliquot of media during *in vitro* culture of cartilage explants stimulated with IL-1, the presence of active ADAMTS-5 could be selectively detected through cleavage of FRET substrates **5** and **39**. To detect this fluorescence signal however, the collected media would need to fulfil the following requirements: i) contain no serum, as this possesses a high concentration of proteinase inhibitors, which would interfere with FRET substrate cleavage (79); ii) contain a sufficient concentration of ADAMTS-5, which could be attained by leaving the culture experiment for at least 5–7 days, as this is where aggrecanase activity is typically at its highest (78, 80); and iii) contain no phenol red, as this would interfere with signal detection due to its absorbance between 400–450 nm (220), which overlaps with the emission maximum of the Abz fluorophore at 420 nm (157). Separate assays could be carried out with 5-FAM-AE~LQGRPISIAK-TAMRA (64) and PGA-P-18 (140) to selectively detect ADAMTS-4 and MMP-13 activity, respectively. The other distinct advantage that these selective substrates have over more generic ones is that they are cleaved more rapidly by the proteinase of interest (higher $k_{\text{cat}}/K_{\text{m}}$). This means that they possess greater detection

sensitivity for that proteinase, so lower concentrations can be used to detect the same level of activity—a fact that is also useful in situations where very low proteinase activity is present. A case in point are lead substrates **5** and **39**, which can be used at final concentrations of 5 μ M compared to the literature substrate **2**, which is currently used at 20 μ M. These substrates can also detect activity from concentrations as low as 0.25 nM ADAMTS-5.

Overall, applying the above-mentioned selective substrates to an *in vitro* model of osteoarthritis would generate a picture of the relative activities of ADAMTS-5, ADAMTS-4, and MMP-13 at different stages in the progression of this disease, even if this were only a model system.

6.4.3. Characterisation of pharmacological inhibitors

Using a cell culture model of osteoarthritis similar to that mentioned above, small molecule inhibitors of ADAMTS-5 could be rapidly screened for their efficacy. Aliquots of media would be taken from two sample groups: one with inhibitor and one without. The activity between the two groups would then be compared using **5** or **39**. In the presence of a selective ADAMTS-5 inhibitor, cleavage of either substrate would result in a decrease in fluorescence intensity compared to the control group. More generally, the inhibition constants of these inhibitors could also be determined more accurately, as lower concentrations of ADAMTS-5 could be used.

6.4.4. High-throughput screening for drugs

Large chemogenomic libraries of small molecules could be screened against the ADAMTS-5 proteinase to discover novel inhibitors (221). The FRET substrates **5** and **39** would be used to assess the inhibition constant of the best hits, which could be applied

to diseases characterised by high ADAMTS-5 expression, including colorectal cancer (222) and glioblastoma (223)—although this mRNA expression does not necessarily indicate a high level of proteinase activity (224).

6.4.5. ADAMTS-5 activity as a biomarker of human osteoarthritis

ADAMTS-5 activity could also be employed as a potential clinical biomarker for human osteoarthritis. A considerable number of synovial fluid biomarkers have been identified for the diagnosis of knee osteoarthritis (122). These have included MMP-1 and MMP-13, whose protein levels were found to be significantly elevated in the synovial fluid of patients with disease compared to those without disease (122, 225, 226). Therefore, the activity of ADAMTS-5 would be measured in the synovial fluid of patients with and without osteoarthritis *via* cleavage of **5** or **39**. In addition, a control experiment consisting of different concentrations of ADAMTS-5 and a fixed concentration of **5** or **39** would enable the synovial fluid concentration of ADAMTS-5 to be determined by virtue of a standard curve. The validity of this approach has been demonstrated by measurement of the ADAMTS-4 concentration in the synovial fluid of patients with a range of joint pathologies using the Sensolyte 520 ADAMTS-4 FRET substrate assay kit (227).

6.4.6. Monitoring ADAMTS-5 activity in animal models of osteoarthritis

The FRET probes **5** and **39** could also be used to detect and localise ADAMTS-5 activity in a pre-clinical animal model of osteoarthritis. Unfortunately, this was not carried out in this thesis due to time constraints. Nevertheless, it would involve intra-articular injection of the probe into the animal *in vivo*, culling of the animal, dissection to isolate the articular cartilage and subsequent freezing. The cartilage samples would then be cryosectioned and ADAMTS-5 activity localised through a fluorescent microscope.

This could potentially reveal varying activities in the different regions and zones of articular cartilage mentioned in the introduction.

Unfortunately, intra-articular injection of imaging probes *in vivo* results in rapid clearance from cartilage unless pharmacokinetic parameters are adjusted (228). This likely explains why many *in vivo* studies with MMP FRET substrates have utilised intravenous (132, 135, 139, 140) or intraorbital (134) injections to administer probes. However, one of the ways in which the cartilage residence time of FRET substrates could be increased would be to attach them to type II collagen through the targeting peptide WYRGRL (229). This would be connected to the FRET substrate at the C-terminus *via* a short flexible PEG linker such as 8-amino-3,6-dioxaoctanoic acid (AEEA) to give **5/39**-AEEA-WYRGRL. Once the FRET probe was injected into the mouse, it would effectively become “anchored” in the ECM of cartilage *via* the WYRGRL targeting moiety. As a study has shown that aggrecan interacts with collagen fibrils *via* the KS region (230), it may be protecting type II collagen from degradation (80, 230). Therefore, attachment of the probe to the more stable type II collagen macromolecule would increase its residence time and potentially improve detection of ADAMTS-5 activity.

6.4.7. Protease-activated prodrugs

Another potential application for **5** and **39** would be in the field of protease-activated prodrugs (PAPs) (231). A PAP essentially consists of three key elements: a prodrug connected to a protease-cleavable linker attached to a targeting entity (231). Once the targeting entity has delivered the compound to the required location, cleavage of the linker by a specific protease releases the drug into the tissue compartment (231). This approach has led to the clinical approval of two PAPs with antibody targeting entities: brentuximab vedotin, which was used to treat late-stage Hodgkin’s lymphoma

in combination with chemotherapy (232); and polatuzumab vedotin, which was used to treat B-cell lymphoma in combination with immunochemotherapy (233).

In the case of the ADAMTS-5 FRET substrates described in this thesis, the fluorophore and quencher of **5** and **39** would be replaced with a lysine and tyrosine, respectively, to give KYTESERGGKIYYKKG (**5**) and KYSENERGGKIYYKKG (**39**). To enable the delivery of drugs to specific tissues, these sequences would form the protease-cleavable linker in the PAP in combination with a targeting entity. For example, osteoarthritic cartilage could be targeted with a PAP consisting of the aforementioned type II collagen binding peptide WYRGRL (229) or any anti-neo-epitope antibody for aggrecan degradation fragments described in Chapter 1 (*e.g.* anti-ARGS). As for the drug, it could be the selective ADAMTS-5 inhibitor GLPG1972 (234). Similarly, other tissues exhibiting ADAMTS-5 expression—such as the aorta (235)—could be targeted with peptides that bind to the constituent endothelial cells (*e.g.* CLIRRTSIC or CPRRSHPIC) (236).

6.5. Future outlook

6.5.1. Developing selective substrates for other ADAMTS proteinases

In this thesis, the approach taken towards the development of a selective ADAMTS-5 FRET peptide substrate could also be applied to anabolic proteinases such as ADAMTS-2 and ADAMTS-3. These pro-collagen N-proteinases cleave the N-terminal pro-peptide of type II collagen to generate collagen fibrils, thereby playing an important role in the repair and maintenance of cartilage extracellular matrix (18). As there are currently no published crystal structures of ADAMTS-2 or ADAMTS-3, a homology model could be generated from their sequences using the existing crystal structures of ADAMTS-1, -4, -5 and -13 as a template. This model would be used for initial computational docking of

ADAMTS-2 or ADAMTS-3 substrates followed by the same approach to obtain the final selective probe. The ability to simultaneously monitor the activity of both anabolic (*e.g.* ADAMTS-2) and catabolic (*e.g.* ADAMTS-5) proteinases would thus help uncover the dynamic processes that underpin the maintenance of cartilage homeostasis. For example, co-injection of selective ADAMTS-2 and ADAMTS-5 FRET substrates *in vivo* could indicate the direction in which the homeostatic balance was moving; whether towards disease (*i.e.* increased catabolism) or away from disease (*i.e.* increased anabolism). A selective inhibitor could then be applied to reverse any metabolic imbalance and shift this towards a healthy system.

6.5.2. Concluding remarks

This thesis has covered the evolution of two FRET peptide substrates for selectively monitoring ADAMTS-5 proteinase activity. The development of these fluorescent imaging agents was inherently multidisciplinary; combining the fields of molecular biology, biochemistry, computation, synthetic chemistry, mass spectrometry and molecular imaging. Whilst aspects of the approach require further optimisation, the impact of this work lies in the ability to utilise it to develop more selective substrates for the comprehensive and targeted interrogation of proteinase activity in specific biological settings.

Chapter 7. References

1. N. D. Rawlings, "Protease Families, Evolution and Mechanism of Action" in *Proteases: Structure and Function*, K. Brix, W. Stöcker, Eds. (Springer Vienna, Vienna, 2013), pp. 1-36.
2. J. S. Bond, Proteases: History, discovery, and roles in health and disease. *J. Biol. Chem.* **294**, 1643-1651 (2019).
3. T. L. Vincent, L. Troeberg, "Pathogenesis of osteoarthritis" in *Oxford Textbook of Rheumatology*. (Oxford University Press, Oxford, 2018), chap. 138, pp. 1-22.
4. H. J. Vernon, Inborn Errors of Metabolism: Advances in Diagnosis and Therapy. *JAMA Pediatr.* **169**, 778-782 (2015).
5. D. Chen *et al.*, Osteoarthritis: toward a comprehensive understanding of pathological mechanism. *Bone Res.* **5**, 16044 (2017).
6. N. D. Rawlings, A. J. Barrett, Evolutionary families of peptidases. *Biochem. J.* **290**, 205-218 (1993).
7. A. J. Barrett, J. K. McDonald, Nomenclature: protease, proteinase and peptidase. *Biochem. J.* **237**, 935 (1986).
8. I. Schechter, A. Berger, On Size of Active Site in Proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157-162 (1967).
9. J. Reiser, B. Adair, T. Reinheckel, Specialized roles for cysteine cathepsins in health and disease. *J. Clin. Invest.* **120**, 3421-3431 (2010).
10. H. Nagase, J. F. Woessner, Jr., Matrix metalloproteinases. *J. Biol. Chem.* **274**, 21491-21494 (1999).
11. H. Nagase, R. Visse, G. Murphy, Structure and function of matrix metalloproteinases and TIMPs. *Cardiovasc. Res.* **69**, 562-573 (2006).
12. J. C. Rodríguez-Manzanique, R. Fernández-Rodríguez, F. J. Rodríguez-Baena, M. L. Iruela-Arispe, ADAMTS proteases in vascular biology. *Matrix Biol.* **44-46**, 38-45 (2015).
13. R. Kelwick, I. Desanlis, G. N. Wheeler, D. R. Edwards, The ADAMTS (A Disintegrin and Metalloproteinase with Thrombospondin motifs) family. *Genome Biol.* **16**, 113 (2015).
14. S. Takeda, H. Takeya, S. Iwanaga, Snake venom metalloproteinases: structure, function and relevance to the mammalian ADAM/ADAMTS family proteins. *Biochim. Biophys. Acta* **1824**, 164-176 (2012).
15. S. Takeda, ADAM and ADAMTS Family Proteins and Snake Venom Metalloproteinases: A Structural Overview. *Toxins* **8**, 1-35 (2016).
16. M. Kashiwagi *et al.*, Altered proteolytic activities of ADAMTS-4 expressed by C-terminal processing. *J. Biol. Chem.* **279**, 10109-10119 (2004).
17. C. Gendron *et al.*, Proteolytic activities of human ADAMTS-5: comparative studies with ADAMTS-4. *J. Biol. Chem.* **282**, 18294-18306 (2007).
18. C. Y. Yang, A. Chanalaris, L. Troeberg, ADAMTS and ADAM metalloproteinases in osteoarthritis - looking beyond the 'usual suspects'. *Osteoarthr. Cartil.* **25**, 1000-1009 (2017).
19. S. Santamaria, R. de Groot, ADAMTS proteases in cardiovascular physiology and disease. *Open Biol. J.* **10**, 200333 (2020).
20. S. Z. Karoulias, N. Taye, S. Stanley, D. Hubmacher, The ADAMTS/Fibrillin Connection: Insights into the Biological Functions of ADAMTS10 and ADAMTS17 and Their Respective Sister Proteases. *Biomolecules* **10** (2020).
21. K. Brew, H. Nagase, The tissue inhibitors of metalloproteinases (TIMPs): an ancient family with structural and functional diversity. *Biochim. Biophys. Acta* **1803**, 55-71 (2010).
22. G. Murphy, Tissue inhibitors of metalloproteinases. *Genome Biol.* **12**, 233 (2011).
23. N. H. Lim *et al.*, Reactive-site mutants of N-TIMP-3 that selectively inhibit ADAMTS-4 and ADAMTS-5: biological and structural implications. *Biochem. J.* **431**, 113-122 (2010).
24. W. L. Mock, H. Cheng, Principles of hydroxamate inhibition of metalloproteases: carboxypeptidase A. *Biochemistry* **39**, 13945-13952 (2000).

25. U. K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).
26. M. D. Tortorella, R. Q. Liu, T. Burn, R. C. Newton, E. Arner, Characterization of human aggrecanase 2 (ADAM-TS5): substrate specificity studies and comparison with aggrecanase 1 (ADAM-TS4). *Matrix Biol.* **21**, 499-511 (2002).
27. A. J. S. Fox, A. Bedi, S. A. Rodeo, The basic science of articular cartilage: structure, composition, and function. *Sports Health* **1**, 461-468 (2009).
28. M. Bottini *et al.*, Nanodrugs to target articular cartilage: An emerging platform for osteoarthritis therapy. *Nanomedicine* **12**, 255-268 (2016).
29. R. W. Farndale, C. A. Sayers, A. J. Barrett, A direct spectrophotometric microassay for sulfated glycosaminoglycans in cartilage cultures. *Connect. Tissue Res.* **9**, 247-248 (1982).
30. R. W. Farndale, D. J. Buttle, A. J. Barrett, Improved quantitation and discrimination of sulphated glycosaminoglycans by use of dimethylmethylene blue. *Biochim. Biophys. Acta* **883**, 173-177 (1986).
31. R. E. Neuman, M. A. Logan, The determination of hydroxyproline. *J. Biol. Chem.* **184**, 299-306 (1950).
32. D. D. Cissell, J. M. Link, J. C. Hu, K. A. Athanasiou, A Modified Hydroxyproline Assay Based on Hydrochloric Acid in Ehrlich's Solution Accurately Measures Tissue Collagen Content. *Tissue Eng. Part C Methods* **23**, 243-250 (2017).
33. V. J. Coulson-Thomas, F. G. Tarsis, Dimethylmethylene Blue Assay (DMMB). *Bio-protocol* **4**, e1236 (2014).
34. Sigma-Aldrich, Enzymatic Method for Determining Hydroxyproline (Hydroxyproline Assay), <https://www.sigmaaldrich.com/technical-documents/protocols/biology/hydroxyproline-assay-kit-mak008.html> (accessed March 8th, 2021).
35. F. L. Moseley, K. A. Bicknell, M. S. Marber, G. Brooks, The use of proteomics to identify novel therapeutic targets for the treatment of disease. *J. Pharm. Pharmacol.* **59**, 609-628 (2007).
36. T. Jones, The imaging science of positron emission tomography. *Eur. J. Nucl. Med.* **23**, 807-813 (1996).
37. R. Weissleder, Molecular imaging: Exploring the next frontier. *Radiology* **212**, 609-614 (1999).
38. M. G. Pomper, Molecular imaging: an overview. *Acad. Radiol.* **8**, 1141-1153 (2001).
39. J. M. Hoffman, S. S. Gambhir, Molecular imaging: the vision and opportunity for radiology in the future. *Radiology* **244**, 39-47 (2007).
40. R. Weissleder, U. Mahmood, Molecular imaging. *Radiology* **219**, 316-333 (2001).
41. M. L. James, S. S. Gambhir, A molecular imaging primer: modalities, imaging agents, and applications. *Physiol. Rev.* **92**, 897-965 (2012).
42. PerkinElmer, Quantum GX2 micro-CT imaging system, <https://www.perkinelmer.com/product/quantum-gx2-instrument-cls149276> (accessed February 13th 2021).
43. H. Lusic, M. W. Grinstaff, X-ray-computed tomography contrast agents. *Chem. Rev.* **113**, 1641-1666 (2013).
44. B. Yoo, M. D. Pagel, A PARACEST MRI contrast agent to detect enzyme activity. *J. Am. Chem. Soc.* **128**, 14032-14033 (2006).
45. B. Yoo, M. S. Raam, R. M. Rosenblum, M. D. Pagel, Enzyme-responsive PARACEST MRI contrast agents: a new biomedical imaging approach for studies of the proteasome. *Contrast Media Mol. Imaging* **2**, 189-198 (2007).
46. M. Suchy *et al.*, A paramagnetic chemical exchange-based MRI probe metabolized by cathepsin D: design, synthesis and cellular uptake studies. *Org. Biomol. Chem.* **8**, 2560-2566 (2010).
47. T. Chauvin *et al.*, Detection of enzymatic activity by PARACEST MRI: a general approach to target a large variety of enzymes. *Angew. Chem. Int. Ed.* **47**, 4370-4372 (2008).

48. J. A. Snir *et al.*, An Aspartyl Cathepsin Targeted PET Agent: Application in an Alzheimer's Disease Mouse Model. *J. Alzheimer's Dis.* **61**, 1241-1252 (2018).
49. V. Butsch *et al.*, Design, (Radio)Synthesis, and in Vitro and in Vivo Evaluation of Highly Selective and Potent Matrix Metalloproteinase 12 (MMP-12) Inhibitors as Radiotracers for Positron Emission Tomography. *J. Med. Chem.* **61**, 4115-4134 (2018).
50. J. Toczek *et al.*, Novel Matrix Metalloproteinase 12 Selective Radiotracers for Vascular Molecular Imaging. *J. Med. Chem.* **62**, 9743-9752 (2019).
51. T. Förster, Zwischenmolekulare Energiewanderung Und Fluoreszenz. *Ann. Phys.* **2**, 55-75 (1948).
52. L. L. Gompels, N. H. Lim, T. Vincent, E. M. Paleolog, In vivo optical imaging in arthritis—an enlightening future? *Rheumatology (Oxford)* **49**, 1436-1446 (2010).
53. N. H. Lim, C. Wen, T. L. Vincent, Molecular and structural imaging in surgically induced murine osteoarthritis. *Osteoarthr. Cartil.* **28**, 874-884 (2020).
54. Thermo Fisher Scientific, "Fluorophores and Their Amine-Reactive Derivatives" in *Molecular Probes Handbook*. (2010), chap. 1, pp. 1-97.
55. R. Weissleder, V. Ntziachristos, Shedding light onto live molecular targets. *Nat. Med.* **9**, 123-128 (2003).
56. PerkinElmer, Fluorescent Agents, <https://www.perkinelmer.com/category/fluorescent-agents> (accessed March 11th 2021).
57. R. Weissleder, C. H. Tung, U. Mahmood, A. Bogdanov, Jr., In vivo imaging of tumors with protease-activated near-infrared fluorescent probes. *Nat. Biotechnol.* **17**, 375-378 (1999).
58. C. H. Tung, S. Bredow, U. Mahmood, R. Weissleder, Preparation of a cathepsin D sensitive near-infrared fluorescence probe for imaging. *Bioconjugate Chem.* **10**, 892-896 (1999).
59. C. Bremer, C. H. Tung, R. Weissleder, In vivo molecular target assessment of matrix metalloproteinase inhibition. *Nat. Med.* **7**, 743-748 (2001).
60. J. Chen *et al.*, Near-infrared fluorescent imaging of matrix metalloproteinase activity after myocardial infarction. *Circulation* **111**, 1800-1805 (2005).
61. F. A. Jaffer *et al.*, Optical visualization of cathepsin K activity in atherosclerosis with a novel, protease-activatable fluorescence sensor. *Circulation* **115**, 2292-2298 (2007).
62. L. Troeberg *et al.*, The C-terminal domains of ADAMTS-4 and ADAMTS-5 promote association with N-TIMP-3. *Matrix Biol.* **28**, 463-469 (2009).
63. G. J. Wayne *et al.*, TIMP-3 inhibition of ADAMTS-4 (Aggrecanase-1) is modulated by interactions between aggrecan and the C-terminal domain of ADAMTS-4. *J. Biol. Chem.* **282**, 20991-20998 (2007).
64. R. Hills *et al.*, Identification of an ADAMTS-4 cleavage motif using phage display leads to the development of fluorogenic peptide substrates and reveals matrilin-3 as a novel substrate. *J. Biol. Chem.* **282**, 11101-11109 (2007).
65. M. Cudic, G. D. Burstein, G. B. Fields, J. Lauer-Fields, Analysis of flavonoid-based pharmacophores that inhibit aggrecanases (ADAMTS-4 and ADAMTS-5) and matrix metalloproteinases through the use of topologically constrained peptide substrates. *Chem. Biol. Drug Des.* **74**, 473-482 (2009).
66. GBD, Disease and Injury Incidence and Prevalence Collaborators, Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789-1858 (2018).
67. Versus Arthritis, Osteoarthritis in general practise (2013).
68. Versus Arthritis, State of musculoskeletal health (2019).
69. R. F. Loeser, S. R. Goldring, C. R. Scanzello, M. B. Goldring, Osteoarthritis: a disease of the joint as an organ. *Arthritis Rheum.* **64**, 1697-1707 (2012).
70. R. F. Loeser, J. A. Collins, B. O. Diekman, Ageing and the pathogenesis of osteoarthritis. *Nat. Rev. Rheumatol.* **12**, 412-420 (2016).

71. A. Abhishek, M. Doherty, "Clinical features of osteoarthritis" in Oxford Textbook of Rheumatology. (Oxford University Press, Oxford, 2018), chap. 139, pp. 1-33.
72. M. Fusco, S. D. Skaper, S. Coaccioli, G. Varrassi, A. Paladini, Degenerative joint diseases and neuroinflammation. *Pain Pract.* **17**, 522-532 (2017).
73. D. J. Hunter, S. Bierma-Zeinstra, Osteoarthritis. *Lancet* **393**, 1745-1759 (2019).
74. D. T. Felson, T. Neogi, Osteoarthritis: is it a disease of cartilage or of bone? *Arthritis Rheum.* **50**, 341-344 (2004).
75. K. D. Brandt, P. Dieppe, E. L. Radin, Commentary: is it useful to subset "primary" osteoarthritis? A critique based on evidence regarding the etiopathogenesis of osteoarthritis. *Semin. Arthritis Rheum.* **39**, 81-95 (2009).
76. C. W. Archer, P. Francis-West, The chondrocyte. *Int. J. Biochem. Cell Biol.* **35**, 401-404 (2003).
77. K. Gelse, E. Poschl, T. Aigner, Collagens - structure, function, and biosynthesis. *Adv. Drug Del. Rev.* **55**, 1531-1546 (2003).
78. A. J. Ellis, V. A. Curry, E. K. Powell, T. E. Cawston, The prevention of collagen breakdown in bovine nasal cartilage by TIMP, TIMP-2 and a low molecular weight synthetic inhibitor. *Biochem. Biophys. Res. Commun.* **201**, 94-101 (1994).
79. L. D. Kozaci, D. J. Buttle, A. P. Hollander, Degradation of type II collagen, but not proteoglycan, correlates with matrix metalloproteinase activity in cartilage explant cultures. *Arthritis Rheum.* **40**, 164-174 (1997).
80. M. A. Pratta *et al.*, Aggrecan protects cartilage collagen from proteolytic cleavage. *J. Biol. Chem.* **278**, 45539-45545 (2003).
81. A. P. Hollander *et al.*, Damage to type II collagen in aging and osteoarthritis starts at the articular surface, originates around chondrocytes, and extends into the cartilage with progressive degeneration. *J. Clin. Invest.* **96**, 2859-2869 (1995).
82. S. S. Glasson, M. G. Chambers, W. B. Van Den Berg, C. B. Little, The OARSI histopathology initiative - recommendations for histological assessments of osteoarthritis in the mouse. *Osteoarthr. Cartil.* **18** S17-23 (2010).
83. P. G. Mitchell, L. Lopresti-Morrow, S. A. Yocum, F. J. Sweeney, L. A. Reiter, Inhibition of interleukin-1-stimulated collagen degradation in cartilage explants. *Ann. N. Y. Acad. Sci.* **732**, 395-397 (1994).
84. The Human Tissue Authority, Human Tissue Act 30. 1-63 (2004).
85. C. G. Knight, F. Willenbrock, G. Murphy, A novel coumarin-labelled peptide for sensitive continuous assays of the matrix metalloproteinases. *FEBS Lett.* **296**, 263-266 (1992).
86. G. R. Dodge, A. R. Poole, Immunohistochemical detection and immunochemical analysis of type II collagen degradation in human normal, rheumatoid, and osteoarthritic articular cartilages and in explants of bovine articular cartilage cultured with interleukin 1. *J. Clin. Invest.* **83**, 647-661 (1989).
87. A. P. Hollander *et al.*, Increased damage to type II collagen in osteoarthritic articular cartilage detected by a new immunoassay. *J. Clin. Invest.* **93**, 1722-1732 (1994).
88. G. C. Wolfe *et al.*, Differential in vivo expression of collagenase messenger RNA in synovium and cartilage. Quantitative comparison with stromelysin messenger RNA levels in human rheumatoid arthritis and osteoarthritis patients and in two animal models of acute inflammatory arthritis. *Arthritis Rheum.* **36**, 1540-1547 (1993).
89. L. S. Lohmander, L. A. Hoerrner, M. W. Lark, Metalloproteinases, tissue inhibitor, and proteoglycan fragments in knee synovial fluid in human osteoarthritis. *Arthritis Rheum.* **36**, 181-189 (1993).
90. L. A. Walakovits, V. L. Moore, N. Bhardwaj, G. S. Gallick, M. W. Lark, Detection of stromelysin and collagenase in synovial fluid from patients with rheumatoid arthritis and posttraumatic knee injury. *Arthritis Rheum.* **35**, 35-42 (1992).

91. P. G. Mitchell *et al.*, Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage. *J. Clin. Invest.* **97**, 761-768 (1996).
92. R. C. Billingham *et al.*, Enhanced cleavage of type II collagen by collagenases in osteoarthritic articular cartilage. *J. Clin. Invest.* **99**, 1534-1545 (1997).
93. L. Dahlberg *et al.*, Selective enhancement of collagenase-mediated cleavage of resident type II collagen in cultured osteoarthritic cartilage and arrest with a synthetic inhibitor that spares collagenase 1 (matrix metalloproteinase 1). *Arthritis Rheum.* **43**, 673-682 (2000).
94. M. Balbin *et al.*, Identification and enzymatic characterization of two diverging murine counterparts of human interstitial collagenase (MMP-1) expressed at sites of embryo implantation. *J. Biol. Chem.* **276**, 10253-10262 (2001).
95. C. B. Little *et al.*, Matrix Metalloproteinase 13-deficient mice are resistant to osteoarthritic cartilage erosion but not chondrocyte hypertrophy or osteophyte development. *Arthritis Rheum.* **60**, 3723-3733 (2009).
96. C. T. Appleton, V. Pitelka, J. Henry, F. Beier, Global analyses of gene expression in early experimental osteoarthritis. *Arthritis Rheum.* **56**, 1854-1868 (2007).
97. J. D. Sandy, R. E. Boynton, C. R. Flannery, Analysis of the catabolism of aggrecan in cartilage explants by quantitation of peptides from the three globular domains. *J. Biol. Chem.* **266**, 8198-8205 (1991).
98. J. D. Sandy, P. J. Neame, R. E. Boynton, C. R. Flannery, Catabolism of aggrecan in cartilage explants. Identification of a major cleavage site within the interglobular domain. *J. Biol. Chem.* **266**, 8683-8685 (1991).
99. M. Z. Ilic, C. J. Handley, H. C. Robinson, M. T. Mok, Mechanism of catabolism of aggrecan by articular cartilage. *Arch. Biochem. Biophys.* **294**, 115-122 (1992).
100. P. Loulakis, A. Shrikhande, G. Davis, C. A. Maniglia, N-terminal sequence of proteoglycan fragments isolated from medium of interleukin-1-treated articular-cartilage cultures. Putative site(s) of enzymic cleavage. *Biochem. J.* **284**, 589-593 (1992).
101. M. W. Lark *et al.*, Cell-mediated catabolism of aggrecan. Evidence that cleavage at the "aggrecanase" site (Glu373-Ala374) is a primary event in proteolysis of the interglobular domain. *J. Biol. Chem.* **270**, 2550-2556 (1995).
102. E. C. Arner, C. E. Hughes, C. P. Decicco, B. Caterson, M. D. Tortorella, Cytokine-induced cartilage proteoglycan degradation is mediated by aggrecanase. *Osteoarthr. Cartil.* **6**, 214-228 (1998).
103. K. Kuno *et al.*, Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene. *J. Biol. Chem.* **272**, 556-562 (1997).
104. M. D. Tortorella *et al.*, Purification and cloning of aggrecanase-1: a member of the ADAMTS family of proteins. *Science* **284**, 1664-1666 (1999).
105. I. Abbaszade *et al.*, Cloning and characterization of ADAMTS11, an aggrecanase from the ADAMTS family. *J. Biol. Chem.* **274**, 23443-23450 (1999).
106. M. D. Tortorella *et al.*, Sites of aggrecan cleavage by recombinant human aggrecanase-1 (ADAMTS-4). *J. Biol. Chem.* **275**, 18566-18573 (2000).
107. J. D. Sandy, C. R. Flannery, P. J. Neame, L. S. Lohmander, The structure of aggrecan fragments in human synovial fluid. Evidence for the involvement in osteoarthritis of a novel proteinase which cleaves the Glu 373-Ala 374 bond of the interglobular domain. *J. Clin. Invest.* **89**, 1512-1516 (1992).
108. L. S. Lohmander, P. J. Neame, J. D. Sandy, The structure of aggrecan fragments in human synovial fluid. Evidence that aggrecanase mediates cartilage degradation in inflammatory joint disease, joint injury, and osteoarthritis. *Arthritis Rheum.* **36**, 1214-1222 (1993).

109. M. W. Lark *et al.*, Aggrecan degradation in human cartilage. Evidence for both matrix metalloproteinase and aggrecanase activity in normal, osteoarthritic, and rheumatoid joints. *J. Clin. Invest.* **100**, 93-106 (1997).
110. C. E. Hughes, B. Caterson, A. J. Fosang, P. J. Roughley, J. S. Mort, Monoclonal antibodies that specifically recognize neoepitope sequences generated by 'aggrecanase' and matrix metalloproteinase cleavage of aggrecan: application to catabolism in situ and in vitro. *Biochem. J.* **305**, 799-804 (1995).
111. J. D. Sandy, C. Verscharen, Analysis of aggrecan in human knee cartilage and synovial fluid indicates that aggrecanase (ADAMTS) activity is responsible for the catabolic turnover and loss of whole aggrecan whereas other protease activity is required for C-terminal processing in vivo. *Biochem. J.* **358**, 615-626 (2001).
112. A. Struglics *et al.*, Human osteoarthritis synovial fluid and joint cartilage contain both aggrecanase- and matrix metalloproteinase-generated aggrecan fragments. *Osteoarthr. Cartil.* **14**, 101-113 (2006).
113. M. D. Tortorella, A. M. Malfait, C. Deccico, E. Arner, The role of ADAM-TS4 (aggrecanase-1) and ADAM-TS5 (aggrecanase-2) in a model of cartilage degradation. *Osteoarthr. Cartil.* **9**, 539-552 (2001).
114. A. Burleigh *et al.*, Joint immobilization prevents murine osteoarthritis and reveals the highly mechanosensitive nature of protease expression in vivo. *Arthritis Rheum.* **64**, 2278-2288 (2012).
115. S. S. Glasson, T. J. Blanchet, E. A. Morris, The surgical destabilization of the medial meniscus (DMM) model of osteoarthritis in the 129/SvEv mouse. *Osteoarthr. Cartil.* **15**, 1061-1069 (2007).
116. S. S. Glasson *et al.*, Characterization of and osteoarthritis susceptibility in ADAMTS-4-knockout mice. *Arthritis Rheum.* **50**, 2547-2558 (2004).
117. S. S. Glasson *et al.*, Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis. *Nature* **434**, 644-648 (2005).
118. C. J. East, H. Stanton, S. B. Golub, F. M. Rogerson, A. J. Fosang, ADAMTS-5 deficiency does not block aggrecanolysis at preferred cleavage sites in the chondroitin sulfate-rich region of aggrecan. *J. Biol. Chem.* **282**, 8632-8640 (2007).
119. J. Larkin *et al.*, Translational development of an ADAMTS-5 antibody for osteoarthritis disease modification. *Osteoarthr. Cartil.* **23**, 1254-1266 (2015).
120. R. H. Song *et al.*, Aggrecan degradation in human articular cartilage explants is mediated by both ADAMTS-4 and ADAMTS-5. *Arthritis Rheum.* **56**, 575-585 (2007).
121. H. M. Ismail *et al.*, Interleukin-1 Acts via the JNK-2 Signaling Pathway to Induce Aggrecan Degradation by Human Chondrocytes. *Arthritis Rheumatol.* **67**, 1826-1836 (2015).
122. A. Boffa *et al.*, Synovial Fluid Biomarkers in Knee Osteoarthritis: A Systematic Review and Quantitative Evaluation Using BIPEDs Criteria. *Cartilage In Press*, 1-22 (2020).
123. V. B. Kraus, M. A. Karsdal, Osteoarthritis: Current Molecular Biomarkers and the Way Forward. *Calcif. Tissue Int. In Press* (2020).
124. NICE, Osteoarthritis: care and management in adults, clinical guideline CG177 (2014).
125. F. W. Roemer, F. Eckstein, D. Hayashi, A. Guermazi, The role of imaging in osteoarthritis. *Best Pract. Res. Clin. Rheumatol.* **28**, 31-60 (2014).
126. J. H. Kellgren, J. S. Lawrence, Radiological assessment of osteo-arthritis. *Ann. Rheum. Dis.* **16**, 494-502 (1957).
127. A. J. Carr *et al.*, Knee replacement. *Lancet* **379**, 1331-1340 (2012).
128. J. S. Mort, M. C. Magny, E. R. Lee, Cathepsin B: an alternative protease for the generation of an aggrecan 'metalloproteinase' cleavage neoepitope. *Biochem. J.* **335**, 491-494 (1998).
129. W. Kafienah, D. Bromme, D. J. Buttle, L. J. Croucher, A. P. Hollander, Human cathepsin K cleaves native type I and II collagens at the N-terminal end of the triple helix. *Biochem. J.* **331**, 727-732 (1998).

130. E. Kozawa *et al.*, Osteoarthritic change is delayed in a Ctsk-knockout mouse model of osteoarthritis. *Arthritis Rheum.* **64**, 454-464 (2012).
131. A. Baici, D. Horler, A. Lang, C. Merlin, R. Kissling, Cathepsin B in osteoarthritis: zonal variation of enzyme activity in human femoral head cartilage. *Ann. Rheum. Dis.* **54**, 281-288 (1995).
132. E. A. Vermeij *et al.*, In vivo molecular imaging of cathepsin and matrix metalloproteinase activity discriminates between arthritic and osteoarthritic processes in mice. *Mol. Imaging* **13**, 1-10 (2014).
133. A. A. Leahy *et al.*, Analysis of the trajectory of osteoarthritis development in a mouse model by serial near-infrared fluorescence imaging of matrix metalloproteinase activities. *Arthritis Rheumatol.* **67**, 442-453 (2015).
134. T. Fukui, E. Tenborg, J. H. Yik, D. R. Haudenschild, In-vitro and in-vivo imaging of MMP activity in cartilage and joint injury. *Biochem. Biophys. Res. Commun.* **460**, 741-746 (2015).
135. P. B. Satkunananthan *et al.*, In vivo fluorescence reflectance imaging of protease activity in a mouse model of post-traumatic osteoarthritis. *Osteoarthr. Cartil.* **22**, 1461-1469 (2014).
136. S. Lee *et al.*, Dark quenched matrix metalloproteinase fluorogenic probe for imaging osteoarthritis development in vivo. *Bioconjugate Chem.* **19**, 1743-1747 (2008).
137. J. H. Ryu *et al.*, Optimization of matrix metalloproteinase fluorogenic probes for osteoarthritis imaging. *Amino Acids* **41**, 1113-1122 (2011).
138. N. H. Lim *et al.*, In vivo imaging of matrix metalloproteinase 12 and matrix metalloproteinase 13 activities in the mouse model of collagen-induced arthritis. *Arthritis Rheumatol.* **66**, 589-598 (2014).
139. N. H. Lim, E. Meinjohanns, M. Meldal, G. Bou-Gharios, H. Nagase, In vivo imaging of MMP-13 activity in the murine destabilised medial meniscus surgical model of osteoarthritis. *Osteoarthr. Cartil.* **22**, 862-868 (2014).
140. A. Duro-Castano *et al.*, In vivo imaging of MMP-13 activity using a specific polymer-FRET peptide conjugate detects early osteoarthritis and inhibitor efficacy. *Adv. Funct. Mater.* **28**, 1-9 (2018).
141. T. Kodadek, The rise, fall and reinvention of combinatorial chemistry. *Chem. Commun.* **47**, 9757-9763 (2011).
142. M. Lebl, Parallel personal comments on "classical" papers in combinatorial chemistry. *J. Comb. Chem.* **1**, 3-24 (1999).
143. R. B. Merrifield, Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154 (1963).
144. Á. Furka, I. Hargittai, The concealed side of the history of combinatorial chemistry. *Period. Polytech. Chem. Eng.* **48**, 13-19 (2004).
145. Á. Furka, F. Sebestyen, M. Asgedom, G. Dibo, General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **37**, 487-493 (1991).
146. K. S. Lam *et al.*, A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **354**, 82-84 (1991).
147. C. L. Chen, P. Strop, M. Lebl, K. S. Lam, One bead-one compound combinatorial peptide library: Different types of screening. *Methods Enzymol.* **267**, 211-219 (1996).
148. K. S. Lam, M. Lebl, V. Krchnak, The "one-bead-one-compound" combinatorial library method. *Chem. Rev.* **97**, 411-448 (1997).
149. R. Liu, X. Li, K. S. Lam, Combinatorial chemistry in drug discovery. *Curr. Opin. Chem. Biol.* **38**, 117-126 (2017).
150. R. Epton, *Innovation and perspectives in solid phase synthesis: peptides, proteins and nucleic acids: biological and biomedical applications* (Mayflower Worldwide, Birmingham, 1994).

151. M. Meldal, Pega - a Flow Stable Polyethylene-Glycol Dimethyl Acrylamide Copolymer for Solid-Phase Synthesis. *Tetrahedron Lett.* **33**, 3077-3080 (1992).
152. M. Meldal, F. I. Auzanneau, O. Hindsgaul, M. M. Palcic, A Pega Resin for Use in the Solid-Phase Chemical-Enzymatic Synthesis of Glycopeptides. *J. Chem. Soc., Chem. Commun.*, 1849-1850 (1994).
153. F. I. Auzanneau, M. Meldal, K. Bock, Synthesis, characterization and biocompatibility of PEGA resins. *J. Pept. Sci.* **1**, 31-44 (1995).
154. M. Renil, M. Meldal, Synthesis and Application of a Pega Polymeric Support for High-Capacity Continuous-Flow Solid-Phase Peptide-Synthesis. *Tetrahedron Lett.* **36**, 4647-4650 (1995).
155. M. Renil, M. Ferreras, J. M. Delaisse, N. T. Foged, M. Meldal, PEGA supports for combinatorial peptide synthesis and solid-phase enzymatic library assays. *J. Pept. Sci.* **4**, 195-210 (1998).
156. M. Meldal, Smart combinatorial assays for the determination of protease activity and inhibition. *QSAR & Comb. Sci.* **24**, 1141-1148 (2005).
157. J. O. Duus, M. Meldal, J. R. Winkler, Fluorescence energy-transfer probes of conformation in peptides: The 2-aminobenzamide/nitrotyrosine pair. *J. Phys. Chem. B* **102**, 6413-6418 (1998).
158. P. M. St Hilaire, M. Willert, M. A. Juliano, L. Juliano, M. Meldal, Fluorescence-quenched solid phase combinatorial libraries in the characterization of cysteine protease substrate specificity. *J. Comb. Chem.* **1**, 509-523 (1999).
159. V. Ferro, M. Meldal, K. Bock, Synthesis of 2'-Acylated and 2''-O-Acylated Maltotriosides as Potential Fluorescence-Quenched Substrates for Alpha-Amylase. *J. Chem. Soc. Perk. T. 1*, 2169-2176 (1994).
160. M. Meldal, I. Svendsen, K. Breddam, F. I. Auzanneau, Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 3314-3318 (1994).
161. E. Del Nery *et al.*, Characterization of the substrate specificity of the major cysteine protease (cruzipain) from *Trypanosoma cruzi* using a portion-mixing combinatorial library and fluorogenic peptides. *Biochem. J.* **323**, 427-433 (1997).
162. A. Davy *et al.*, Substrate specificity of barley cysteine endoproteases EP-A and EP-B. *Plant Physiol.* **117**, 255-261 (1998).
163. J. C. Spetzler, V. Westphal, J. R. Winther, M. Meldal, Fluorescent quenched libraries for monitoring chemical and enzymatic disulfide bond formation and cleavage. *Peptides*, 63-66 (1998).
164. P. M. St Hilaire *et al.*, The substrate specificity of a recombinant cysteine protease from *Leishmania mexicana*: Application of a combinatorial peptide library approach. *ChemBioChem* **1**, 115-122 (2000).
165. L. C. Alves *et al.*, Substrate specificity of recombinant cysteine proteinase, CPB, of *Leishmania mexicana*. *Mol. Biochem. Parasitol.* **116**, 1-9 (2001).
166. H. R. Stennicke, M. Renatus, M. Meldal, G. S. Salvesen, Internally quenched fluorescent peptide substrates disclose the subsite preferences of human caspases 1, 3, 6, 7 and 8. *Biochem. J.* **350**, 563-568 (2000).
167. K. Breddam, M. Meldal, Substrate Preferences of Glutamic-Acid-Specific Endopeptidases Assessed by Synthetic Peptide-Substrates Based on Intramolecular Fluorescence Quenching. *Eur. J. Biochem.* **206**, 103-107 (1992).
168. B. J. Battersby *et al.*, Toward larger chemical libraries: Encoding with fluorescent colloids in combinatorial chemistry. *J. Am. Chem. Soc.* **122**, 2138-2139 (2000).
169. H. Fenniri *et al.*, Towards the DRED of Resin-Supported Combinatorial Libraries: A Non-Invasive Methodology Based on Bead Self-Encoding and Multispectral Imaging *Angew. Chem. Int. Ed.* **39**, 4483-4485 (2000).

170. K. Braeckmans *et al.*, Encoding microcarriers by spatial selective photobleaching. *Nat. Mater.* **2**, 169-173 (2003).
171. M. Meldal, S. F. Christensen, Microparticle Matrix Encoding of Beads. *Angew. Chem. Int. Ed.* **49**, 3473-3476 (2010).
172. J. E. Rasmussen *et al.*, Small-Molecule Affinity Ligands for Protein Purification: Combined Computational Enrichment and Automated In-line Screening of an Optically Encoded Library. *Angew. Chem. Int. Ed.* **49**, 3477-3480 (2010).
173. The Cell Division Laboratory, SDS-PAGE gel recipes, <https://celldivisionlab.com/2013/08/12/sds-page-gel-recipes/> (accessed March 27th, 2021).
174. A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850-858 (1996).
175. Abcam, Gelatin zymography protocol, <https://docs.abcam.com/pdf/protocols/gelatin-zymography-protocol-mmp-9.pdf> (accessed March 28th, 2021).
176. J. W. Verhoeven, Glossary of terms used in photochemistry (IUPAC Recommendations 1996). *Pure Appl. Chem.* **68**, 2223-2286 (1996).
177. J. G. Bieth, Theoretical and practical aspects of proteinase inhibition kinetics. *Methods Enzymol.* **248**, 59-84 (1995).
178. Y. Liu *et al.*, Use of a fluorescence plate reader for measuring kinetic parameters with inner filter effect correction. *Anal. Biochem.* **267**, 331-335 (1999).
179. L. Mosyak *et al.*, Crystal structures of the two major aggrecan degrading enzymes, ADAMTS4 and ADAMTS5. *Protein Sci.* **17**, 16-21 (2008).
180. M. Meldal, Multiple Column Synthesis of Quenched Solid-Phase Bound Fluorogenic Substrates for Characterization of Endoprotease Specificity. *Methods* **6**, 417-424 (1994).
181. E. Kaiser, R. L. Colescott, C. D. Bossinger, P. I. Cook, Color test for detection of free terminal amino groups in solid-phase synthesis of peptides. *Anal. Biochem.* **34**, 595-598 (1970).
182. H. X. Hu *et al.*, Sustainable Flow Synthesis of Encoded Beads for Combinatorial Chemistry and Chemical Biology. *ACS Comb. Sci.* **20**, 492-498 (2018).
183. M. Li, C. Hoeck, S. Schoffelen, C. H. Gotfredsen, M. Meldal, Specific Electrostatic Molecular Recognition in Water. *Chem. Eur. J.* **22**, 7206-7214 (2016).
184. A. Basi *et al.* (2015) Protein expression handbook. ed Thermo Fisher Scientific, p 115.
185. M. M. Fowkes, N. H. Lim, Purification and Activity Determination of ADAMTS-4 and ADAMTS-5 and Their Domain Deleted Mutants. *Methods Mol. Biol.* **2043**, 75-91 (2020).
186. R. Behrendt, S. Huber, R. Marti, P. White, New t-butyl based aspartate protecting groups preventing aspartimide formation in Fmoc SPPS. *J. Pept. Sci.* **21**, 680-687 (2015).
187. R. Behrendt, S. Huber, P. White, Preventing aspartimide formation in Fmoc SPPS of Asp-Gly containing peptides—practical aspects of new trialkylcarbinol based protecting groups. *J. Pept. Sci.* **22**, 92-97 (2016).
188. M. Berkmen, Production of disulfide-bonded proteins in Escherichia coli. *Protein Expr. Purif.* **82**, 240-251 (2012).
189. J. C. A. Bardwell, "5 - Disulfide Bond Formation Enzymes" in The Enzymes, R. E. Dalbey, C. M. Koehler, F. Tamanoi, Eds. (Academic Press, 2007), vol. 25, pp. 111-128.
190. J. M. Longpre *et al.*, Characterization of proADAMTS5 processing by proprotein convertases. *Int. J. Biochem. Cell Biol.* **41**, 1116-1126 (2009).
191. H. Karring *et al.*, Evidence against a blood derived origin for transforming growth factor beta induced protein in corneal disorders caused by mutations in the TGFBI gene. *Mol. Vis.* **13**, 997-1004 (2007).
192. M. H. Lee, V. Knauper, J. D. Becherer, G. Murphy, Full-length and N-TIMP-3 display equal inhibitory activities toward TNF-alpha convertase. *Biochem. Biophys. Res. Commun.* **280**, 945-950 (2001).
193. C. G. Knight, Active-site titration of peptidases. *Methods Enzymol.* **248**, 85-101 (1995).

194. Thermo Fisher Scientific, A theoretical and practical guide for spectrophotometric determination of protein concentrations at 280 nm. *Technical note* (2018).
195. M. Kashiwagi, M. Tortorella, H. Nagase, K. Brew, TIMP-3 is a potent inhibitor of aggrecanase 1 (ADAM-TS4) and aggrecanase 2 (ADAM-TS5). *J. Biol. Chem.* **276**, 12501-12504 (2001).
196. M. Kveiborg *et al.*, Selective inhibition of ADAM12 catalytic activity through engineering of tissue inhibitor of metalloproteinase 2 (TIMP-2). *Biochem. J.* **430**, 79-86 (2010).
197. A. Amour *et al.*, TNF-alpha converting enzyme (TACE) is inhibited by TIMP-3. *FEBS Lett.* **435**, 39-44 (1998).
198. E. H. Kerkvliet *et al.*, Low molecular weight inhibitors of matrix metalloproteinases can enhance the expression of matrix metalloproteinase-2 (gelatinase A) without inhibiting its activation. *Cancer* **97**, 1582-1588 (2003).
199. K. Fushimi, L. Troeberg, H. Nakamura, N. H. Lim, H. Nagase, Functional differences of the catalytic and non-catalytic domains in human ADAMTS-4 and ADAMTS-5 in aggrecanolytic activity. *J. Biol. Chem.* **283**, 6706-6716 (2008).
200. S. Santamaria, H. Nagase, Measurement of Protease Activities Using Fluorogenic Substrates. *Methods Mol. Biol.* **1731**, 107-122 (2018).
201. S. Santamaria *et al.*, Exosites in Hypervariable Loops of ADAMTS Spacer Domains control Substrate Recognition and Proteolysis. *Sci. Rep.* **9**, 1-12 (2019).
202. J. Westling *et al.*, ADAMTS4 cleaves at the aggrecanase site (Glu373-Ala374) and secondarily at the matrix metalloproteinase site (Asn341-Phe342) in the aggrecan interglobular domain. *J. Biol. Chem.* **277**, 16059-16066 (2002).
203. J. L. Lauer-Fields *et al.*, Substrate conformation modulates aggrecanase (ADAMTS-4) affinity and sequence specificity. Suggestion of a common topological specificity for functionally diverse proteases. *J. Biol. Chem.* **282**, 142-150 (2007).
204. Y. C. Yu, M. Tirrell, G. B. Fields, Minimal lipidation stabilizes protein-like molecular architecture. *J. Am. Chem. Soc.* **120**, 9979-9987 (1998).
205. M. Maier *et al.*, Highly Site Specific, Protease Cleavable, Hydrophobic Peptide-Polymer Nanoparticles. *Macromolecules* **44**, 6258-6267 (2011).
206. J. M. Berg, J. L. Tymoczko, L. Stryer, G. J. Gatto, *Biochemistry* (W.H. Freeman and Company, Palgrave Macmillan, New York, Basingstoke, ed. 7th, 2012).
207. K. Holmberg, Interactions between surfactants and hydrolytic enzymes. *Colloids Surf. B. Biointerfaces* **168**, 169-177 (2018).
208. H. I. Park, S. Lee, A. Ullah, Q. Cao, Q. X. Sang, Effects of detergents on catalytic activity of human endometase/matrilysin 2, a putative cancer biomarker. *Anal. Biochem.* **396**, 262-268 (2010).
209. D. J. Matthews, J. A. Wells, Substrate phage: selection of protease substrates by monovalent phage display. *Science* **260**, 1113-1117 (1993).
210. G. P. Smith, Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317 (1985).
211. J. Bazan, I. Calkosinski, A. Gamian, Phage display—a powerful technique for immunotherapy: 1. Introduction and potential of therapeutic applications. *Hum. Vaccin. Immunother.* **8**, 1817-1828 (2012).
212. K. C. Desch *et al.*, Probing ADAMTS13 substrate specificity using phage display. *PLoS One* **10**, e0122931 (2015).
213. J. Zhou *et al.*, Deep profiling of protease substrate specificity enabled by dual random and scanned human proteome substrate phage libraries. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 25464-25475 (2020).
214. L. Michaelis, M. L. Menten, Die Kinetik der Invertinwirkung. *Biochem. Z.* **49**, 333-369 (1913).
215. A. M. Embaby, L. Lelieveldt, F. Diness, M. Meldal, Metallo-Organozymes with Specific Proteolytic Activity. *Chem. Eur. J.* **24**, 17424-17428 (2018).

216. M. O. Palmier, S. R. Van Doren, Rapid determination of enzyme kinetics from fluorescence: overcoming the inner filter effect. *Anal. Biochem.* **371**, 43-51 (2007).
217. D. Madsen *et al.*, Design and Combinatorial Development of Shield-1 Peptide Mimetics Binding to Destabilized FKBP12. *ACS Comb. Sci.* **22**, 156-164 (2020).
218. M. Uhlen *et al.*, Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **28**, 1248-1250 (2010).
219. The Human Protein Atlas, ADAMTS-5 (accessed 25th February 2021).
220. L. Rovati, P. Fabbri, F. Pilati, "Plastic Optical Fiber pH Sensor Using a Sol-Gel Sensing Matrix" in *Fiber Optic Sensors*, M. Yasin, Ed. (IntechOpen, 2012), pp. 415-438.
221. M. Bredel, E. Jacoby, Chemogenomics: an emerging strategy for rapid target and drug discovery. *Nat. Rev. Genet.* **5**, 262-275 (2004).
222. N. Haraguchi *et al.*, High expression of ADAMTS5 is a potent marker for lymphatic invasion and lymph node metastasis in colorectal cancer. *Mol. Clin. Oncol.* **6**, 130-134 (2017).
223. M. Nakada *et al.*, Human glioblastomas overexpress ADAMTS-5 that degrades brevican. *Acta Neuropathol.* **110**, 239-246 (2005).
224. S. Santamaria, ADAMTS-5: A difficult teenager turning 20. *Int. J. Exp. Pathol.* **101**, 4-20 (2020).
225. F. Monibi *et al.*, Identification of Synovial Fluid Biomarkers for Knee Osteoarthritis and Correlation with Radiographic Assessment. *J. Knee Surg.* **29**, 242-247 (2016).
226. B. J. Heard *et al.*, Matrix metalloproteinase protein expression profiles cannot distinguish between normal and early osteoarthritic synovial fluid. *BMC Musculoskelet. Disord.* **13**, 126 (2012).
227. S. Roberts *et al.*, ADAMTS-4 activity in synovial fluid as a biomarker of inflammation and effusion. *Osteoarthr. Cartil.* **23**, 1622-1626 (2015).
228. H. Y. Hu *et al.*, DOTAM derivatives as active cartilage-targeting drug carriers for the treatment of osteoarthritis. *Bioconjugate Chem.* **26**, 383-388 (2015).
229. D. A. Rothenfluh, H. Bermudez, C. P. O'Neil, J. A. Hubbell, Biofunctional polymer nanoparticles for intra-articular targeting and retention in cartilage. *Nat. Mater.* **7**, 248-254 (2008).
230. H. Hedlund *et al.*, Association of the aggrecan keratan sulfate-rich region with collagen in bovine articular cartilage. *J. Biol. Chem.* **274**, 5777-5781 (1999).
231. M. Poreba, Protease-activated prodrugs: strategies, challenges, and future directions. *FEBS J.* **287**, 1936-1969 (2020).
232. J. M. Connors *et al.*, Brentuximab Vedotin with Chemotherapy for Stage III or IV Hodgkin's Lymphoma. *N. Engl. J. Med.* **378**, 331-344 (2018).
233. H. Tilly *et al.*, Polatuzumab vedotin in combination with immunochemotherapy in patients with previously untreated diffuse large B-cell lymphoma: an open-label, non-randomised, phase 1b-2 study. *Lancet Oncol.* **20**, 998-1010 (2019).
234. F. Brebion *et al.*, Discovery of GLPG1972/S201086, a Potent, Selective, and Orally Bioavailable ADAMTS-5 Inhibitor for the Treatment of Osteoarthritis. *J. Med. Chem.* **64**, 2937-2952 (2021).
235. M. Fava *et al.*, Role of ADAMTS-5 in Aortic Dilatation and Extracellular Matrix Remodeling. *Arterioscler. Thromb. Vasc. Biol.* **38**, 1537-1548 (2018).
236. J. Chung *et al.*, Discovery of novel peptides targeting pro-atherogenic endothelium in disturbed flow regions -Targeted siRNA delivery to pro-atherogenic endothelium *in vivo*. *Sci. Rep.* **6**, 1-15 (2016).

Chapter 8. Appendices

8.1. Chapter 3 Appendix

8.1.1. Figures

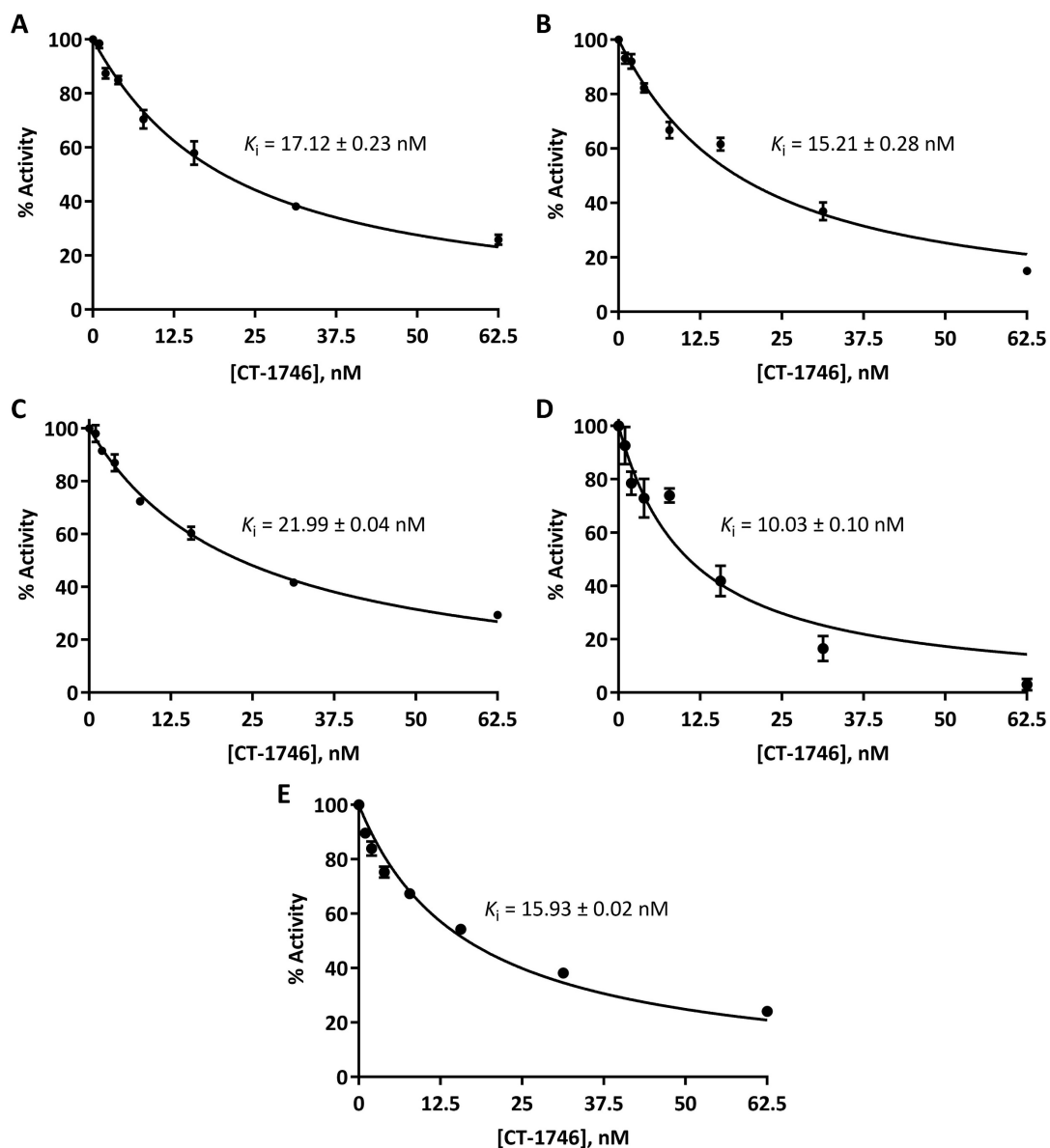


Figure 8.1. Titration curves for ADAMTS-4 data shown in Table 3.1. Entries A–E for the TS4-4 proteinase. All rates are expressed as a function of percentage activity without inhibitor using the 5-FAM-AE~LQGRPISIAK-TAMRA substrate (0.5 μ M). All values are given as mean percentage activity $\pm$ SEM with $n = 3$ wells used for each CT-1746 concentration at 37 $^{\circ}$ C.

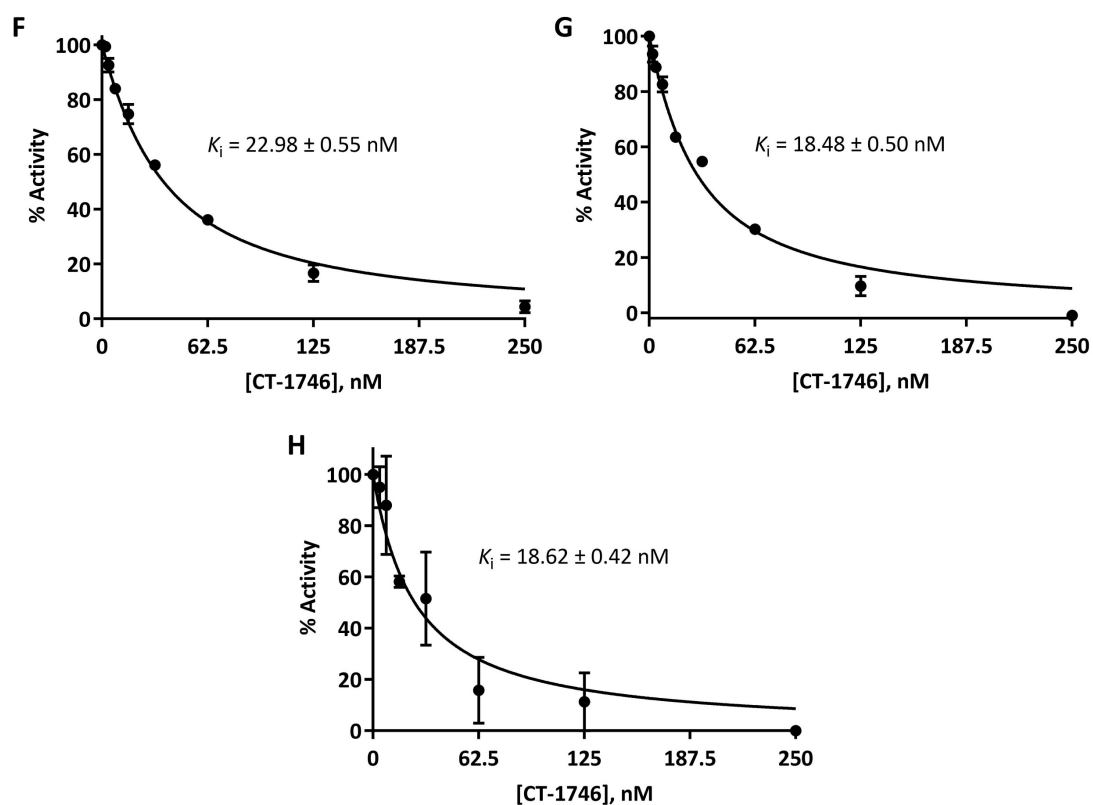


Figure 8.2. Titration curves for ADAMTS-5 data shown in Table 3.1. Entries F–H for the TS5-5 proteinase. All rates are expressed as a function of percentage activity without inhibitor using the Abz-TESE~SRGAIY-Dpa-KK-NH₂ substrate (20 μ M). All values are given as mean percentage activity $\pm$ SEM with $n = 3$ wells used for each CT-1746 concentration at 37 $^{\circ}$ C.

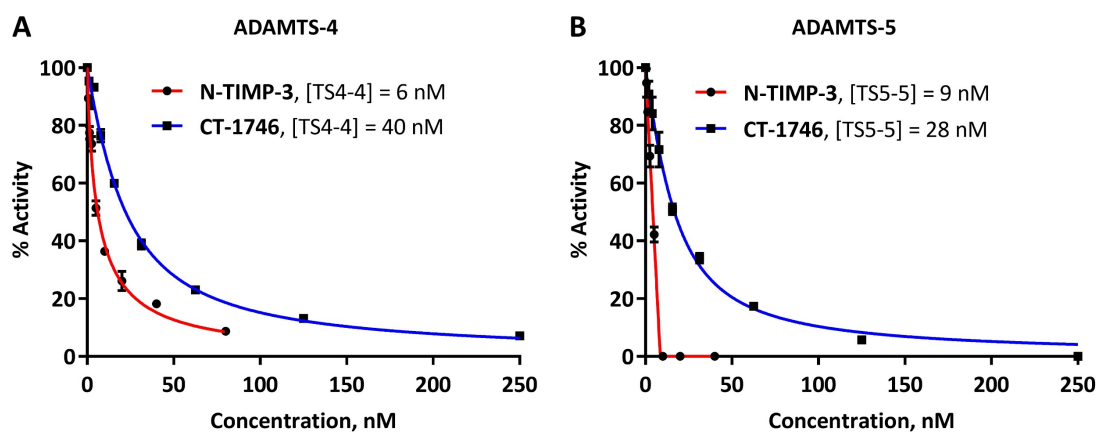


Figure 8.3. Determination of ADAMTS-4 and ADAMTS-5 concentrations listed in Table 3.3 through active-site titration with N-TIMP-3 and CT-1746. Curves for active-site titration of TS4-4 (A) and TS5-5 (B) with N-TIMP-3 (red) and CT-1746 (blue). Rates are expressed as a function of percentage activity without inhibitor using the 5-FAM-AE~LQGRPISIAK-TAMRA (0.5 μ M) and Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μ M) substrates for TS4-4 and TS5-5, respectively. Values are given as mean percentage activity $\pm$ SEM. For each proteinase, $n = 3$ wells were used for each CT-1746 concentration at 37 $^{\circ}$ C.

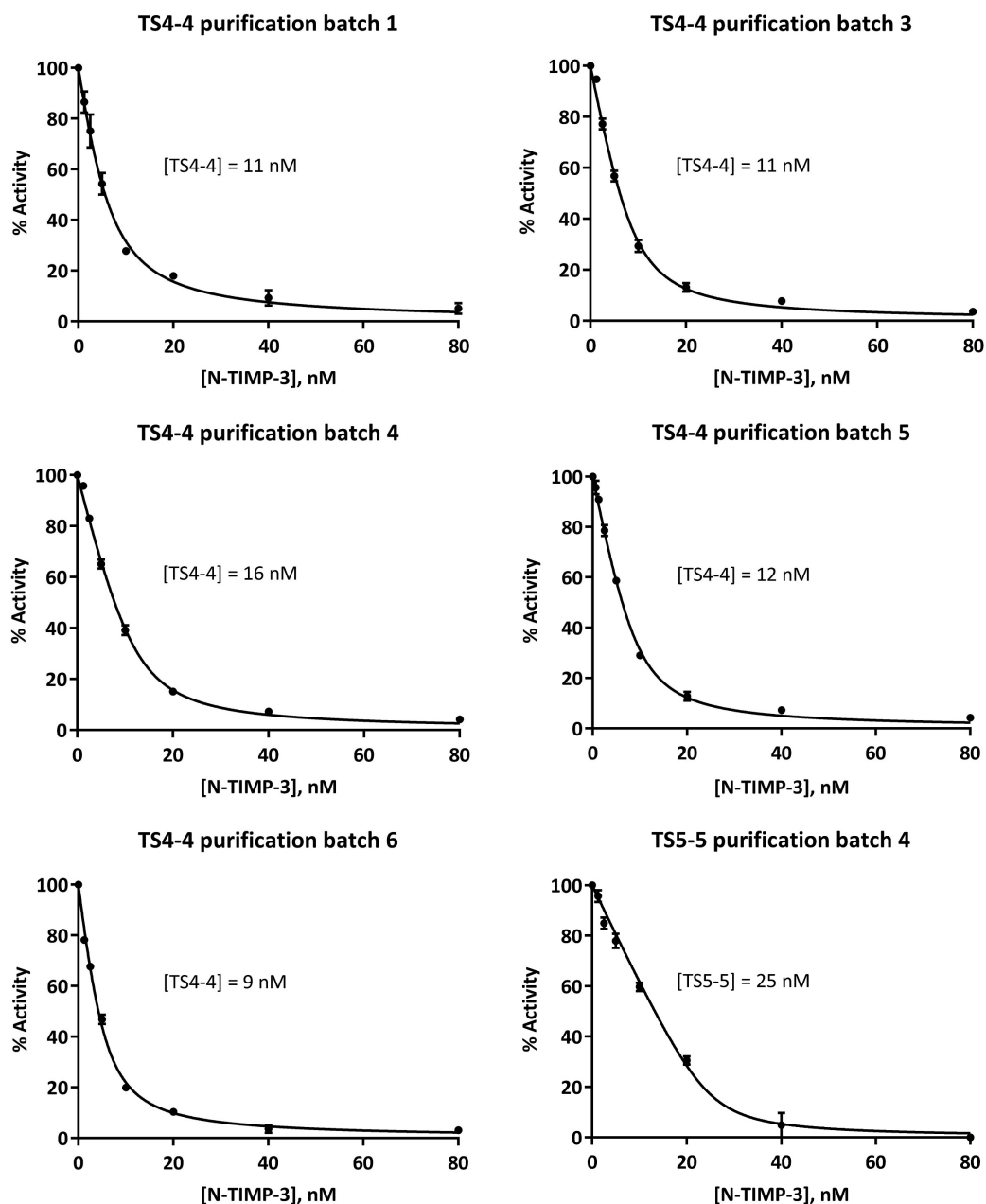


Figure 8.4. Determination of ADAMTS-4 and ADAMTS-5 amounts from different purification batches listed in Table 3.4. Curves for active-site titration of TS4-4 and TS5-5 purification batches with N-TIMP-3. Rates are expressed as a function of percentage activity without inhibitor using the 5-FAM-AE~LQGRPISIAK-TAMRA (0.5 μ M) and Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μ M) substrates for TS4-4 and TS5-5, respectively. Values are given as mean percentage activity $\pm$ SEM. For each proteinase, $n = 3$ wells were used for each N-TIMP-3 concentration at 37 $^{\circ}$ C.

8.2. Chapter 4 Appendix

8.2.1. Figures and tables

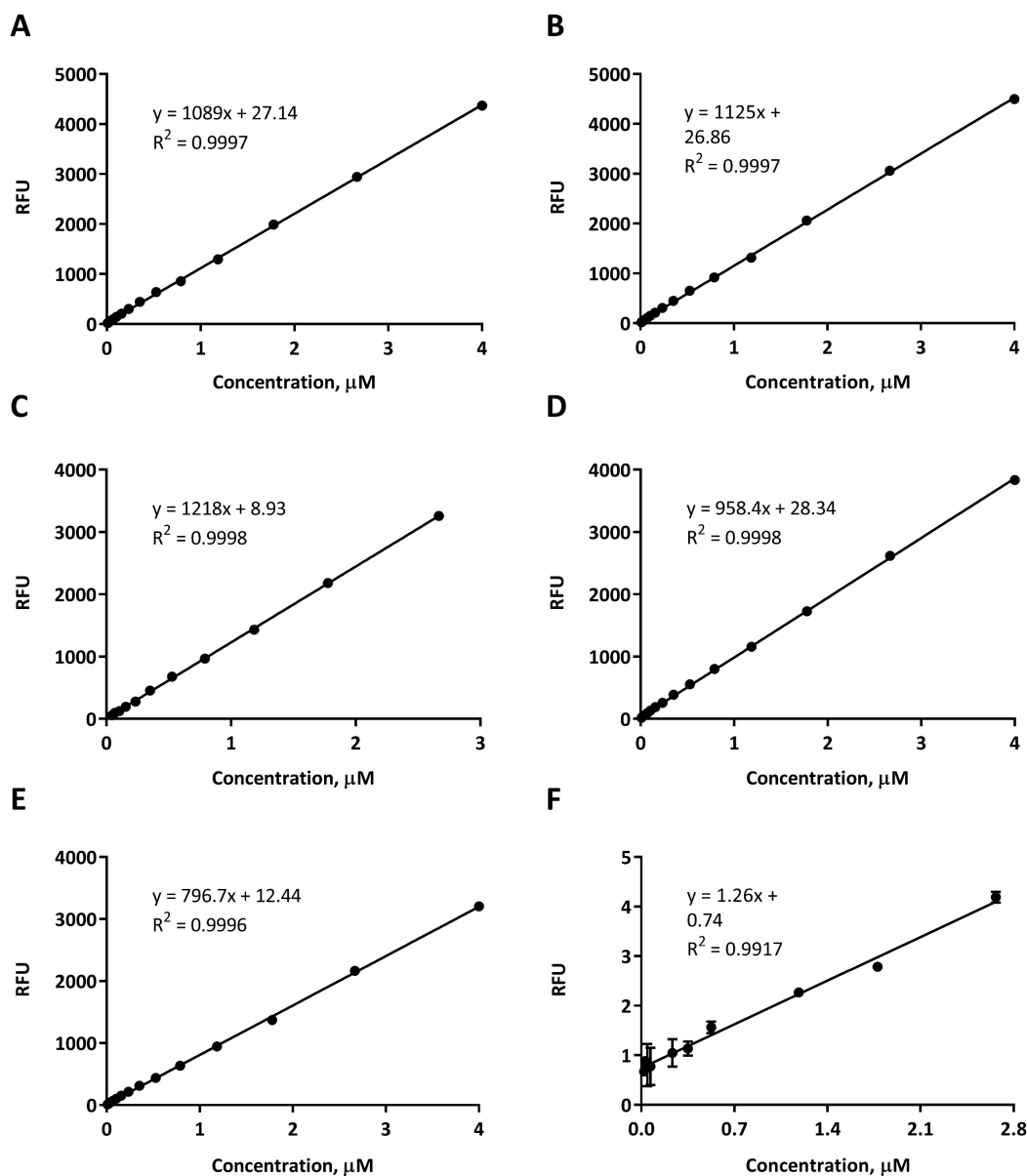


Figure 8.5. Calibration curves for the 5-FAM fluorophore at different concentrations of Brij-35 surfactant. For each calibration curve, the relative fluorescence intensity (RFU) was measured at each fluorophore concentration (μM) and is expressed as the mean $\pm$ SEM for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. Brij-35 concentrations (v/v): (A) 0%; (B) 0.0005%; (C) 0.005%; (D) 0.02%; (E) 0.05%; (F) 5%.

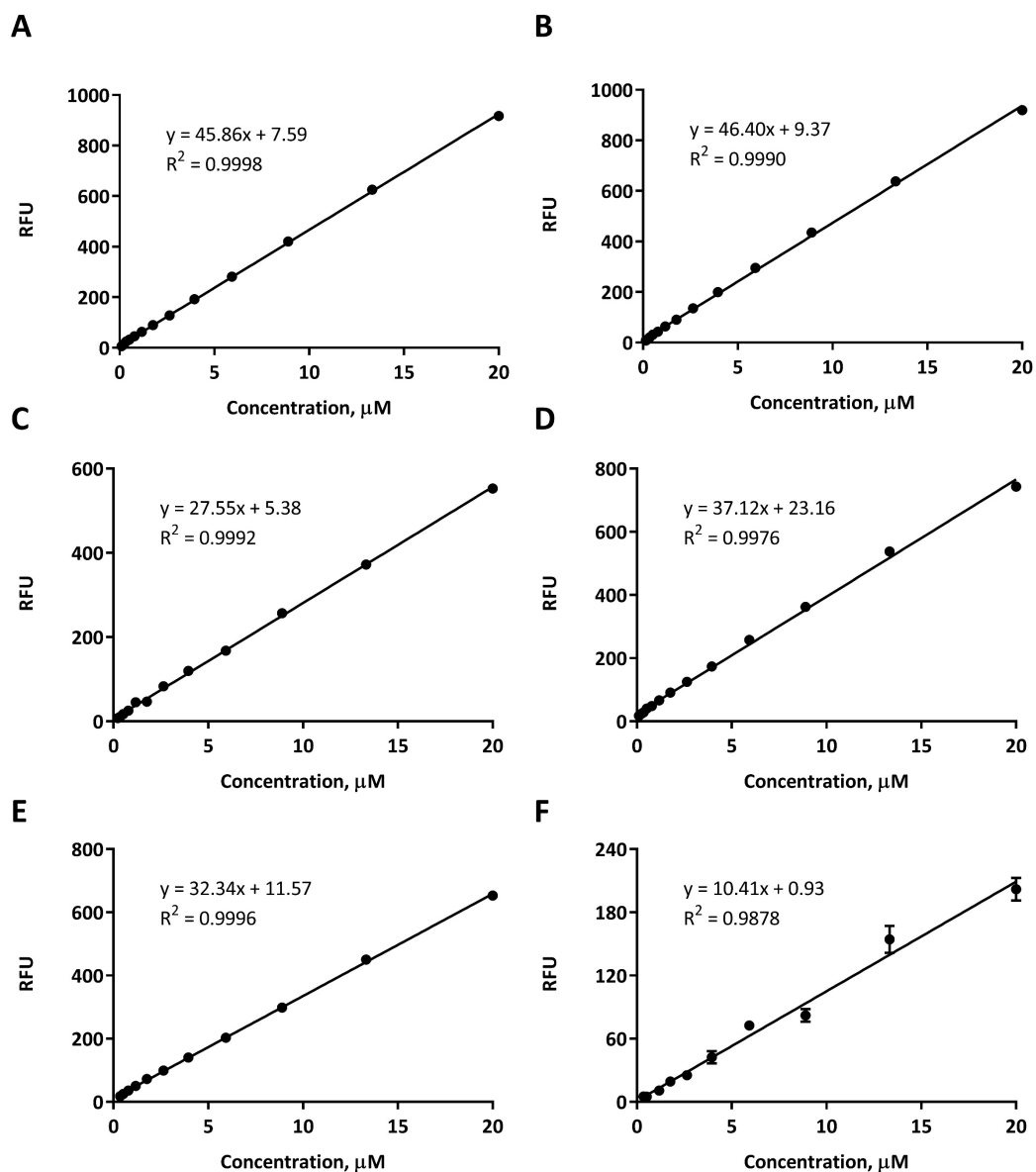


Figure 8.6. Calibration curves for the Abz fluorophore at different concentrations of Brij-35 surfactant. For each calibration curve, the relative fluorescence intensity (RFU) was measured at each fluorophore concentration (μM) and is expressed as the mean ± SEM for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. Brij-35 concentrations (v/v): (A) 0%; (B) 0.0005%; (C) 0.005%; (D) 0.02%; (E) 0.05%; (F) 5%.

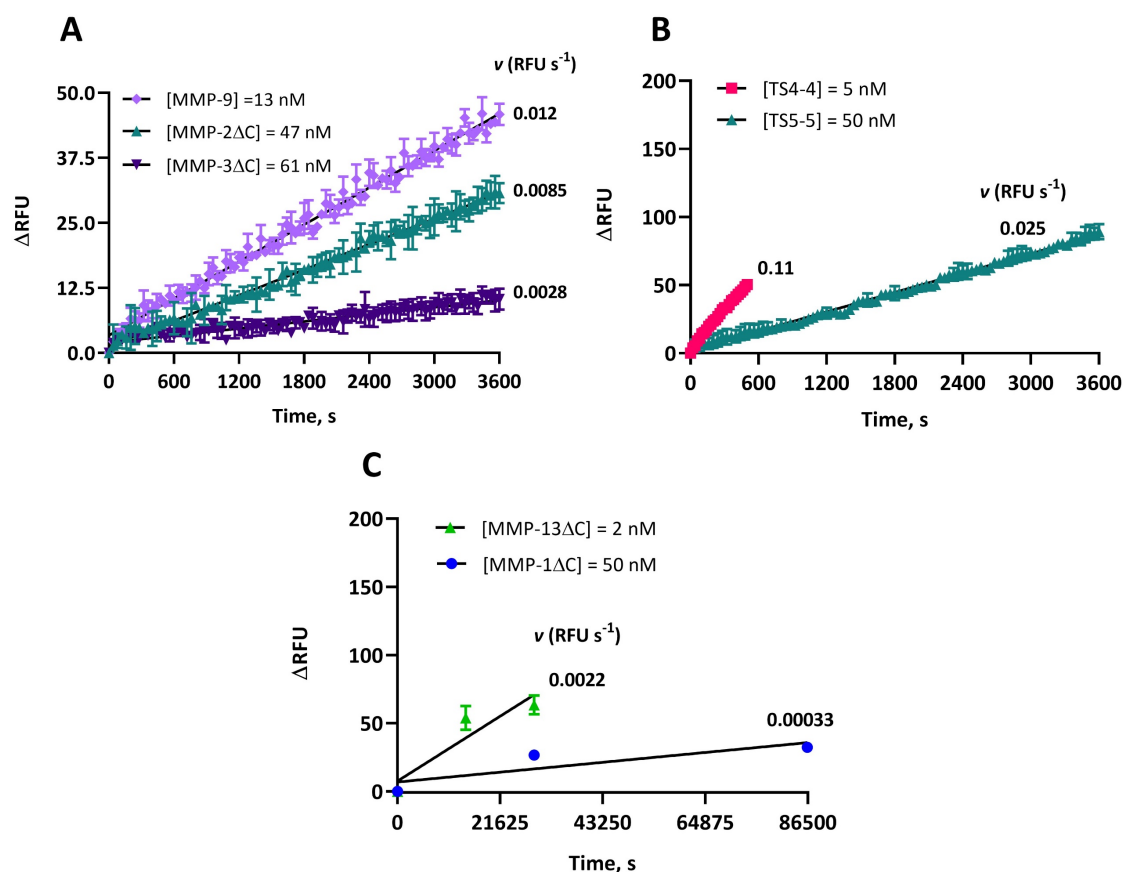


Figure 8.7. Cleavage of FRET substrate 1 by ADAMTS and MMP metalloproteinases. Representative rate data for cleavage of 5-FAM-AE~LQGRPISIAK-TAMRA ($0.5 \mu\text{M}$) by (A) MMP-2 Δ C, MMP-3 Δ C and MMP-9; (B) TS4-4 and TS5-5; and (C) MMP-1 Δ C and MMP-13 Δ C, all fitted by linear regression. The rate (v) is given in RFU s^{-1} . Each time point is expressed as ΔRFU ($\text{RFU at time } t - \text{RFU at time } 0$) $\pm$ SEM for $n = 3$ wells from a single experiment performed in triplicate at 37°C . Cat Dis = catalytic and disintegrin domain, ΔC = catalytic domain, MMP-9 = full length. *Please note that data is shown for when the cleavage rate is linearly proportional to time.*

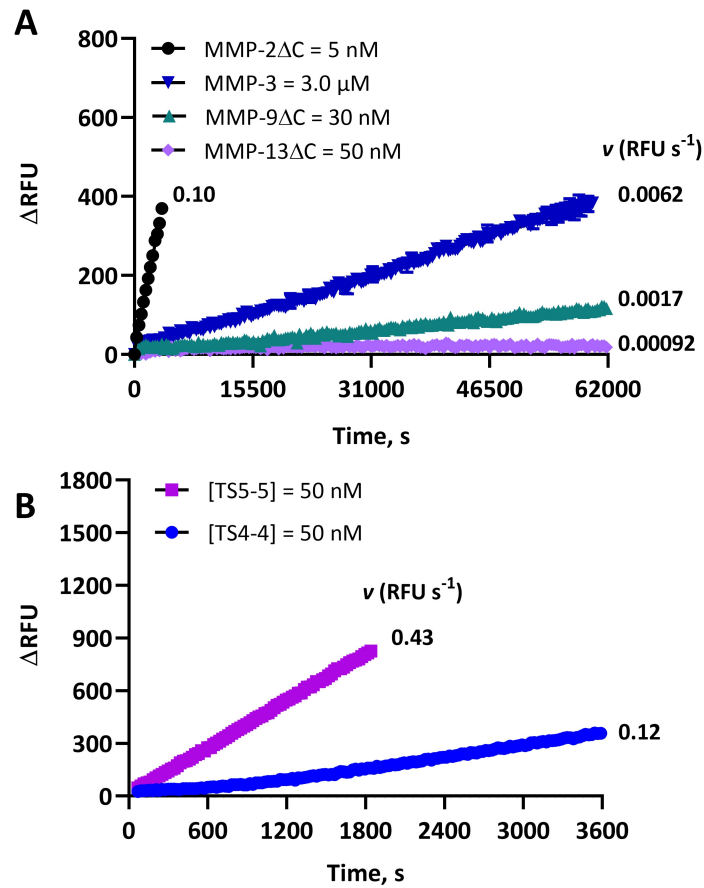


Figure 8.8. Cleavage of FRET substrate 2 by ADAMTS and MMP metalloproteinases. Representative rate data for cleavage of Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μM) by (A) MMP-2 ΔC , MMP-3, MMP-9 ΔC and MMP-13 ΔC ; (B) TS4-4 and TS5-5, all fitted by linear regression. No cleavage was observed after 72 h with MMP-1 at a concentration of 92 nM. The rate (v) is given in RFU s^{-1} . Each time point is expressed as ΔRFU ($\text{RFU at time } t - \text{RFU at time } 0$) $\pm$ SEM for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. Cat Dis = catalytic and disintegrin domain, ΔC = catalytic domain, MMP-1 and MMP-3 = full length. *Please note that data is shown for when the cleavage rate is linearly proportional to time.*

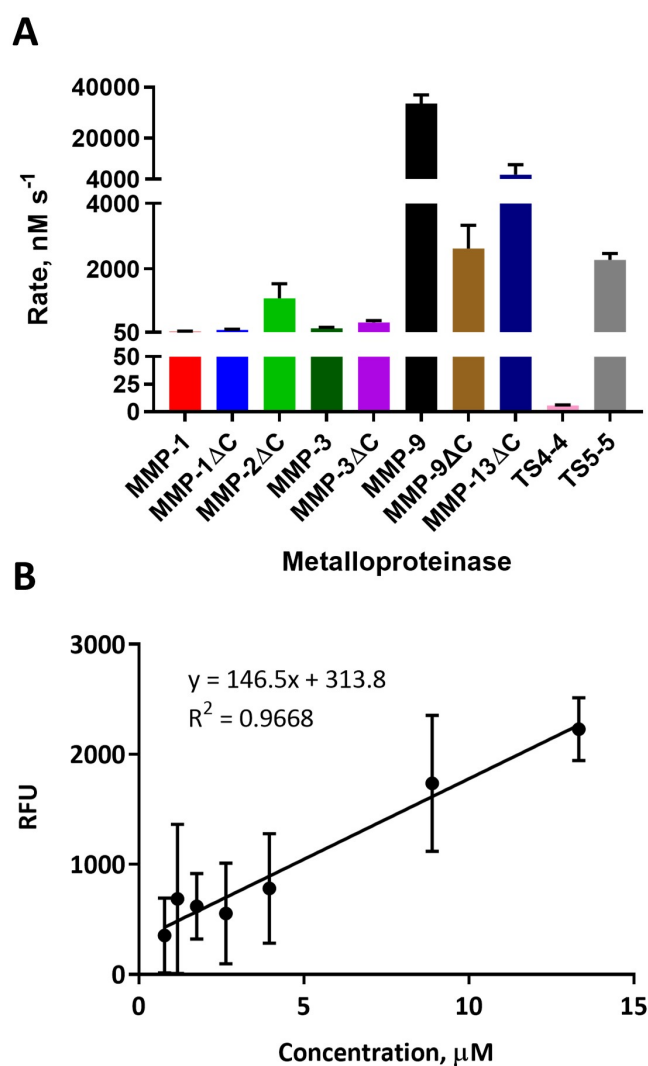


Figure 8.9. Activity of metalloproteinases from ADAMTS-4 and ADAMTS-5 selectivity studies. (A) rate data for cleavage of Mca-PLG~L-Dpa-AR-NH₂ (1.5 μM) by MMPs, 5-FAM-AE~LQGRPISIAK-TAMRA (0.5 μM) by TS4-4 and Abz-TESE~SRGAIY-Dpa-KK-NH₂ (20 μM) by TS5-5. (B) Calibration curve for the Mca fluorophore at a concentration of 0.005% (v/v) Brij-35 surfactant. The relative fluorescence intensity (RFU) was measured at each fluorophore concentration (μM) and is expressed as the mean ± SEM for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. All rates in (A) are expressed in nM s⁻¹ using the calibration curves from Figure 8.5C, 8.6C and 8.9B. Each rate is expressed as the mean ± SEM for $n = 6$ -24 wells from 2-8 independent experiments performed in triplicate at 37 °C. Cat Dis = catalytic and disintegrin domain, ΔC = catalytic domain, MMP-1 and MMP-3 = full length.

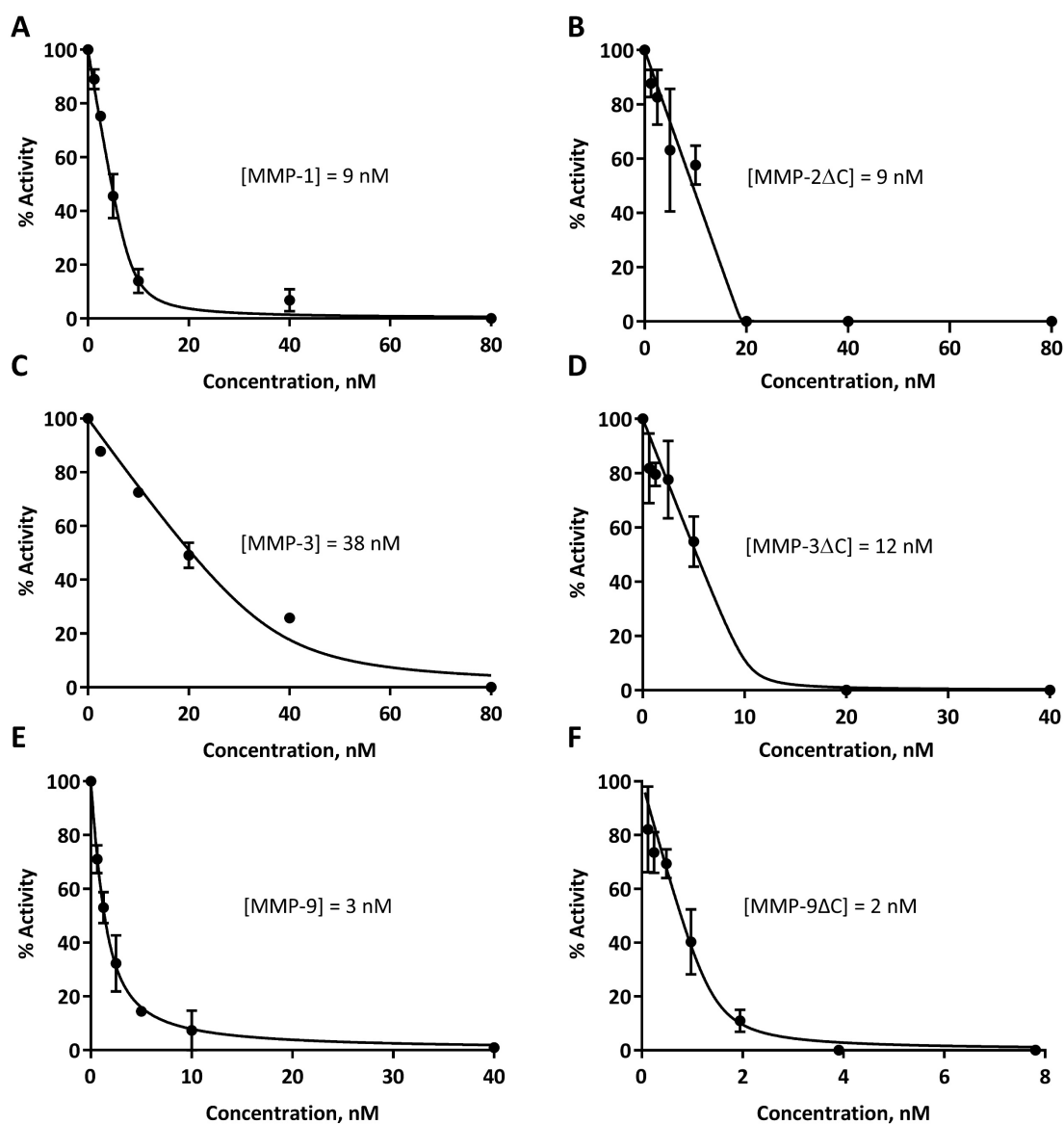


Figure 8.10. Determination of MMP concentrations from ADAMTS-4 and ADAMTS-5 selectivity studies through active-site titration with N-TIMP-3 or CT-1746. Curves for active-site titration of MMPs with N-TIMP-3 or CT-1746 using the Mca-PLG~L-Dpa-AR-NH₂ FRET substrate (1.5 μ M). Rates are expressed as a function of percentage activity without inhibitor. Values are given as mean percentage activity $\pm$ SEM. For each proteinase, $n = 3$ wells were used for each N-TIMP-3 or CT-1746 concentration at 37 $^{\circ}$ C.

Table 8.1. Known peptide substrates cleaved by ADAMTS-5 compiled from CutDB, BRENDA and MEROPS online databases.

P4	P3	P2	P1	P1'	P2'	P3'	P4'	P5'	P6'	P7'	P8'	P9'
R	I	P	E	G	A	T	H	I	K	V	R	Q
T	E	G	E	A	R	G	S	V	I	Dpa	K	K
G	E	L	E	G	R	G	T	I	D	I	S	G
S	E	L	E	G	R	G	T	I	G	I	S	G
K	E	E	E	G	L	G	S	V	E	L	S	G
I	S	Q	E	L	G	Q	R	P	P	V	T	H
T	A	Q	E	A	G	E	G	P	S	G	I	L
T	E	G	E	A	R	G	S	V	I	L	T	V
S	D	V	M	G	R	G	H	A	R	L	V	H
D	V	Q	E	F	R	G	V	T	A	V	I	R
T	K(Mca)	G	E	L	E	G	R	G	T	K(Dnp)	G	I
K(Mca)	G	E	L	E	G	R	G	T	K(Dnp)	G	I	S
G	E	L	E	G	R	G	T	K(Dnp)	G	I	S	G
L	E	G	R	G	T	K(Dnp)	G	I	S	G	P	Hyp
K(Mca)	K	Q	E	F	R	G	Q	T	K(Dnp)-NH ₂	-	-	-
R	M	N	M	C	I	E	M	G	G	N	P	L
A	F	K	<u>C</u>	L	K	D	G	A	G	D	V	A
D	A	M	S	L	D	G	G	F	V	Y	I	A
A	F	R	<u>C</u>	L	V	E	K	G	D	V	A	K
S	V	A	C	V	K	K	A	S	Y	L	D	C

P4	P3	P2	P1	P1'	P2'	P3'	P4'	P5'	P6'	P7'	P8'	P9'
P	L	G	E	V	V	G	E	D	P	M	A	E
P	S	R	S	F	Y	R	Q	N	G	E	P	Y
L	G	E	V	V	G	E	D	P	M	A	E	L
V	K	L	W	S	L	N	P	D	T	G	L	W
F	K	F	E	N	Q	R	R	N	K	R	E	D
K	F	E	N	Q	R	R	N	K	R	E	D	R
F	E	N	Q	R	R	N	K	R	E	D	R	T
Q	R	R	N	K	R	E	D	R	T	F	L	V
L	V	T	A	T	L	G	G	E	E	L	E	P
P	L	P	A	T	V	G	V	T	Q	P	Y	L
T	A	T	L	G	G	E	E	L	E	P	A	P
P	A	T	V	G	V	T	Q	P	Y	L	D	R
A	L	V	T	A	T	L	G	G	E	E	L	E
Q	V	L	G	D	Q	T	V	S	D	N	E	L
K	Y	V	N	K	E	I	Q	N	A	V	N	G
P	G	A	A	G	R	V	G	P	P	G	S	N
P	P	G	P	S	G	K	D	G	P	K	G	A
P	G	P	S	G	K	D	G	P	K	G	A	R
G	A	A	G	V	K	D	G	R	G	E	T	G
S	P	G	P	A	G	P	T	G	K	Q	G	D
P	G	P	Q	G	H	A	G	A	Q	G	P	P
P	G	I	N	G	S	P	G	G	K	G	E	M

P4	P3	P2	P1	P1 [†]	P2'	P3'	P4'	P5'	P6'	P7'	P8'	P9'
G	L	D	K	V	P	K	D	L	P	P	D	T
D	T	T	L	L	D	L	Q	N	K	I	T	E
D	V	Q	E	F	R	G	V	T	A	V	I	R
G	P	A	Y	A	Y	G	S	P	P	Q	P	E
K	T	P	F	V	T	H	P	G	Y	D	T	G
G	Q	Q	P	S	V	G	Q	Q	M	I	F	E
A	A	W	D	L	P	Q	Q	A	H	Q	P	T
P	L	E	E	S	Q	D	Q	C	K	C	E	N
T	E	S	E	S	R	G	A	I	Y	Dpa	K	K
V	A	F	H	D	F	S	S	D	L	E	N	V
N	V	P	H	L	Y	R	L	R	L	D	G	N
G	S	Q	C	S	V	D	L	E	S	A	S	G
S	F	Q	E	H	T	V	D	G	E	N	Q	I
D	L	E	S	A	S	G	E	K	D	L	A	P
E	A	A	E	A	R	R	G	Q	F	E	S	V

C = carboxymethylated cysteine; Mca = 7-hydroxy-4-methylcoumarin; Dnp = 2,4-dinitrophenyl; Hyp = 4-hydroxyproline; Dpa = *N*-3-[2,4-dinitrophenyl]-2,3-diaminopropionyl.

[†]The cleavage site for each peptide sequence is located between the two residues coloured in red.

Please note that the order of the substrates in this table is purely arbitrary, as cleavage rates have not been reported for the majority of them.

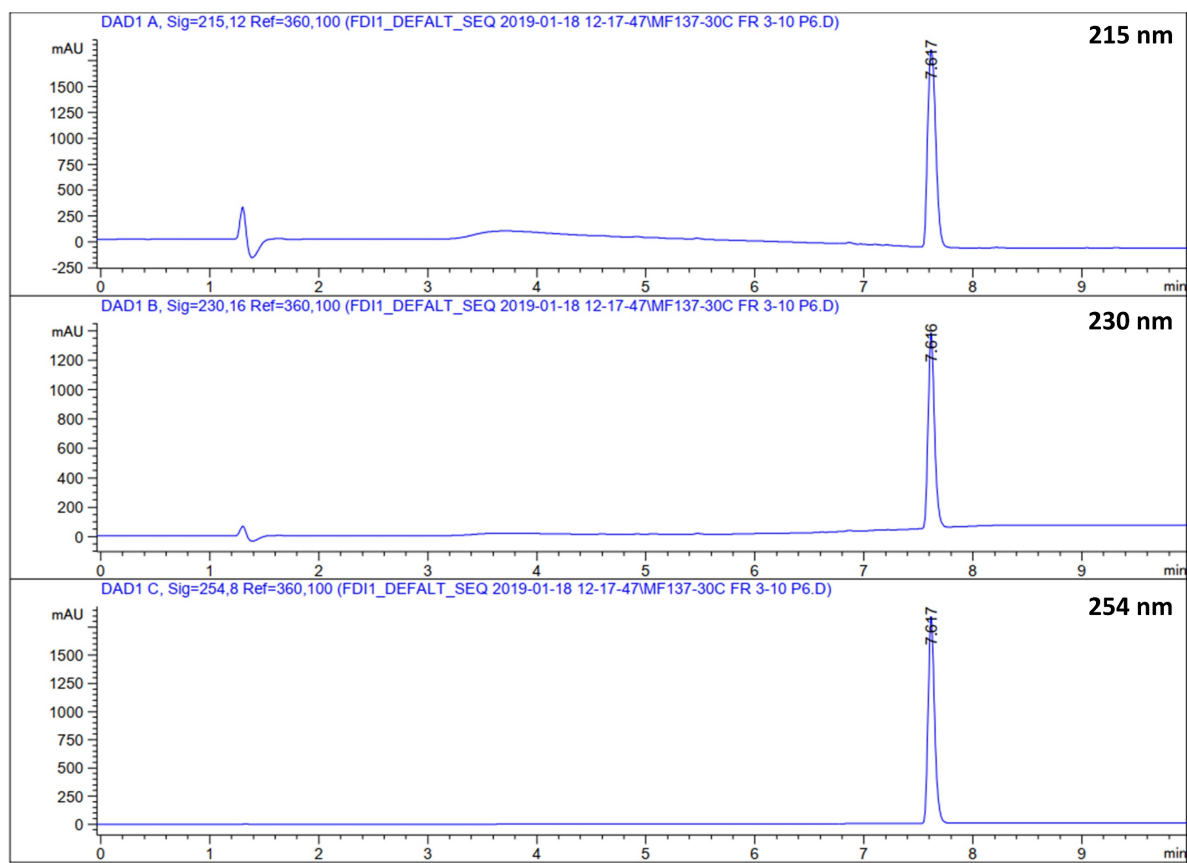


Figure 8.11. HPLC Chromatograms of the small molecule fluorophore building block Fmoc-Lys(*retro*-Abz-Boc)-OH (22). The analytical sample was run at three different wavelengths: 215 nm, 230 nm and 254 nm.

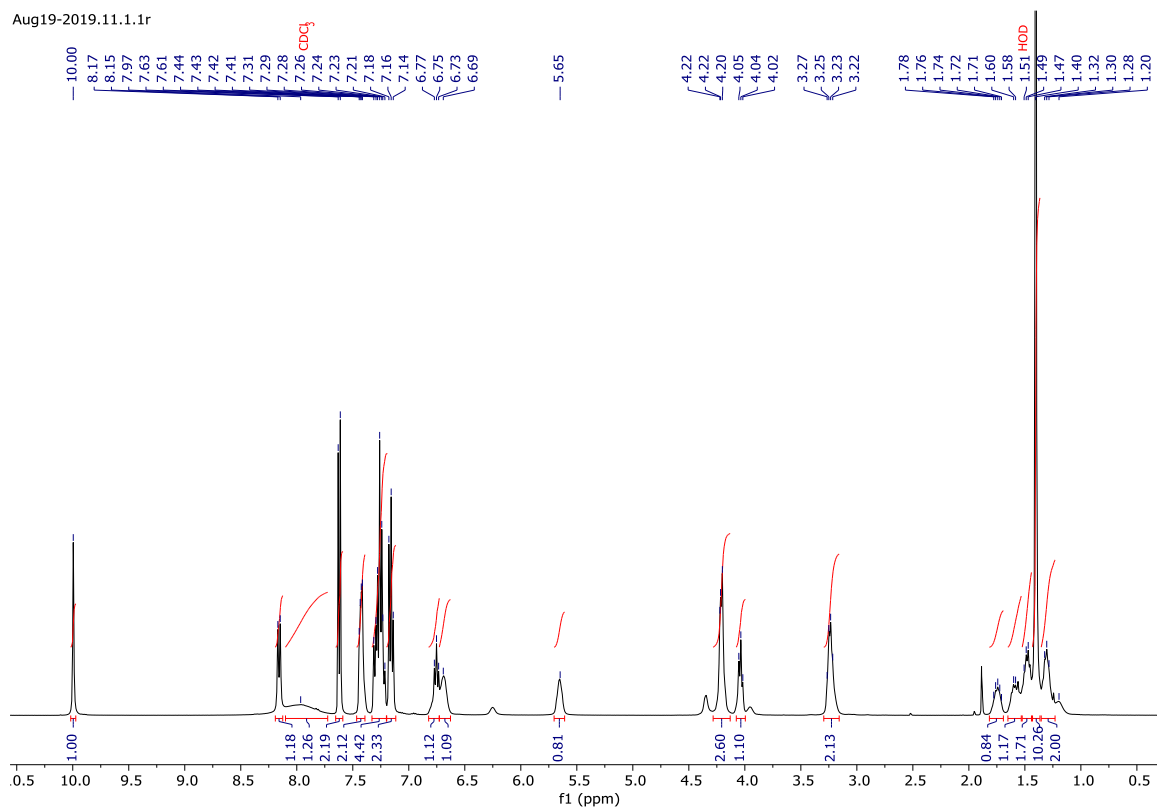


Figure 8.12. ^1H -NMR spectrum of the small molecule fluorophore building block Fmoc-Lys(*retro*-Abz-Boc)-OH (22).

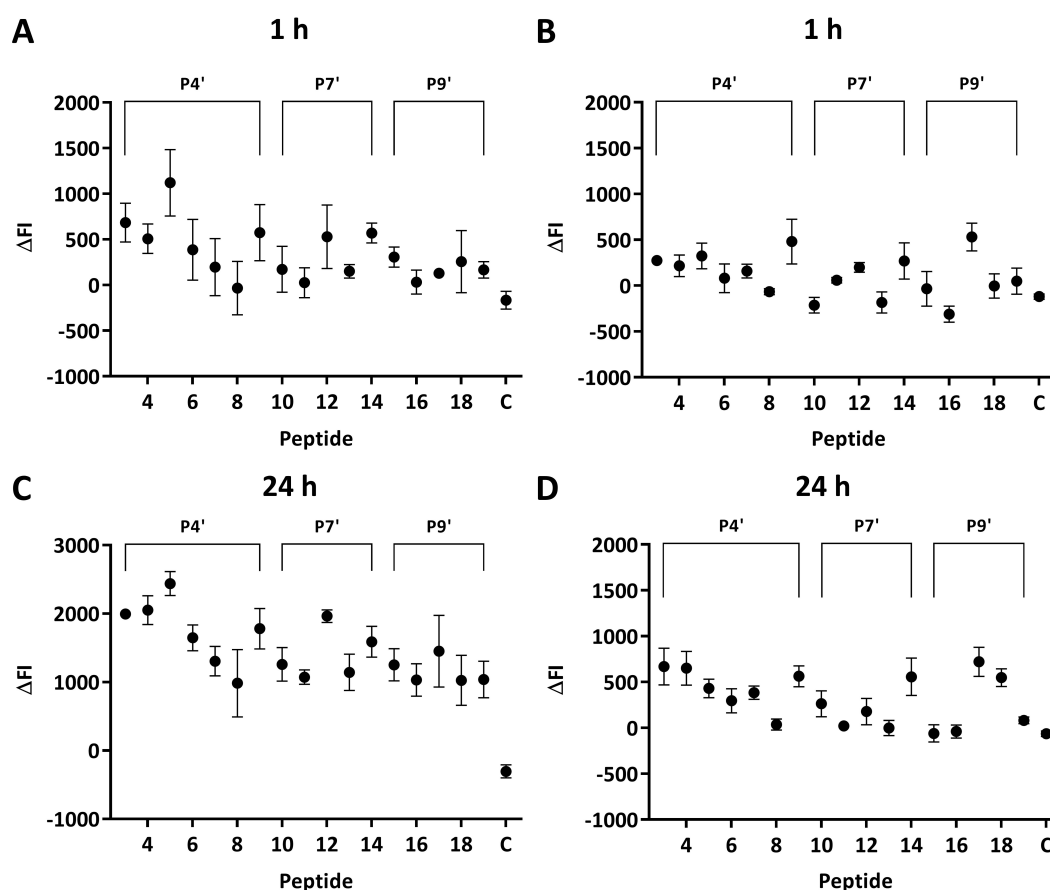


Figure 8.13. Change in fluorescence intensity at different time points after incubation of on-resin FRET peptide substrates in the presence or absence of ADAMTS-5. Individual changes in fluorescence intensity (ΔFI) for P4', P7' and P9' peptide substrates on-resin after 1 h (A) and 24 h (C) in the presence of TS5-5 (1 μM); and after 1 h (B) and 24 h (D) in the absence of TS5-5. The change in fluorescence intensity (ΔFI) is expressed as background-subtracted FI at time t – background-subtracted FI at time 0. For each peptide substrate, ΔFI is expressed as the mean $\pm$ SEM for $n = 3$ wells from three independent experiments performed in single replicate at 37 $^{\circ}C$. In all cases, cleavage of peptide substrates was monitored at $\lambda_{ex} = 350$ nm and $\lambda_{em} = 438$ nm in TNC buffer with 0.005% (v/v) Brij-35. C = control, un-modified PEGA₁₉₀₀ resin with no substrate. *Please note that the variation in ΔFI for control PEGA₁₉₀₀ beads reflects slight differences in the fluorescence intensity readings of different beads within the same well.*

8.3. Chapter 5 Appendix

8.3.1. MALDI and tandem mass spectra of peptide hits

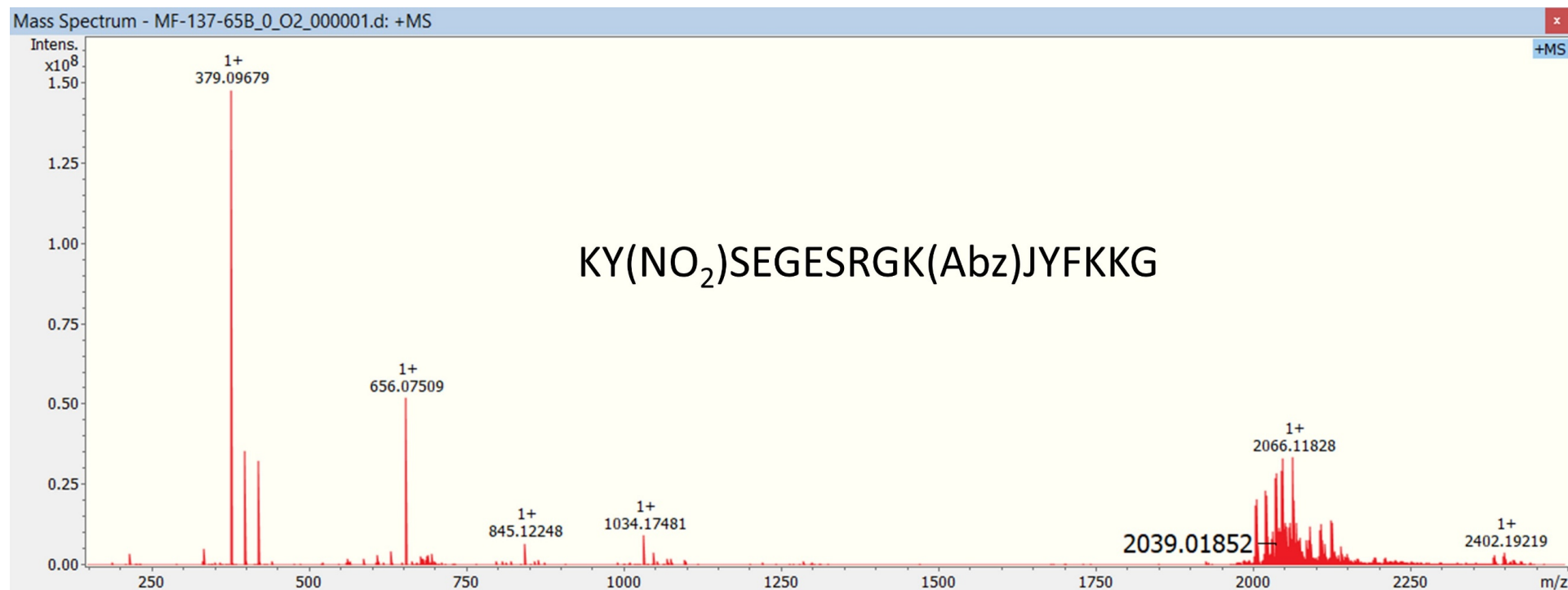


Figure 8.14. Mass spectrum of the peptide hit KY(NO₂)SEGESRGK(Abz)JYFKKG (31) obtained from ADAMTS-5 library screening. MALDI spectrum of the peptide hit KY(NO₂)SEGESRGK(Abz)JYFKKG, with the mass of the quasi-molecular ion [M + H]⁺ (2039.01852 u) indicated in large text.

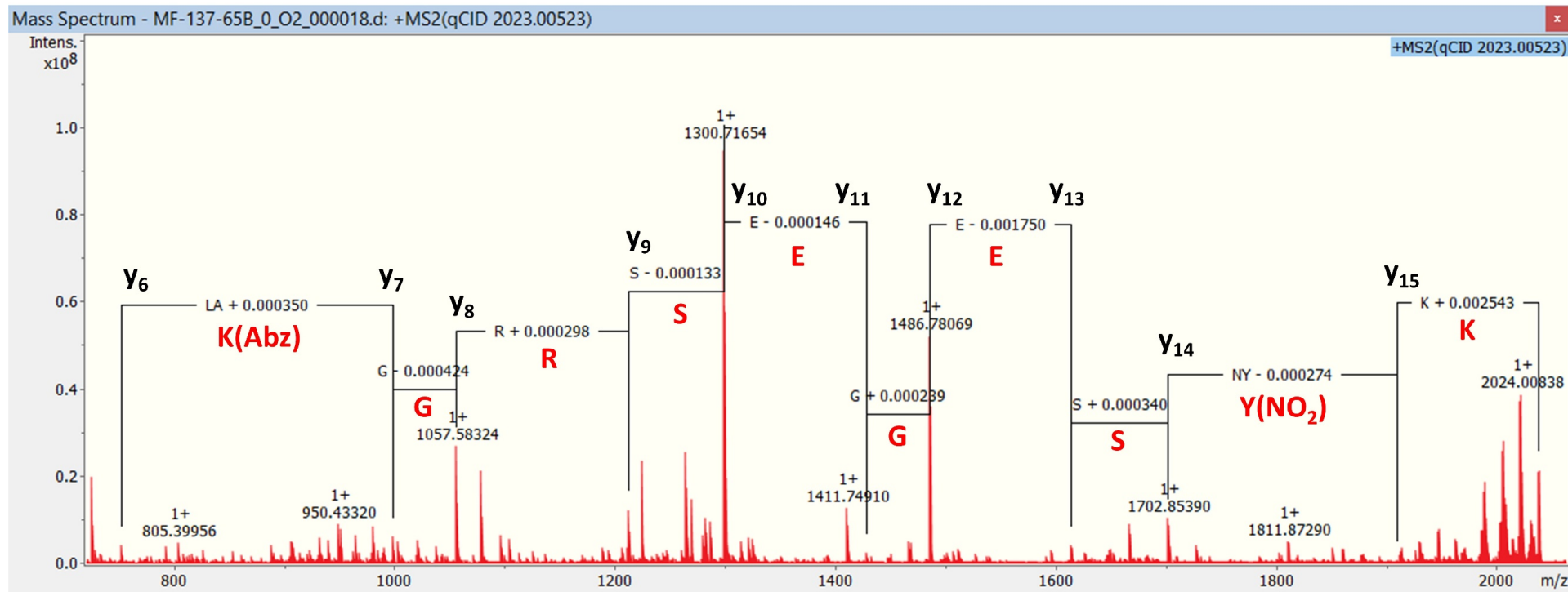


Figure 8.15. Tandem mass spectrum of the peptide hit KY(NO₂)SEGESRGK(Abz)JYFKKG (31) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of $[M - O + H]^+$ (2023.00523 u), where M = mass of the quasi-molecular ion. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

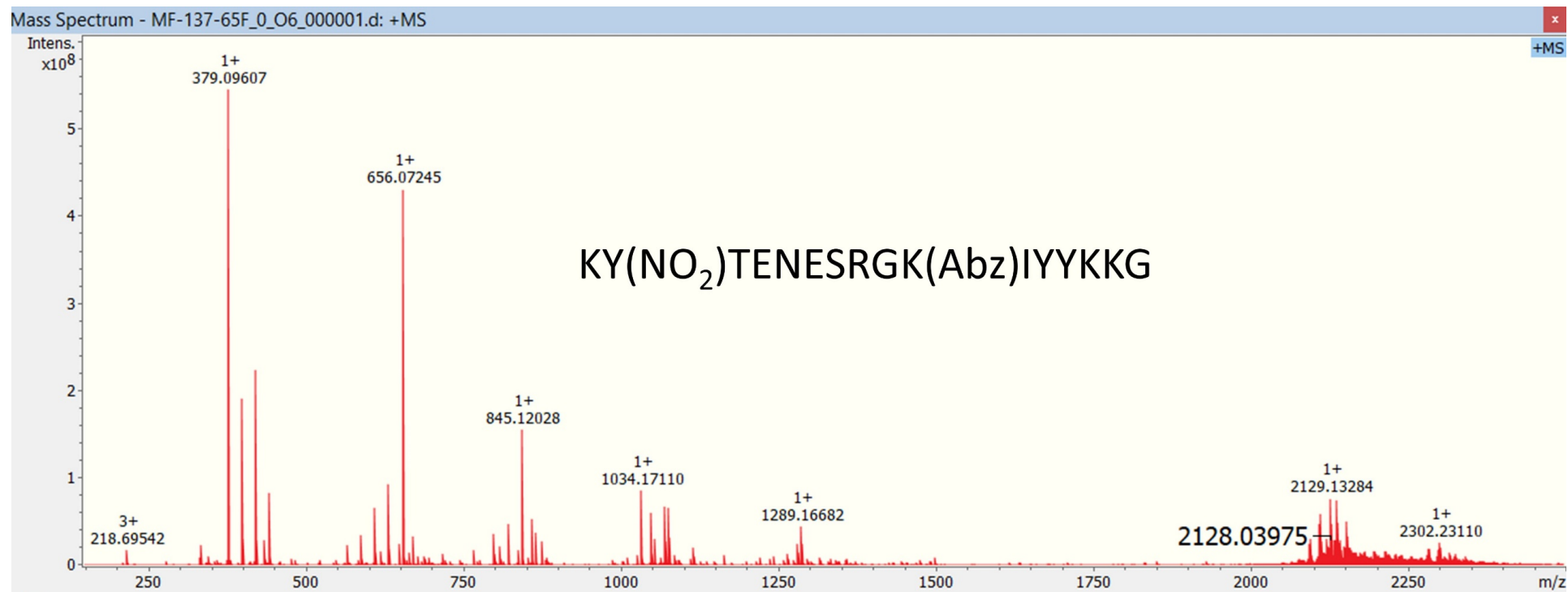


Figure 8.16. Mass spectrum of the peptide hit KY(NO₂)TENESRGK(Abz)IYYKKG (32) obtained from ADAMTS-5 library screening. Initial MALDI spectrum of the peptide hit KY(NO₂)TENESRGK(Abz)IYYKKG, with the mass of the quasi-molecular ion [M + H]⁺ (2128.03975 u) indicated in large text.

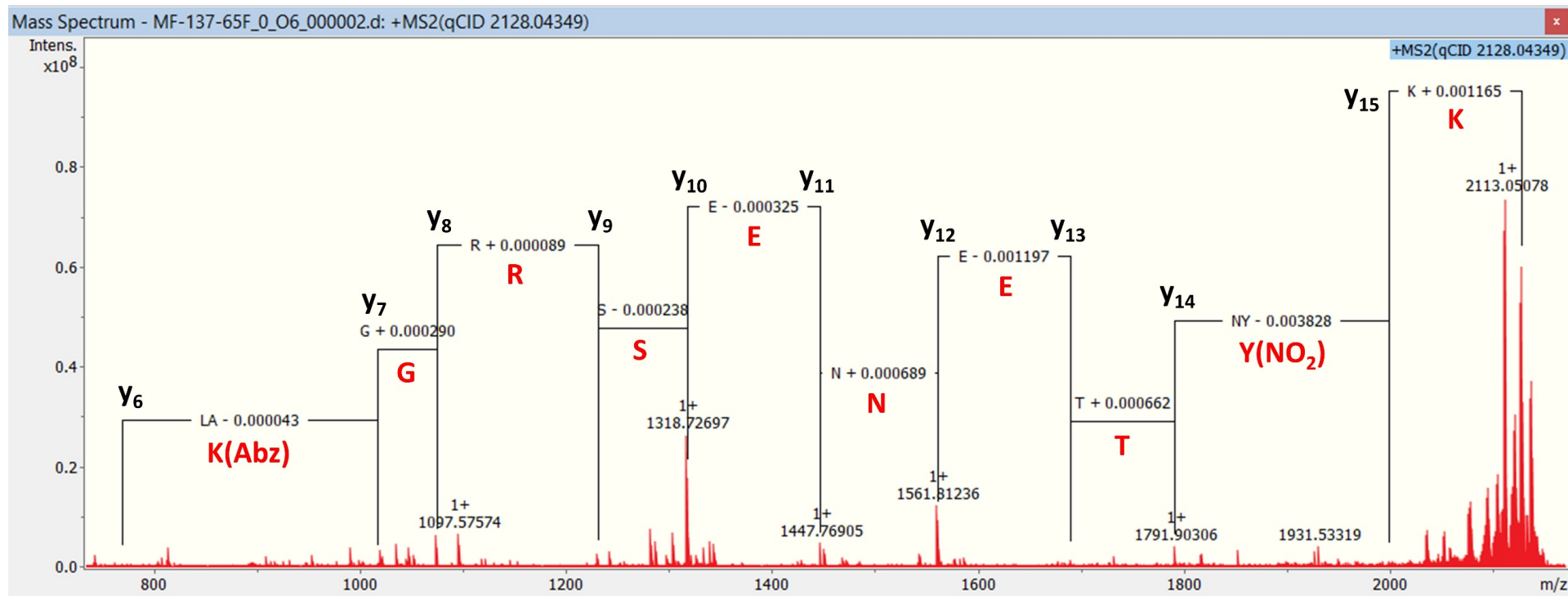


Figure 8.17. Tandem mass spectrum of the peptide hit KY(NO₂)TENESRGK(Abz)IYYKKG (32) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of the quasi-molecular ion $[M + H]^+$ (2128.04349 u), where M = mass of the quasi-molecular ion. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

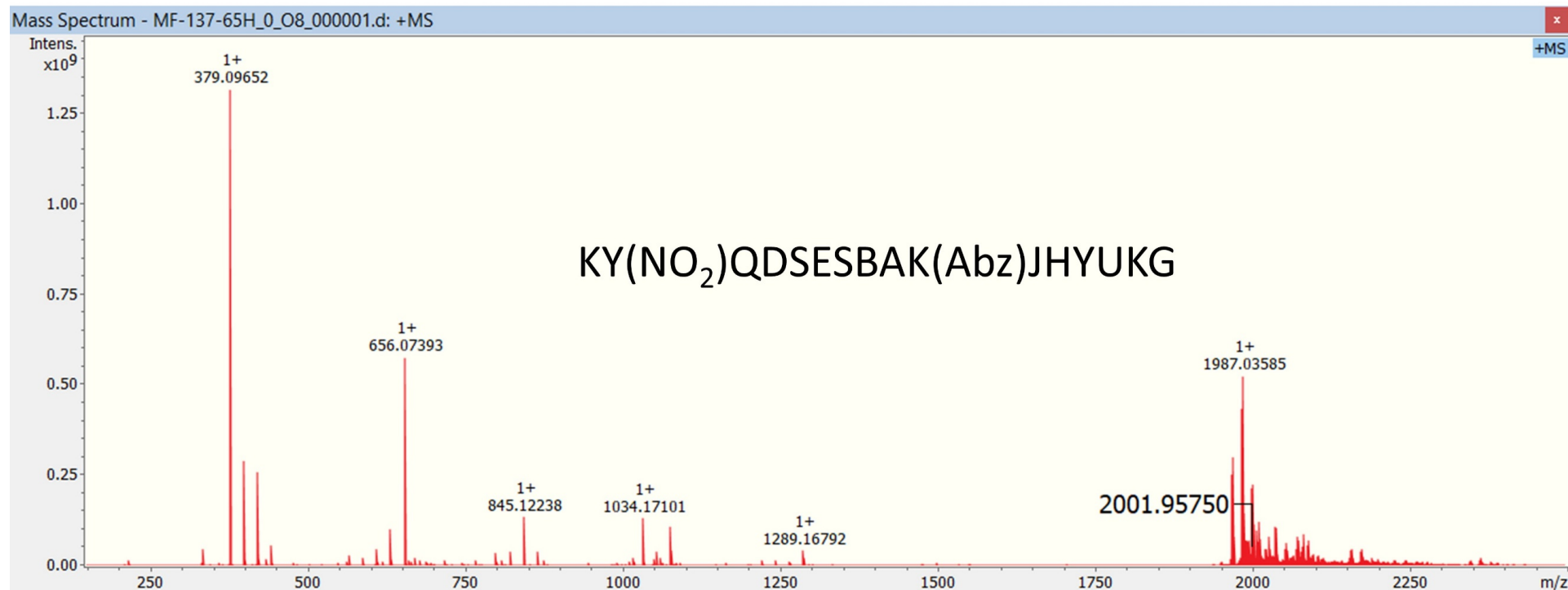


Figure 8.18. Mass spectrum of the peptide hit KY(NO₂)QDSESBK(Abz)JHYUKG (33) obtained from ADAMTS-5 library screening. Initial MALDI spectrum of the peptide hit KY(NO₂)QDSESBK(Abz)JHYUKG, with the mass of the quasi-molecular ion $[M + H]^+$ (2001.95750 u) indicated in large text.

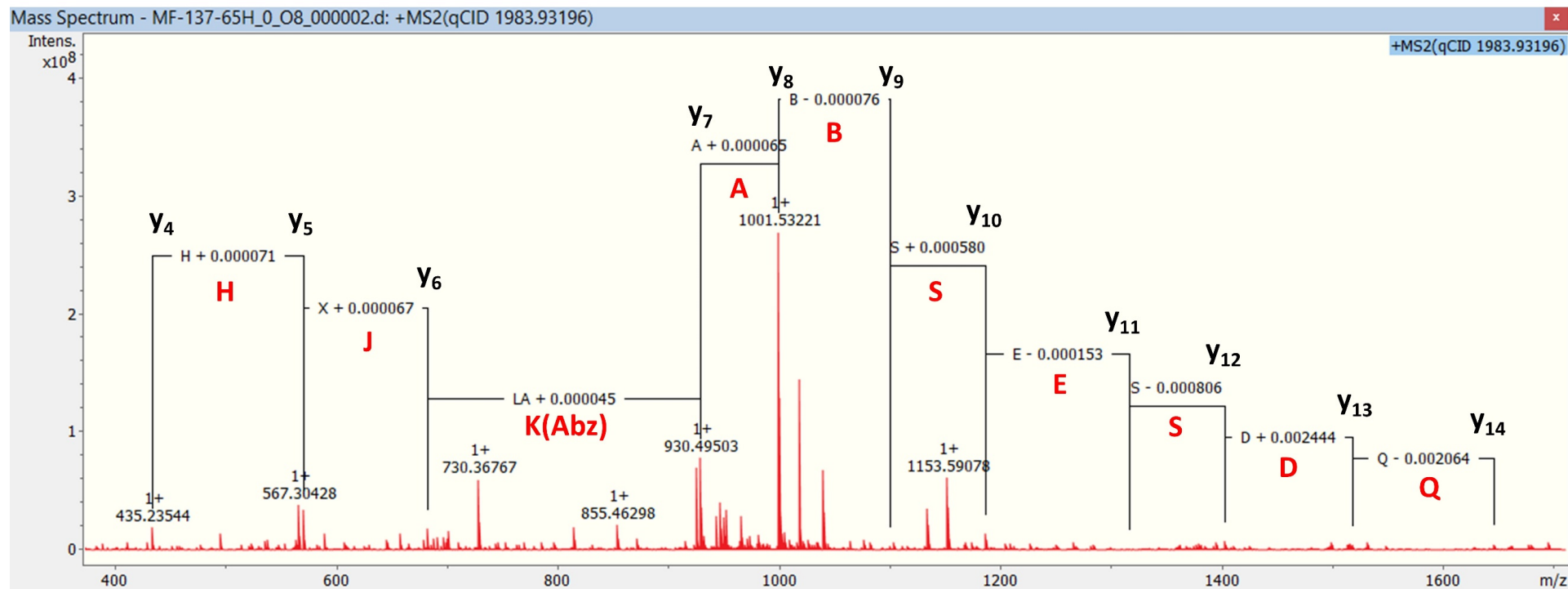


Figure 8.19. Tandem mass spectrum of the peptide hit KY(NO₂)QDESBAK(Abz)JHYUKG (33) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of $[M - OH_2 + H]^+$ (1983.93196 u), where M = mass of the quasi-molecular ion. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

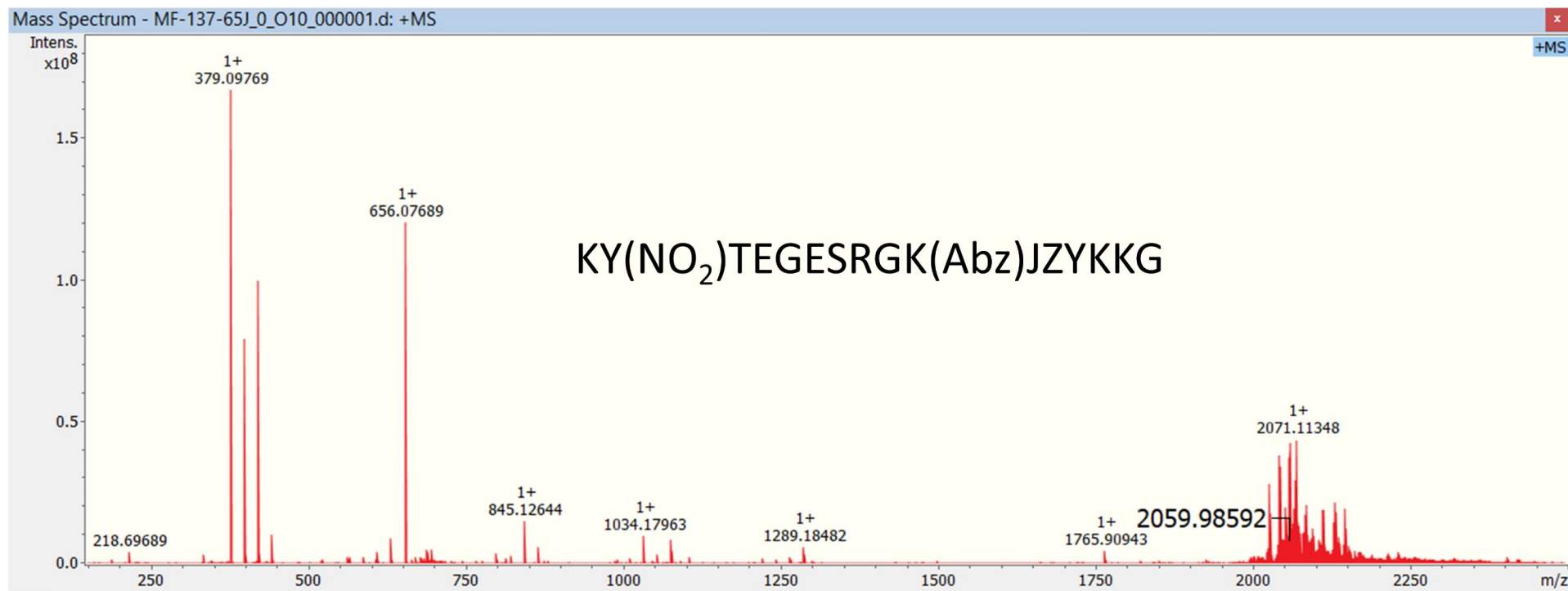


Figure 8.20. Mass spectrum of the peptide hit KY(NO₂)TEGESRGK(Abz)JZYKKG (34) obtained from ADAMTS-5 library screening. Initial MALDI spectrum of the peptide hit KY(NO₂)TEGESRGK(Abz)JZYKKG, with the mass of the quasi-molecular ion [M + H]⁺ (2059.98592 u) indicated in large text.

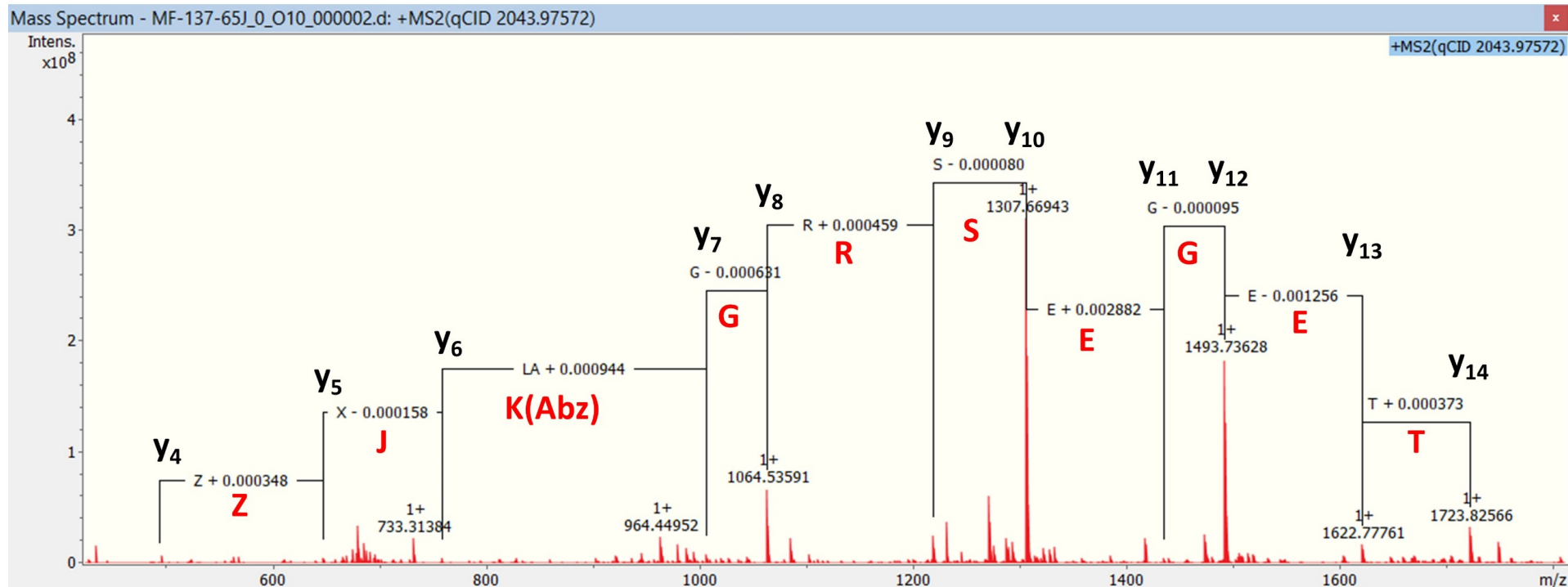


Figure 8.21. Tandem mass spectrum of the peptide hit KY(NO₂)TEGESRGK(Abz)JZYKKG (34) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of $[M - O + H]^+$ (2043.97572 u), where M = mass of the quasi-molecular ion. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

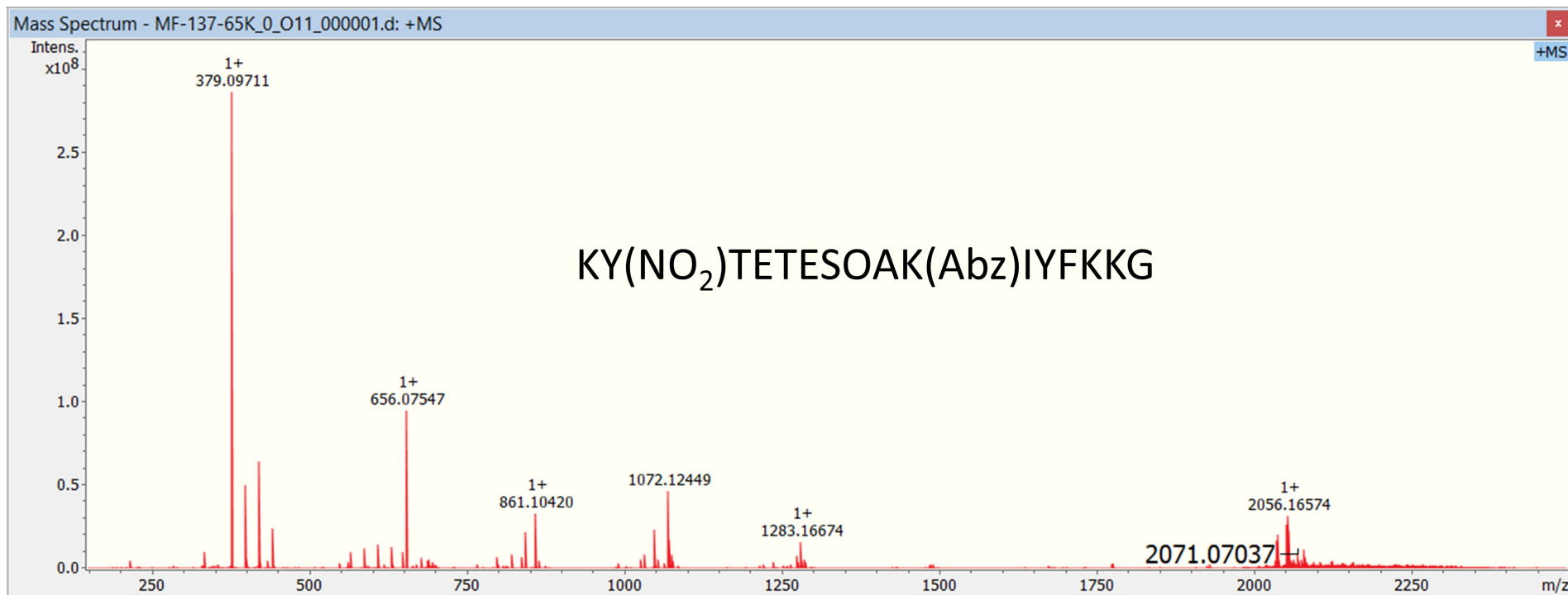


Figure 8.22. Mass spectrum of the peptide hit KY(NO₂)TETESOAK(Abz)IYFKKG (35) obtained from ADAMTS-5 library screening. Initial MALDI spectrum of the peptide hit KY(NO₂)TETESOAK(Abz)IYFKKG, with the mass of the quasi-molecular ion $[M + H]^+$ (2071.07037 u) indicated in large text.

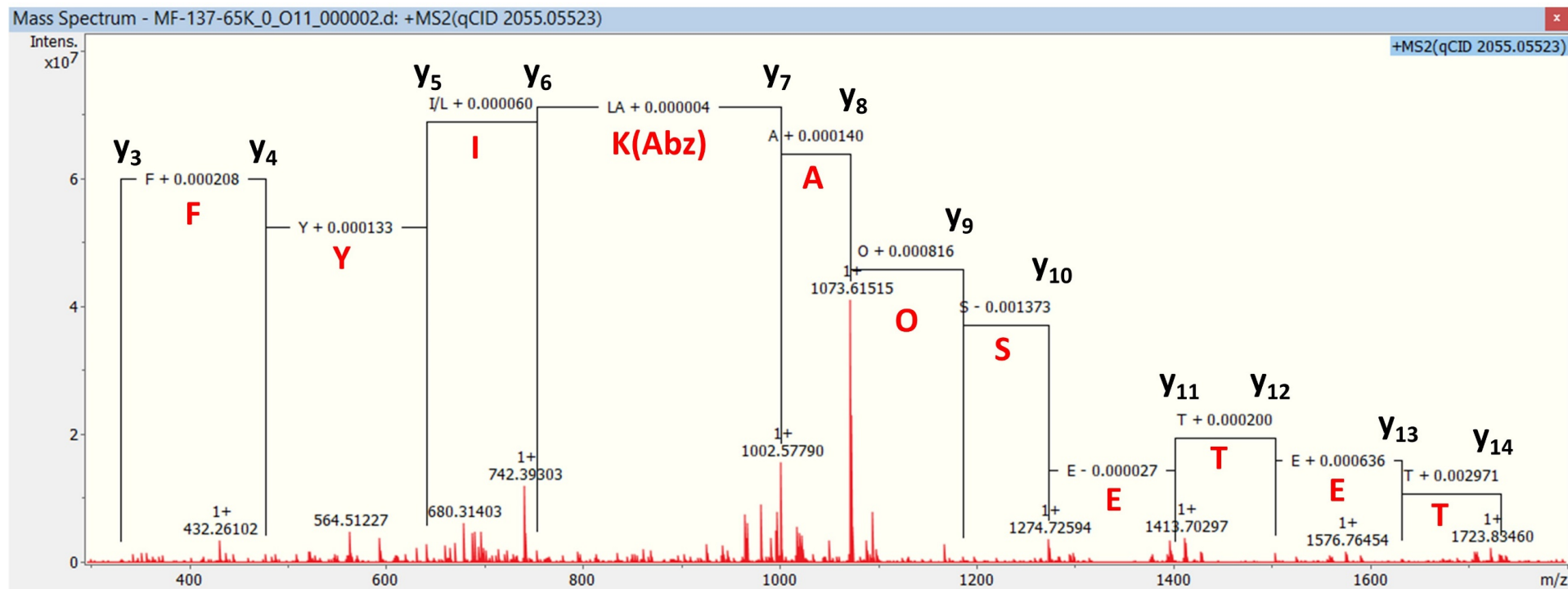


Figure 8.23. Tandem mass spectrum of the peptide hit KY(NO₂)TETESOAK(Abz)IYFKKG (35) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of $[M - O + H]^+$ (2055.05523 u), where M = mass of the quasi-molecular ion. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

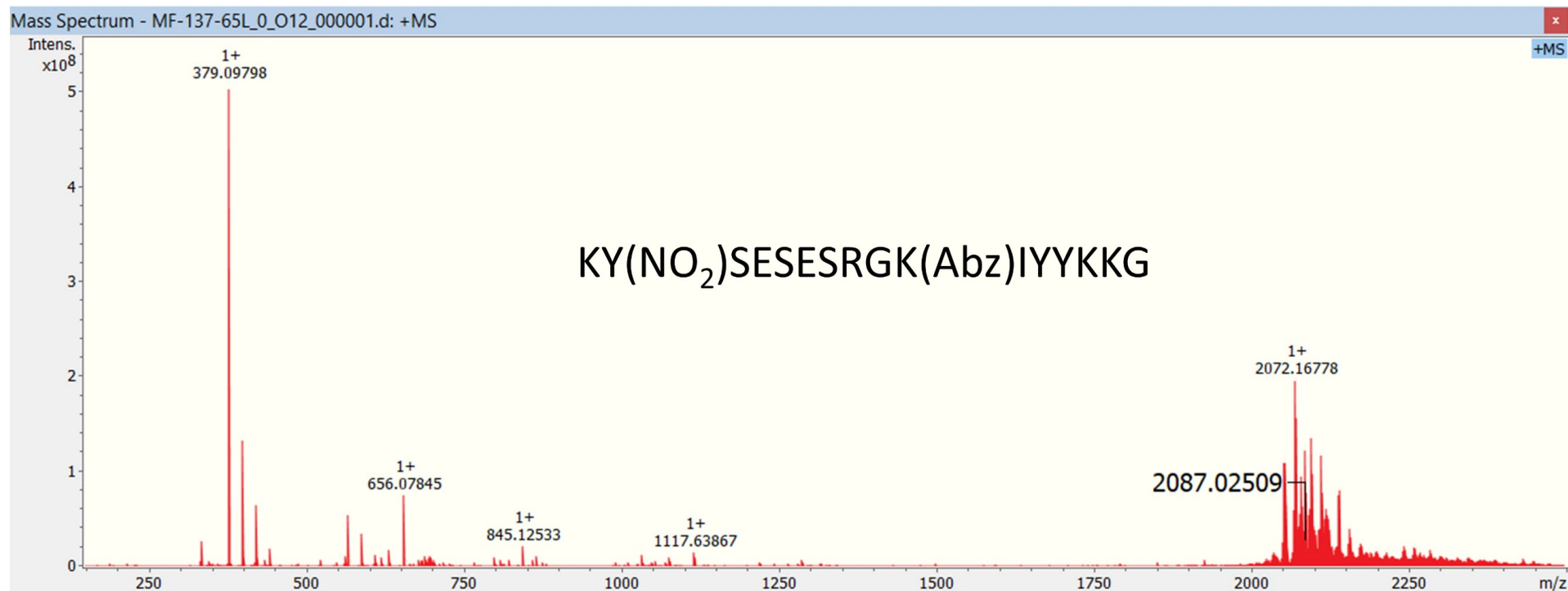


Figure 8.24. Mass spectrum of the peptide hit KY(NO₂)SESESRGK(Abz)IYYKKG (36) obtained from ADAMTS-5 library screening. Initial MALDI spectrum of the peptide hit KY(NO₂)SESESRGK(Abz)IYYKKG, with the mass of the quasi-molecular ion $[M + H]^+$ (2087.02509 u) indicated in large text.

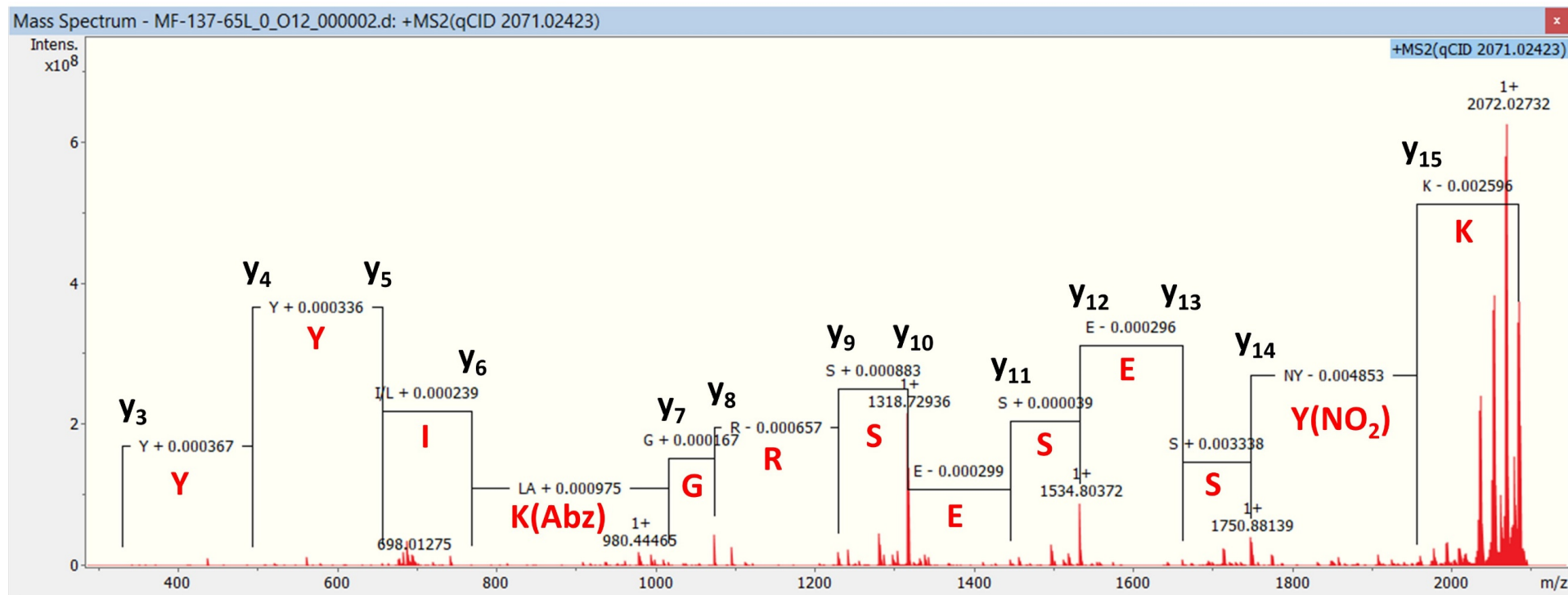


Figure 8.25. Tandem mass spectrum of the peptide hit KY(NO₂)SESESRGK(Abz)IYYKKG (36) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of $[M - O + H]^+$ (2071.02423 u), where M = mass of the quasi-molecular ion. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

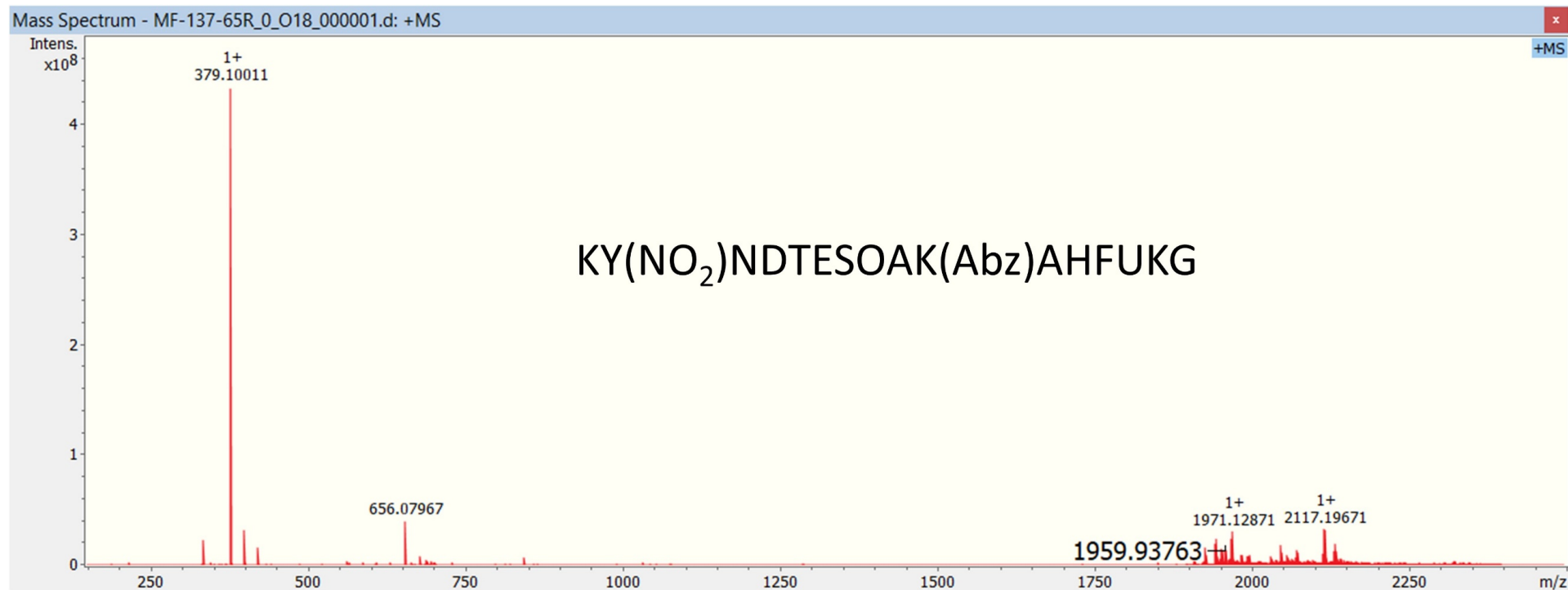


Figure 8.26. Mass spectrum of the peptide hit KY(NO₂)NDTESOAK(Abz)AHFUKG (37) obtained from ADAMTS-5 library screening. Initial MALDI spectrum of the peptide hit KY(NO₂)NDTESOAK(Abz)AHFUKG, with the mass of the quasi-molecular ion [M + H]⁺ (1959.93763 u) indicated in large text.

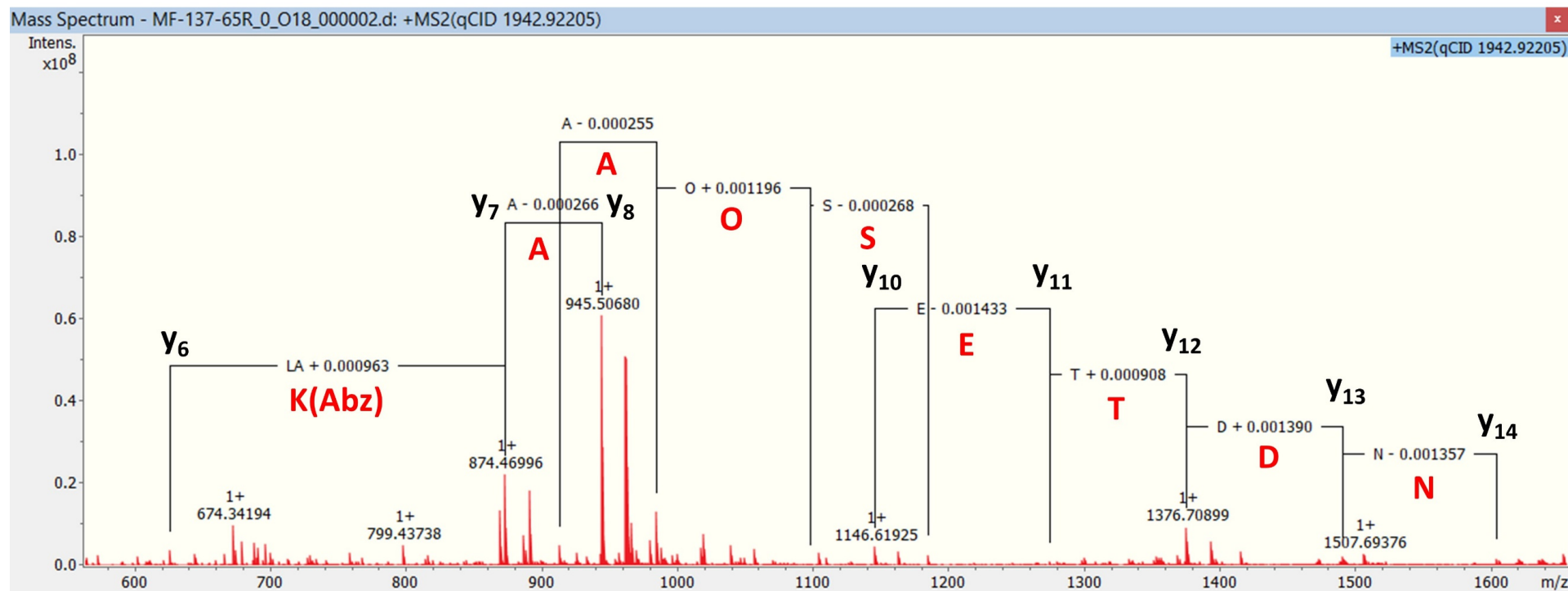


Figure 8.27. Tandem mass spectrum of the peptide hit KY(NO₂)NDTESOAK(Abz)AHFUKG (37) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of $[M - OH + H]^+(1942.92205 \text{ u})$, where M = mass of the quasi-molecular ion. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

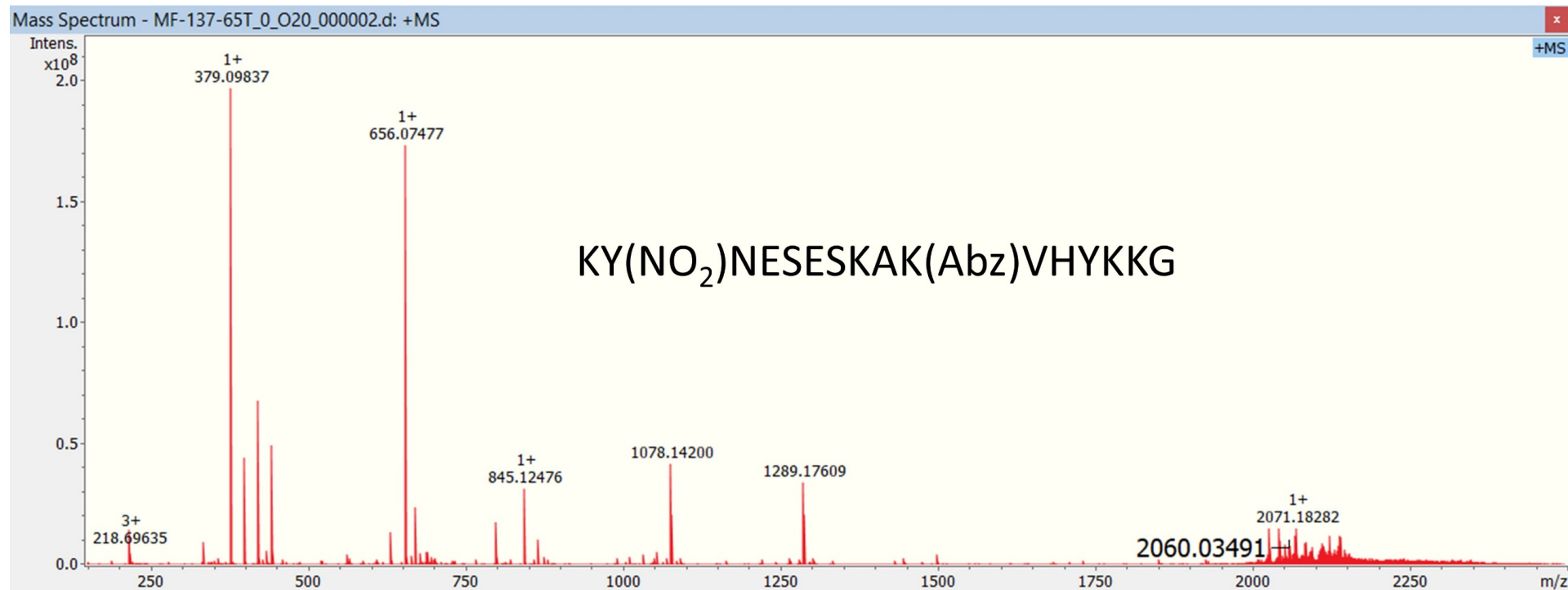


Figure 8.28. Mass spectrum of the peptide hit KY(NO₂)NESEKAK(Abz)VHYKKG (38) obtained from ADAMTS-5 library screening. Initial MALDI spectrum of the peptide hit KY(NO₂)NESEKAK(Abz)VHYKKG, with the mass of the quasi-molecular ion $[M + H]^+$ (2060.03491 u) indicated in large text.

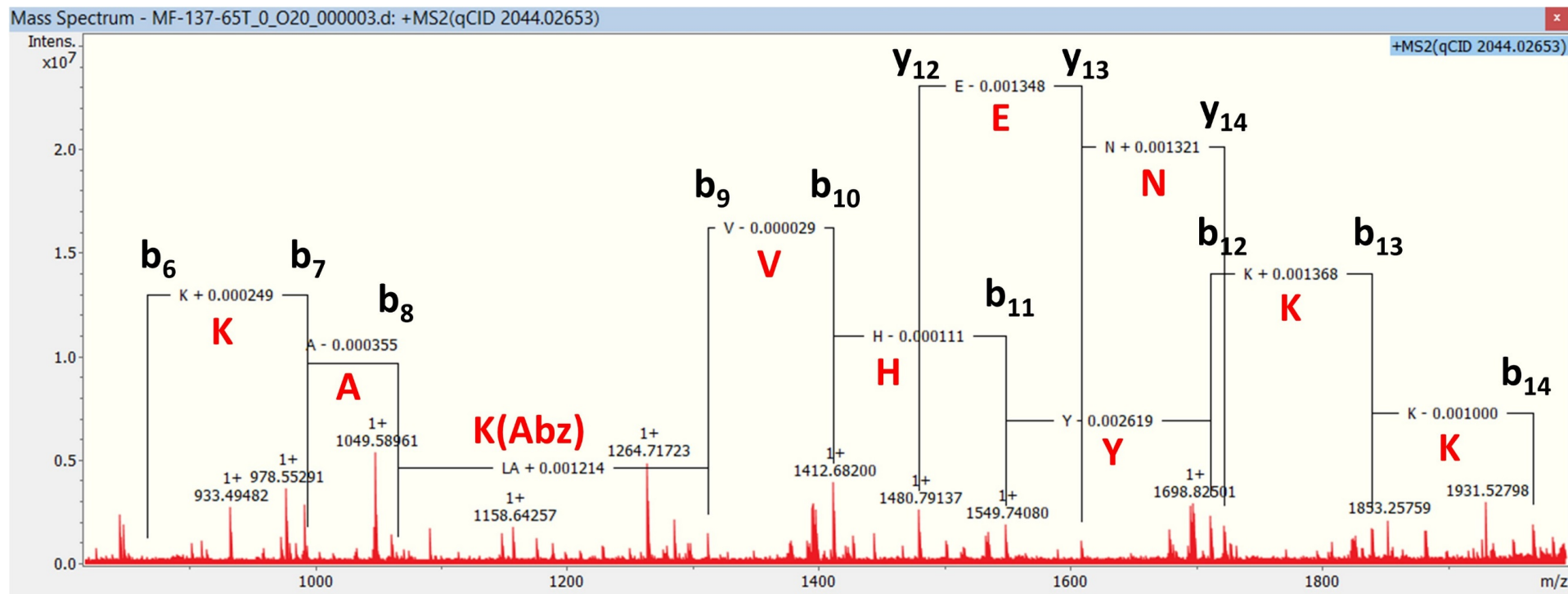


Figure 8.29. Tandem mass spectrum of the peptide hit KY(NO₂)NESESKAK(Abz)VHYKKG (38) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of $[M - O + H]^+$ (2044.02653 u), where M = mass of the quasi-molecular ion. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

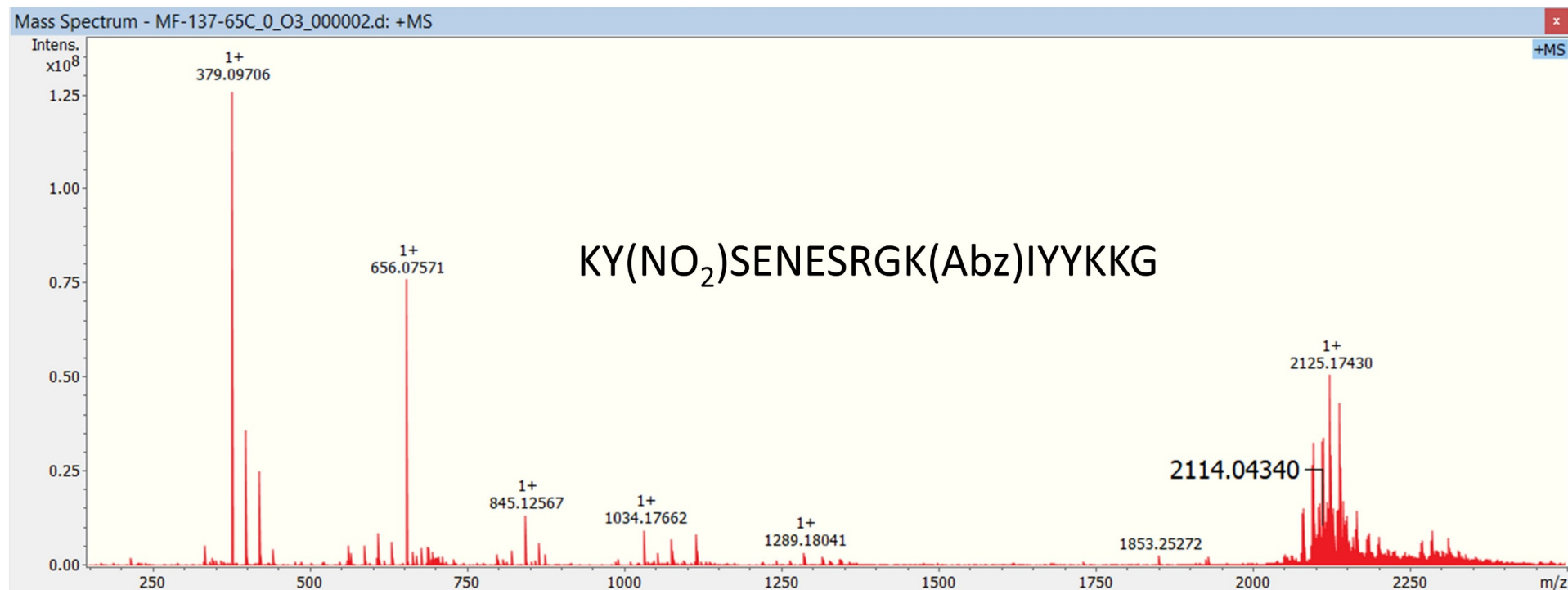


Figure 8.30. Mass spectrum of the peptide hit KY(NO₂)SENERGK(Abz)IYYKKG (39) obtained from ADAMTS-5 library screening. Initial MALDI spectrum of the peptide hit KY(NO₂)SENERGK(Abz)IYYKKG, with the mass of an unknown quasi-molecular ion $[M + H]^+$ (2114.04340 u) indicated in large text.

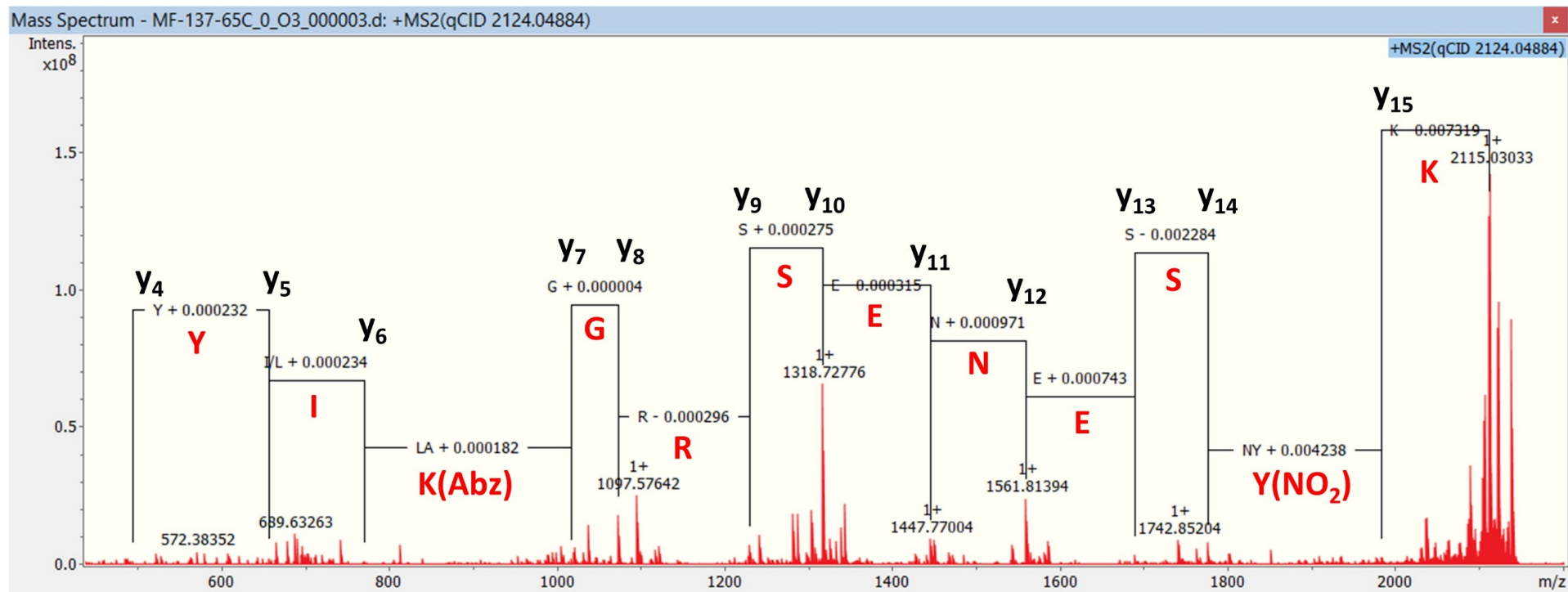
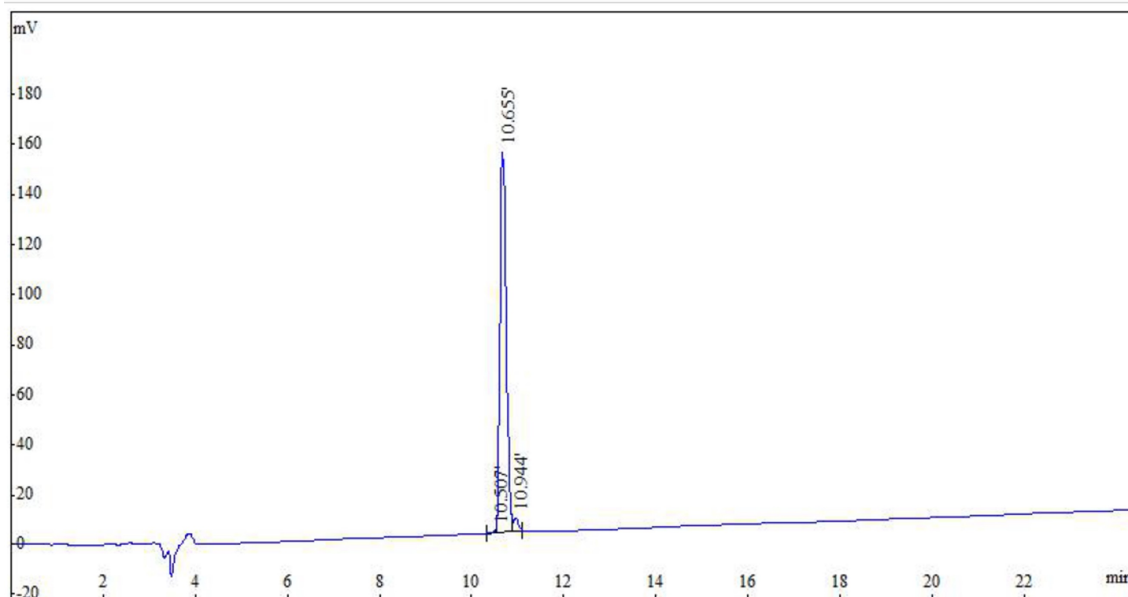


Figure 8.31. Tandem mass spectrum of the peptide hit KY(NO₂)SENE SRGK(Abz)IYYKKG (39) obtained from ADAMTS-5 library screening. Tandem mass spectrum following fragmentation of $[M - O + H]^+$ (2124.04884 u), which is related to the quasi-molecular ion 2140.04590 u that could not be identified. Instead, the fragmentation pattern fit the peptide with quasi-molecular ion mass 2114.04340 u, which is shown in Figure 8.27. For clarity, only selected residue losses are shown and only fragment ion peaks that could be identified as b ions or y ions have been labelled. Spectra were analysed and processed using Bruker Compass DataAnalysis software.

8.3.2. HPLC characterisation data of peptide hits from GL Biochem

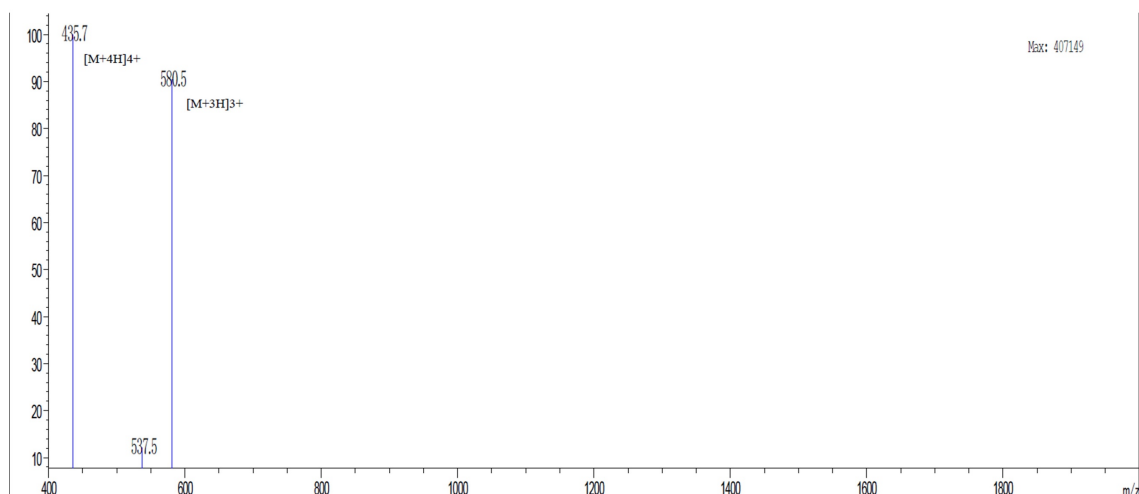
Abz-TESESRGAIY-Dpa-KK-NH₂ (2)

UV trace



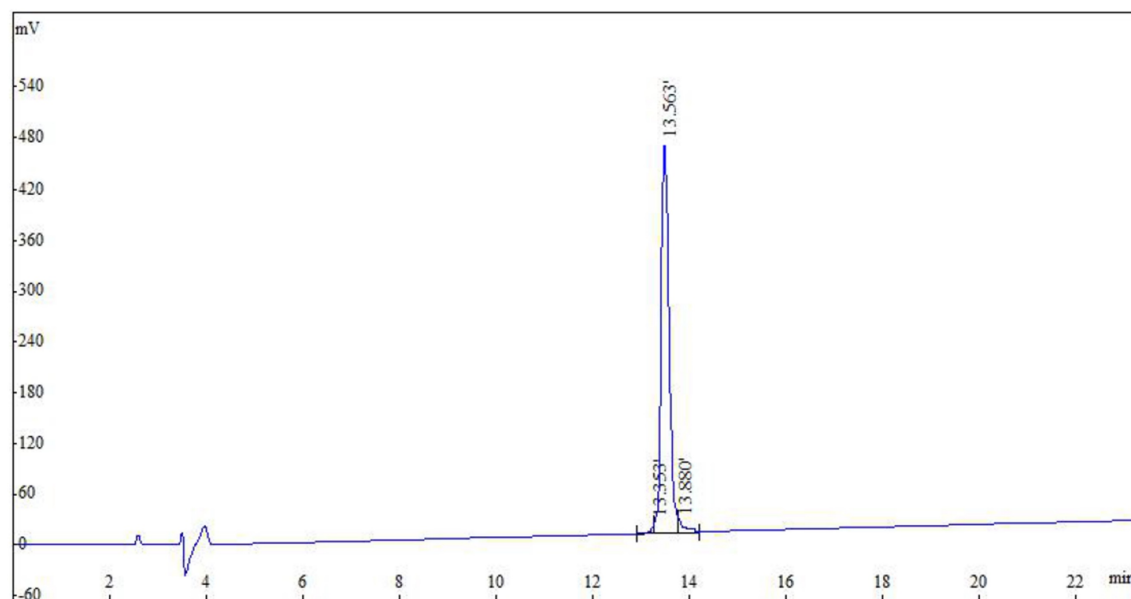
Rank	Time	Conc.	Area	Height
1	10.507	0.6450	8528	1007
2	10.655	96.8862	1281004	151430
3	10.944	2.4688	32642	4835
Total		100	1322174	157272

Mass trace



KY(NO₂)TESESRGK(Abz)IYYKKG (5)

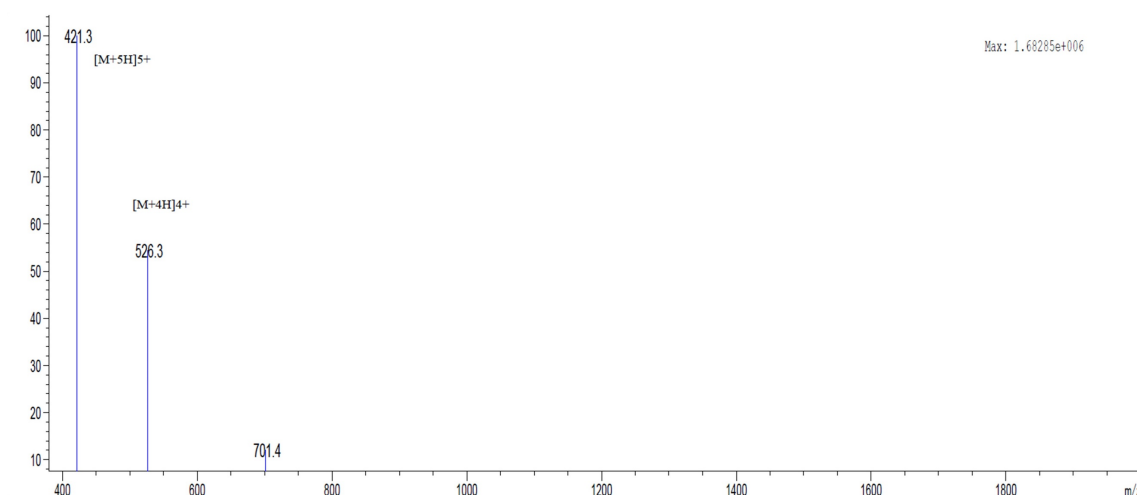
UV trace



Rank	Time	Conc.	Area	Height
1	13.353	1.1360	55943	11168
2	13.563	96.3088	4742794	457251
3	13.880	2.5552	125833	12116

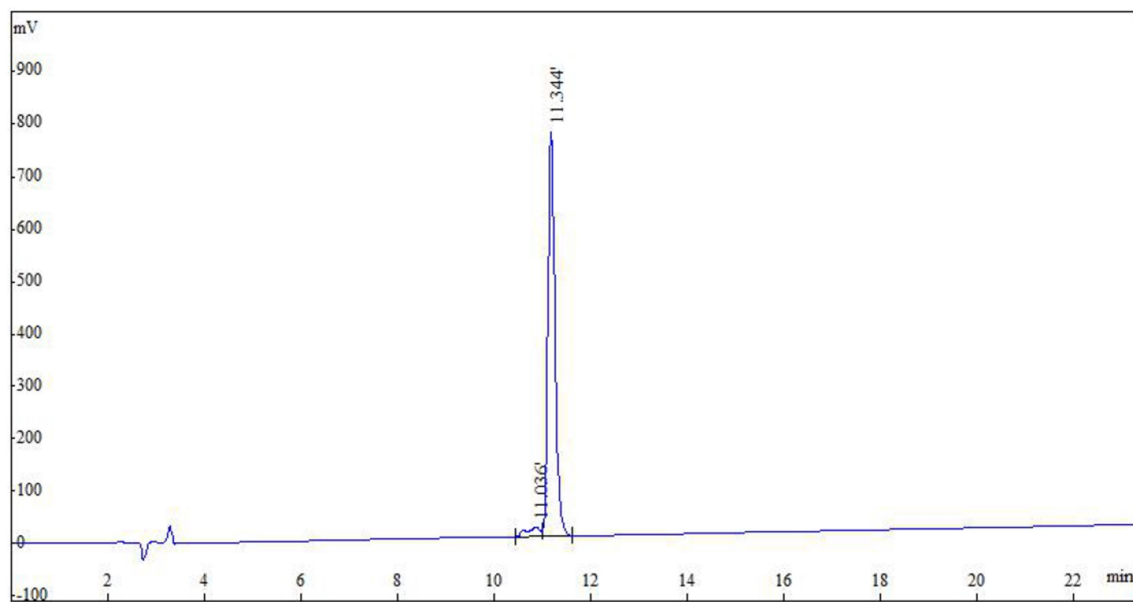
Total 100 4924570 480535

Mass trace



KY(NO₂)SEGESRGK(Abz)JYFKKG (31)

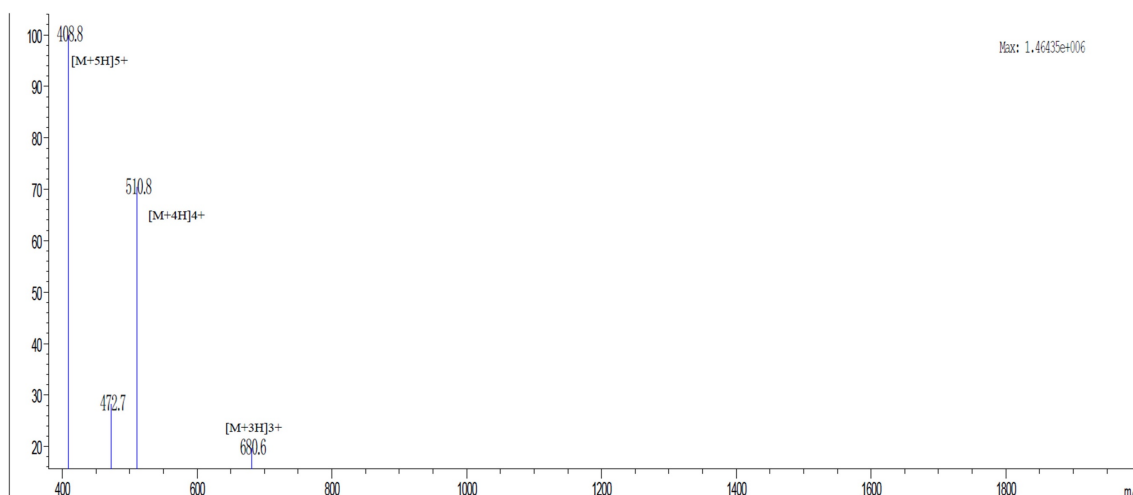
UV trace



Rank	Time	Conc.	Area	Height
1	11.036	4.0557	306300	17721
2	11.344	95.9443	7245970	766772

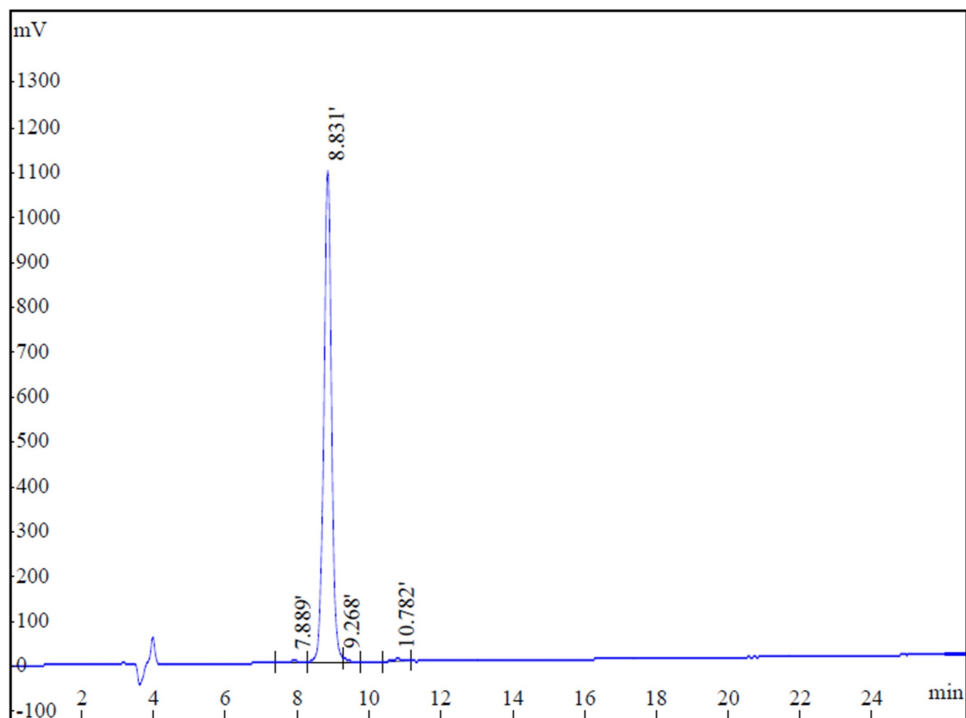
Total 100 7552270 784493

Mass trace



KY(NO₂)TENESRGK(Abz)IYYKKG (32)

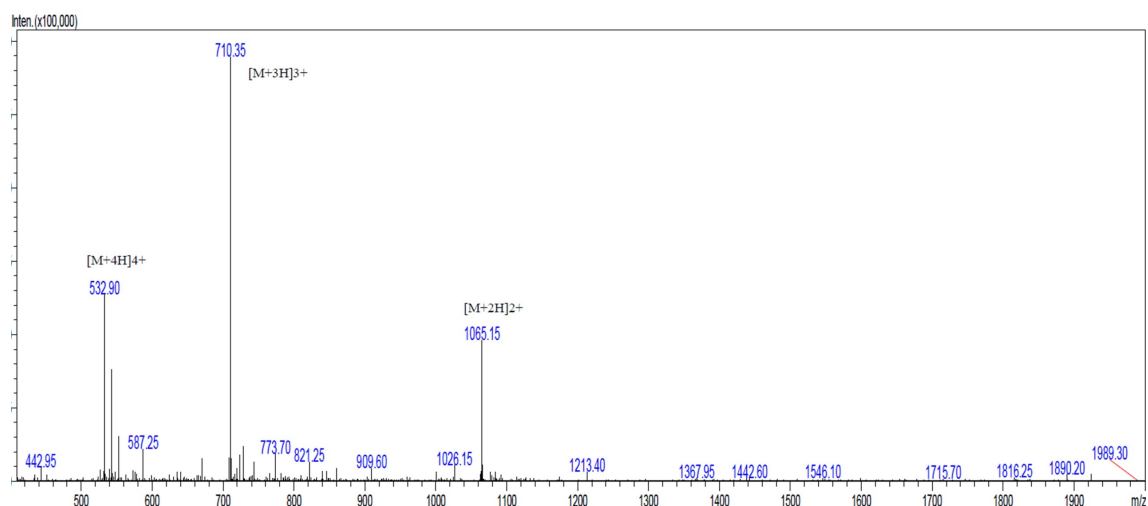
UV trace



Rank	Time	Conc.	Area	Height
1	7.889	0.4987	75813	4937
2	8.831	98.57	14983773	1090223
3	9.268	0.392	59595	8713
4	10.782	0.5434	82603	5358

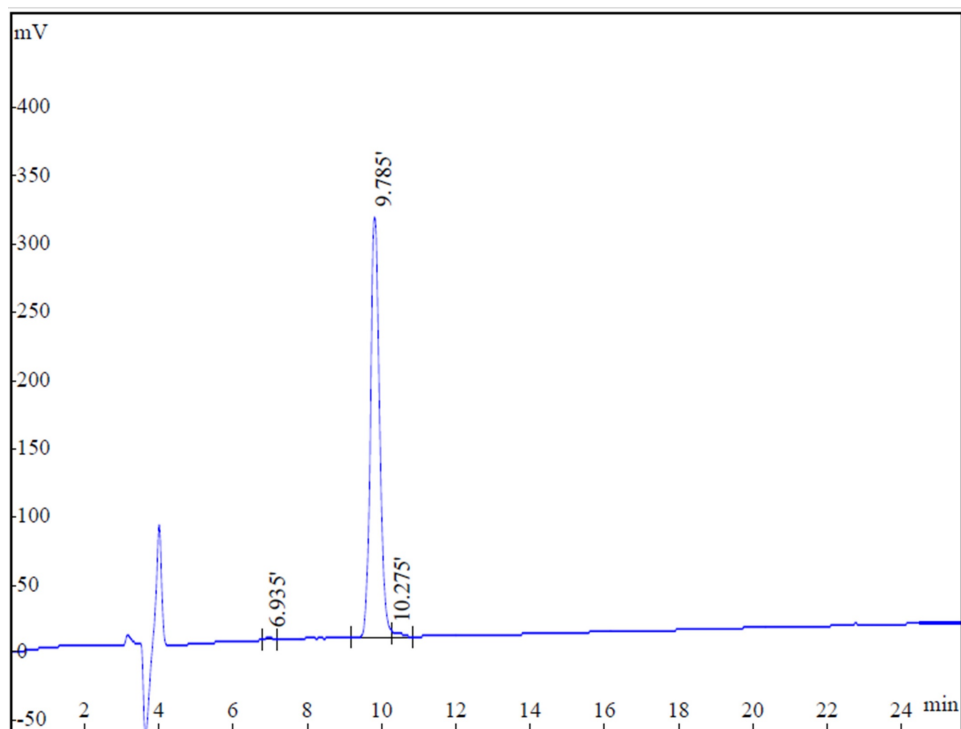
Total 100.0000

Mass trace



KY(NO₂)QDSESAK(Abz)JHYUKG (33)

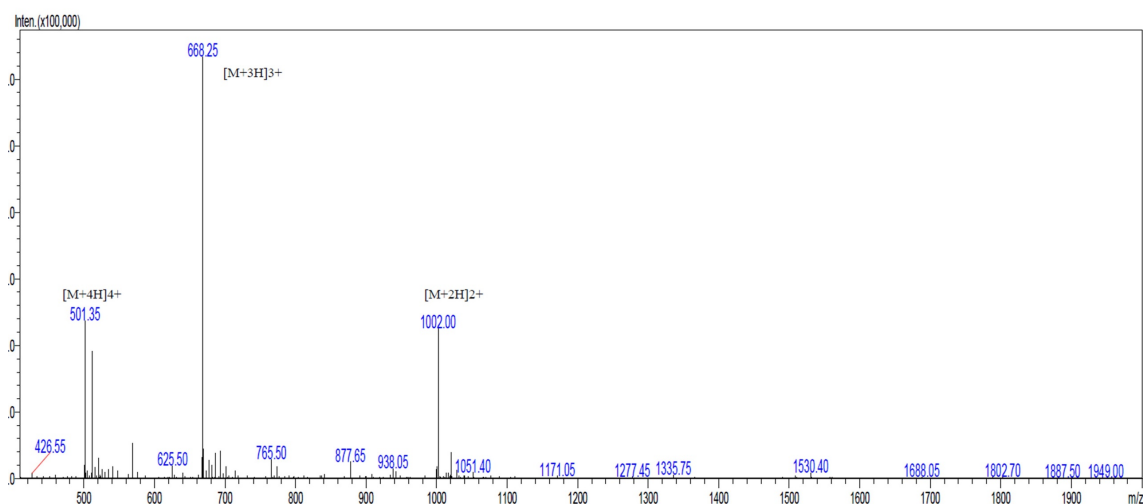
UV trace



Rank	Time	Conc.	Area	Height
1	6.935	0.4779	23701	2445
2	9.785	98.66	4892636	307866
3	10.275	0.8668	42991	3187

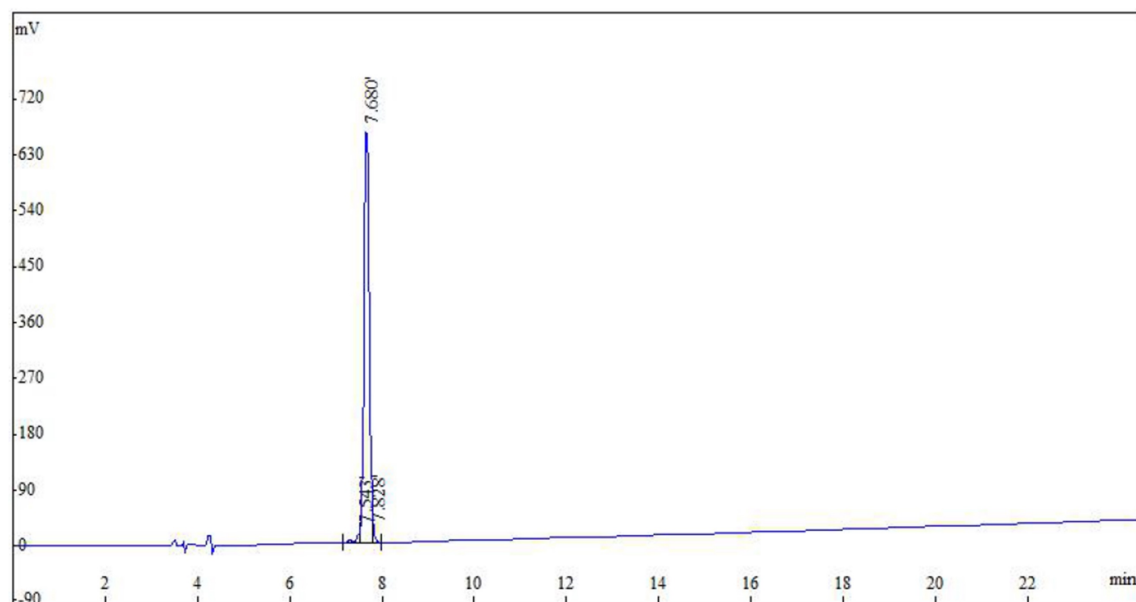
Total 100.0000

Mass trace



KY(NO₂)TEGESRGK(Abz)JZYKKG (34)

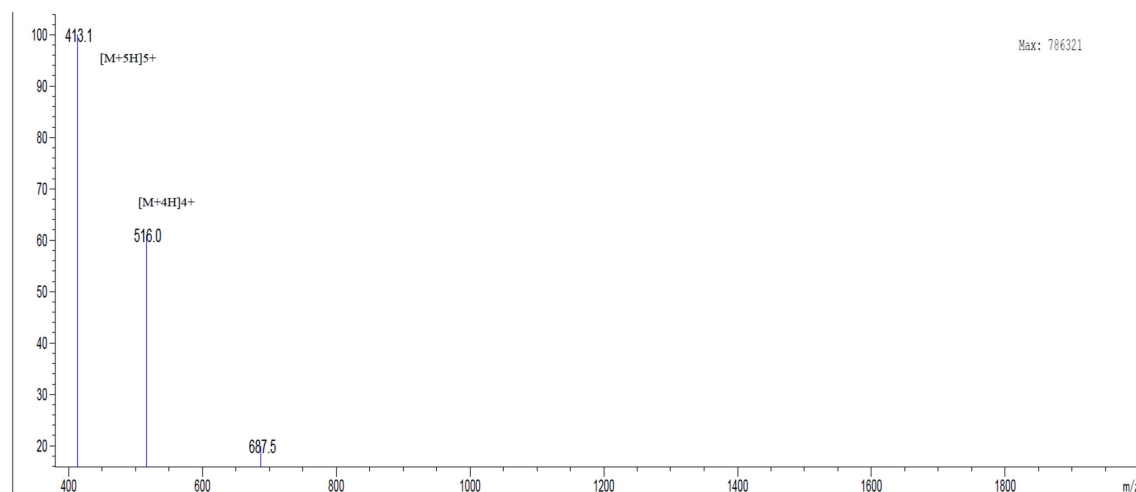
UV trace



Rank	Time	Conc.	Area	Height
1	7.543	1.8241	87205	17936
2	7.680	97.0825	4641206	662396
3	7.828	1.0934	52272	22738

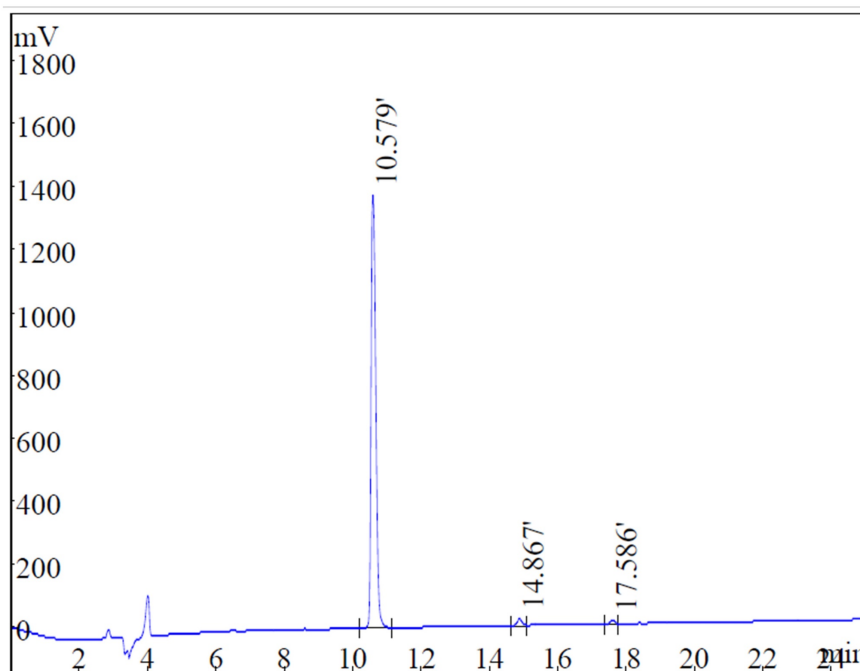
Total 100 4780683 703070

Mass trace



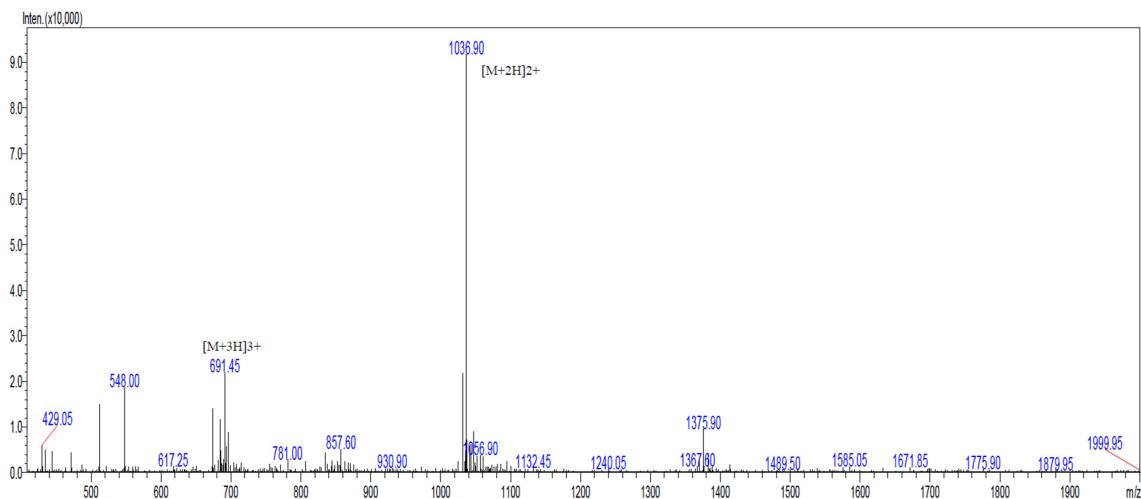
KY(NO₂)TETESOAK(Abz)IYFKKG (35)

UV trace



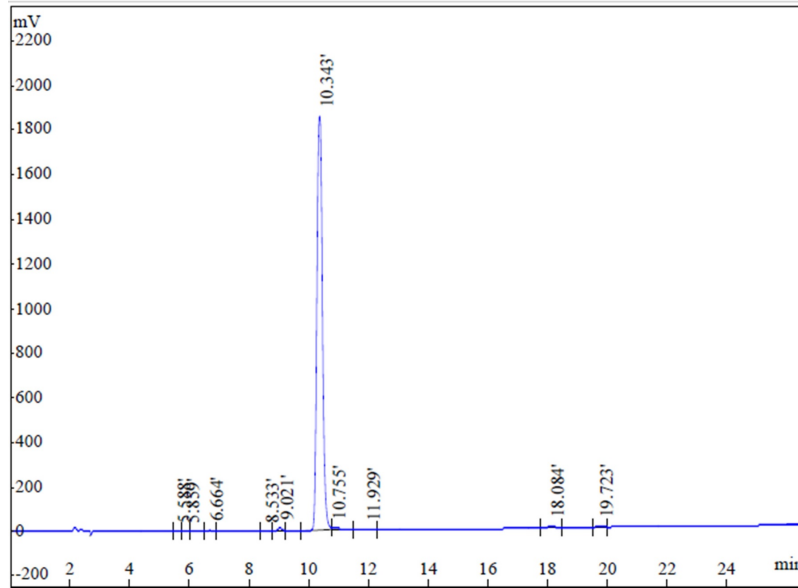
Rank	Time	Conc.	Area	Height
1	10.579	97.83	11728760	1375986
2	14.867	1.413	169350	20453
3	17.586	0.7528	90249	11981
Total		100.0000		

Mass trace



KY(NO₂)SESESRGK(Abz)IYYKKG (36)

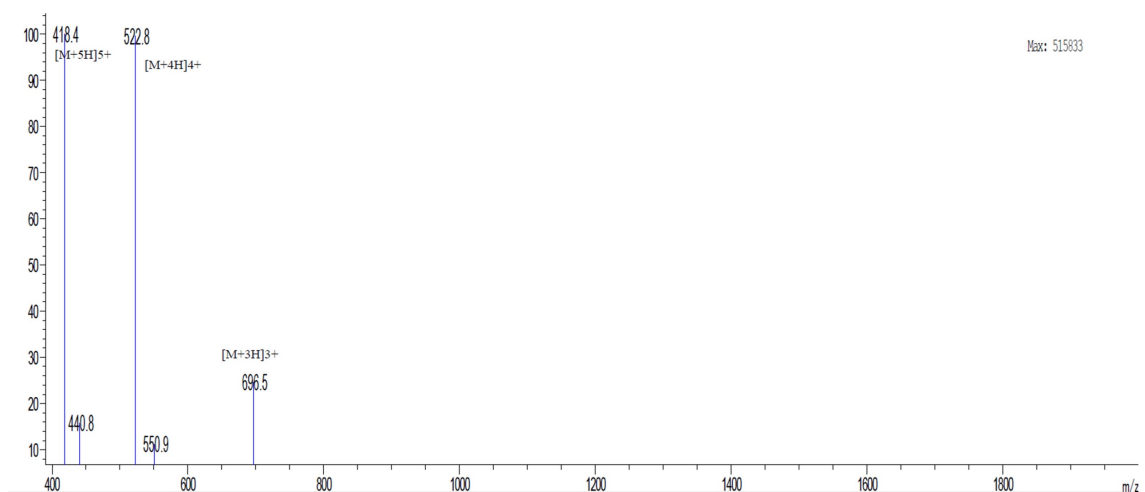
UV trace



Rank	Time	Conc.	Area	Height
1	5.588	0.1759	38845	5453
2	5.859	0.1707	37709	4824
3	6.664	0.1825	40306	5166
4	8.533	0.06911	15264	2050
5	9.021	0.4412	97436	11531
6	10.343	97.14	21455898	1854737
7	10.755	0.9527	210418	9880
8	11.929	0.1152	25438	2476
9	18.084	0.463	102250	6570
10	19.723	0.287	63393	5551

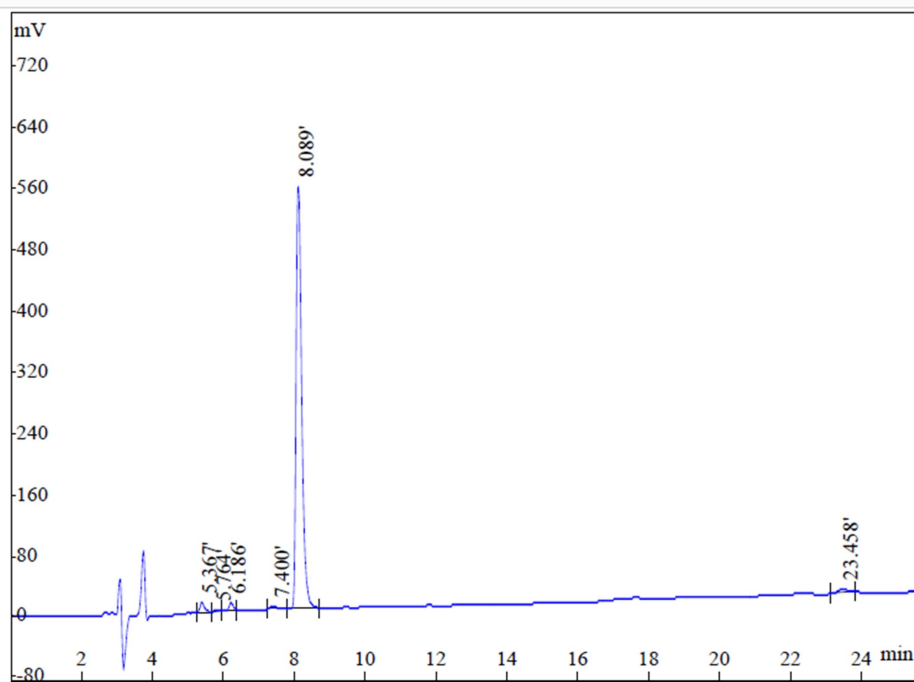
Total 100.0000

Mass trace



KY(NO₂)NDTESOAK(Abz)AHFUKG (37)

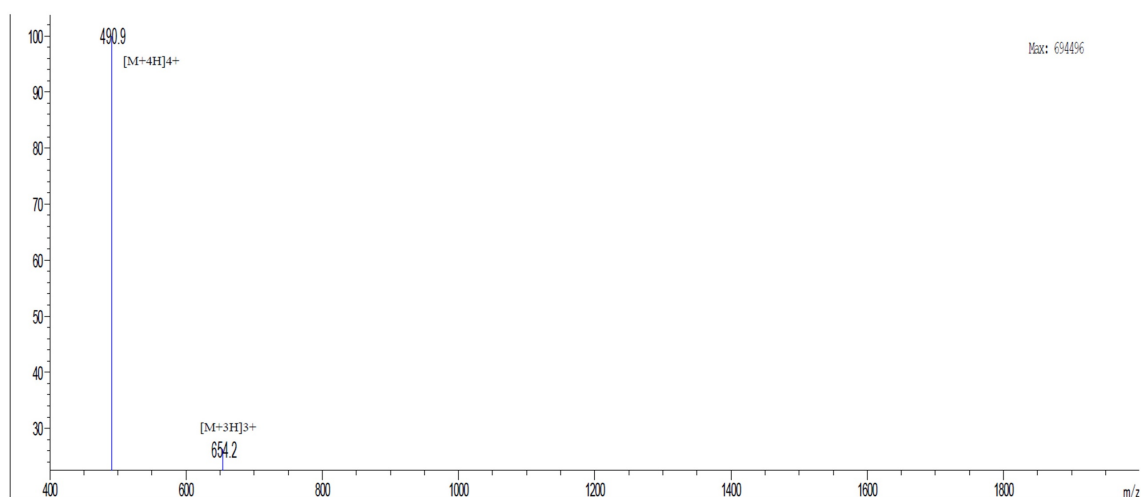
UV trace



Rank	Time	Conc.	Area	Height
1	5.367	1.563	88791	11881
2	5.764	0.1759	9992	1437
3	6.186	1.264	71813	9589
4	7.400	0.4651	26419	2430
5	8.089	95.12	5402438	550326
6	23.458	1.415	80389	4605

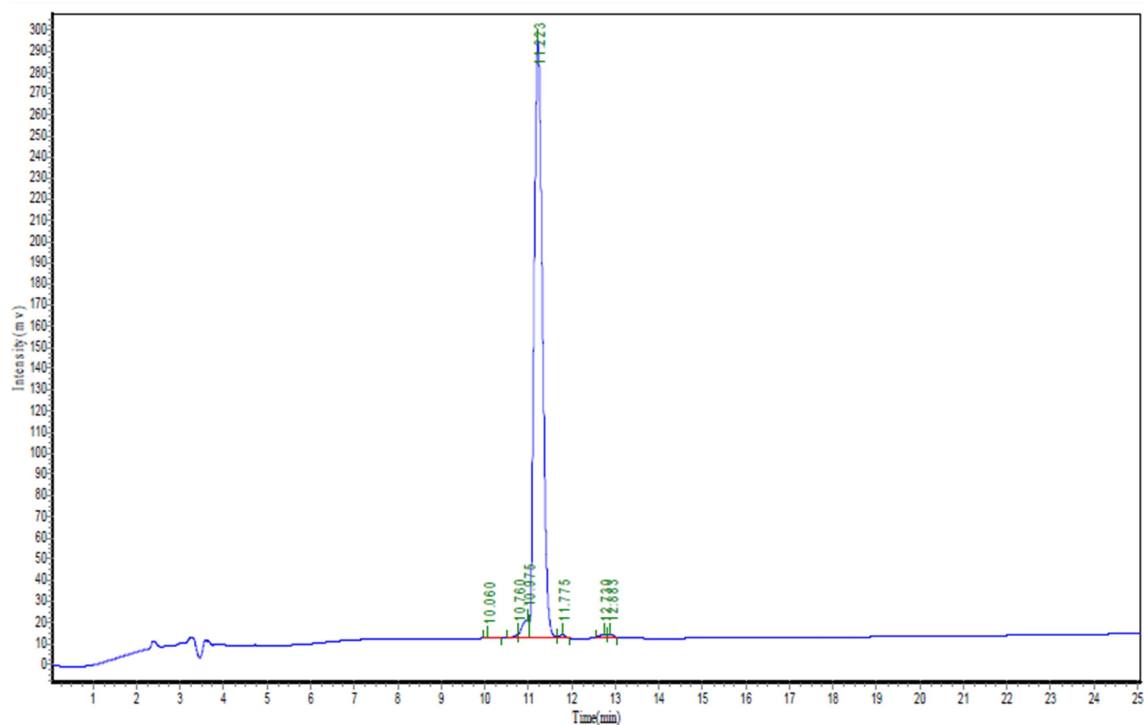
Total 100.0000

Mass trace



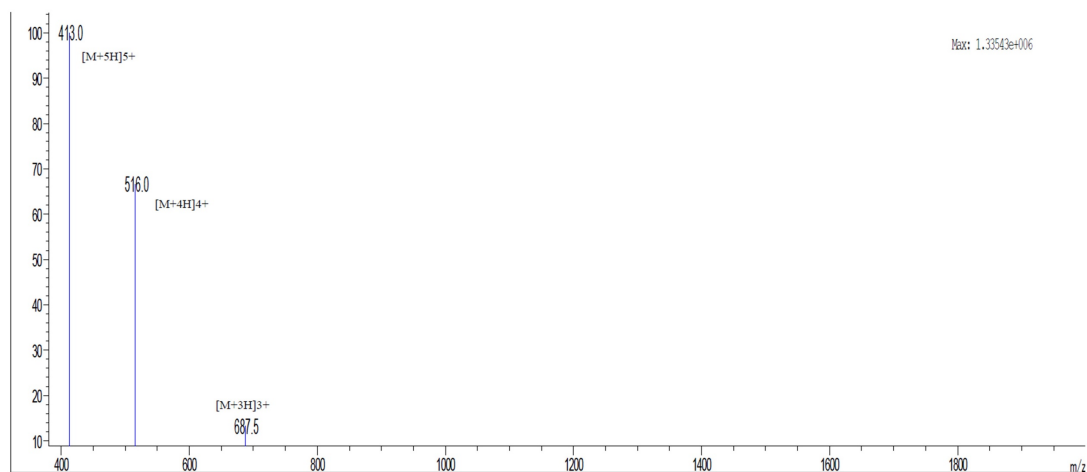
KY(NO₂)NESESKAK(Abz)VHYKKG (38)

UV trace



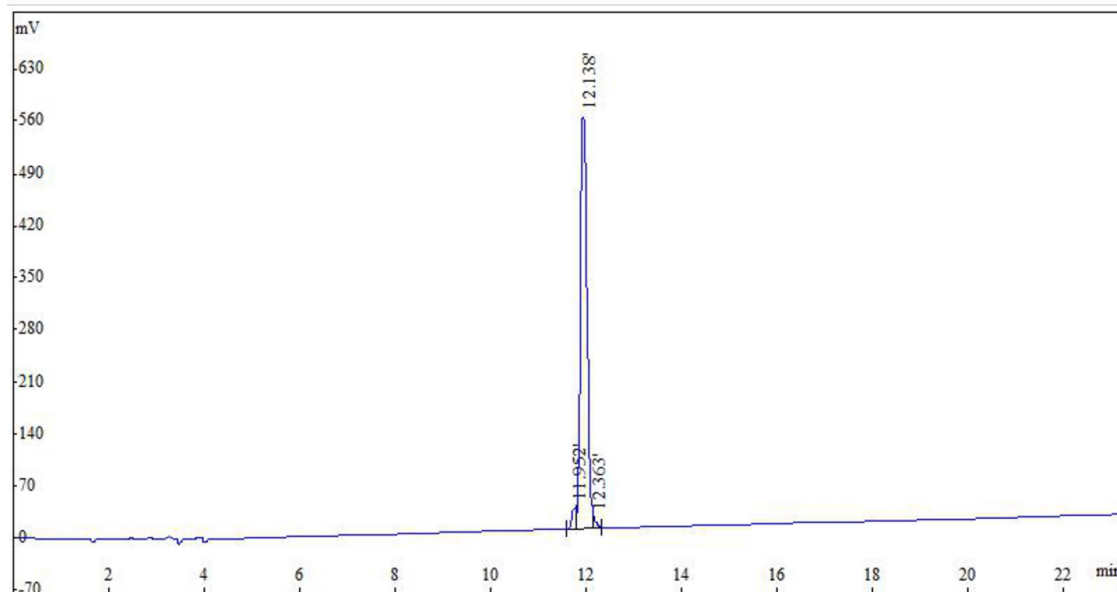
Peak	Time	Height	Area	Conc.
1	10.060	485.004	7658.506	0.1972
2	10.760	1679.758	8693.121	0.2239
3	10.975	8697.797	94104.883	2.4237
4	11.223	282049.031	3729261.500	96.0472
5	11.775	1437.979	15084.591	0.3885
6	12.730	1471.926	14147.716	0.3644
7	12.883	1450.531	13789.783	0.3552
Total				100.000

Mass trace



KY(NO₂)SENESRGK(Abz)IYYKKG (39)

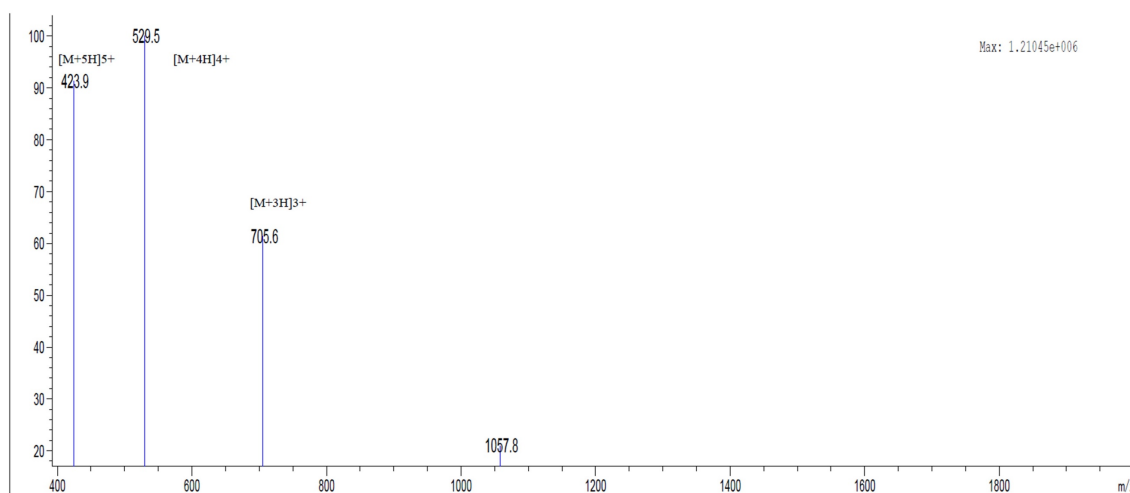
UV trace



Rank	Time	Conc.	Area	Height
1	11.952	3.5895	180849	27966
2	12.138	95.6302	4818110	551791
3	12.363	0.7803	39312	12293

Total 100 5038271 592050

Mass trace



8.3.3. Figures

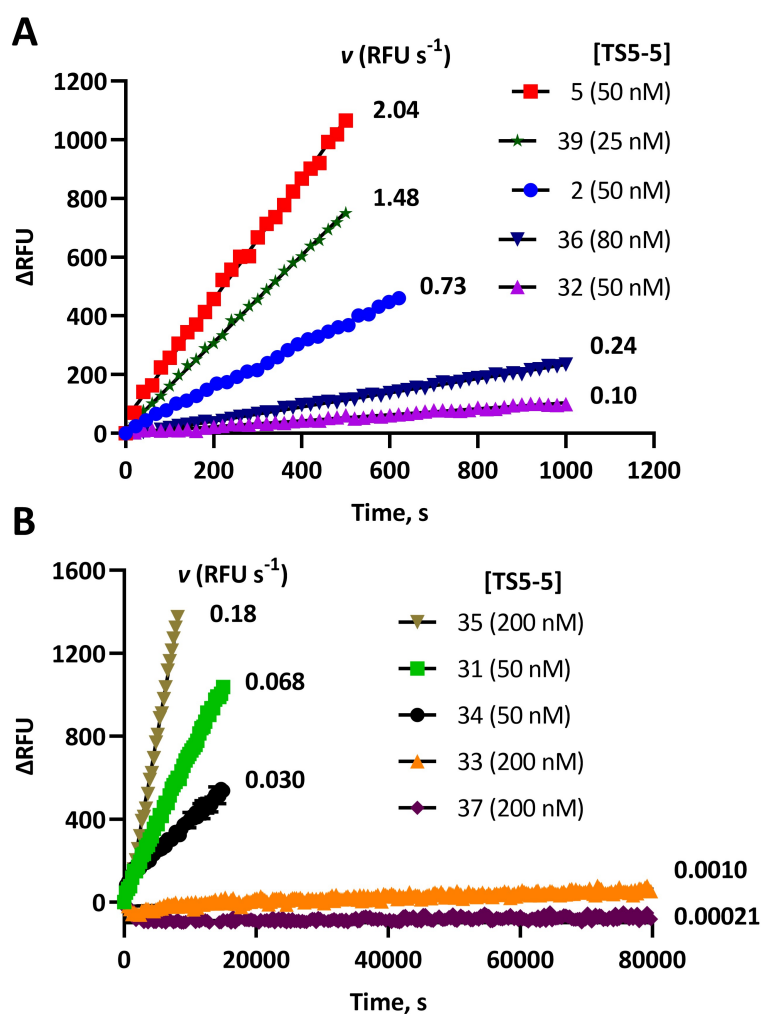


Figure 8.32. Cleavage of the current substrate, library parent compound and peptide hits by the ADAMTS-5 proteinase. (A) Representative individual rate data for cleavage of the current substrate 2, library parent compound 5 and peptide hits 32, 36 and 39 (all 20 μ M) by TS5-5; (B) representative individual rate data for cleavage of peptide hits 31, 33–35 and 37 (all 20 μ M) by TS5-5. All data was fitted by linear regression. The rate (v) is given in RFU s⁻¹. Each time point is expressed as Δ RFU (RFU at time t – RFU at time 0) $\pm$ SEM for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. Please note that data is shown for when the cleavage rate is linearly proportional to time.

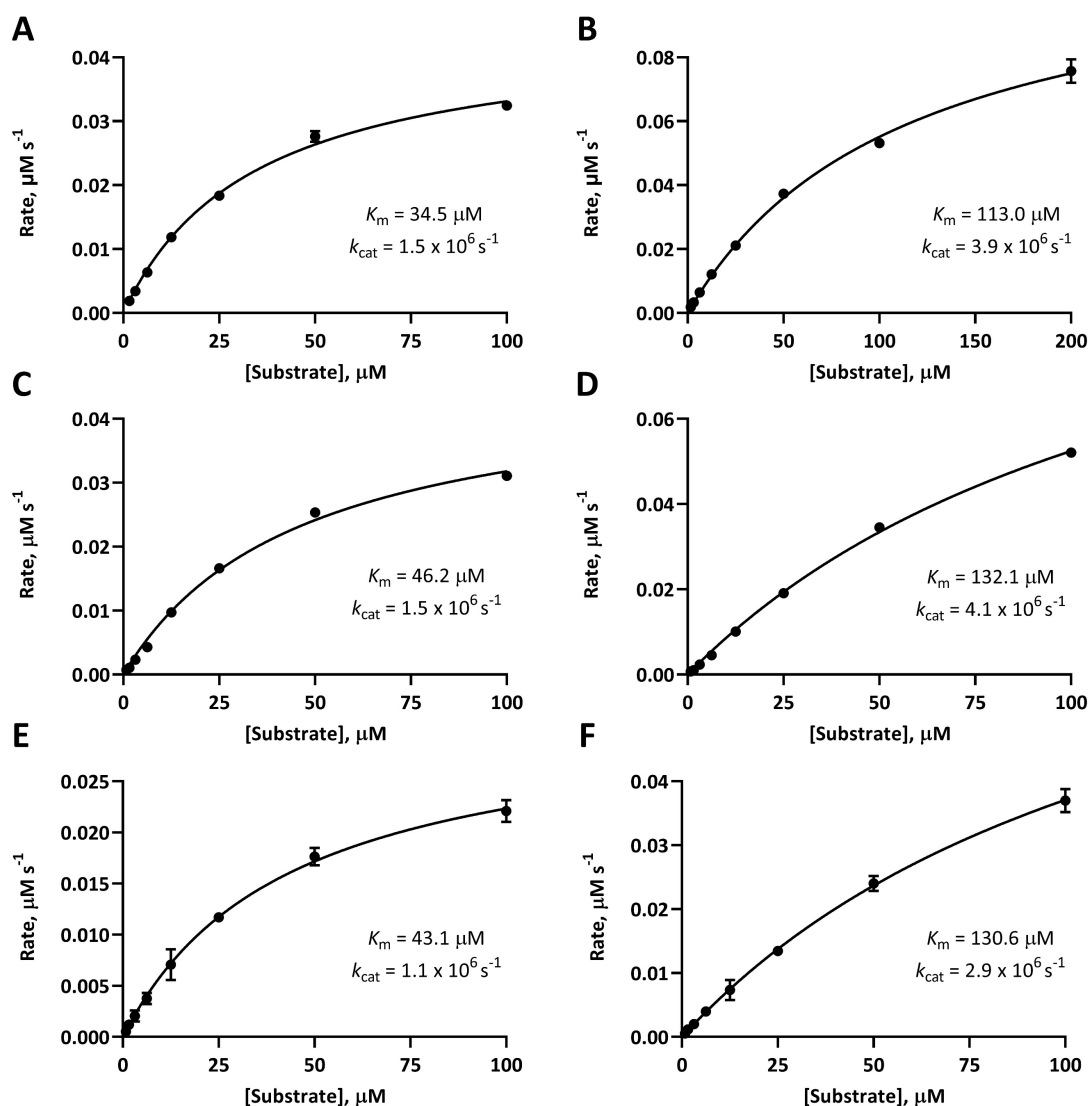


Figure 8.33. Michaelis-Menten saturation curves for cleavage of the current substrate 2 by ADAMTS-5. Individual curves showing cleavage of Abz-TESE~SRGAIY-Dpa-KK-NH₂ at different concentrations by a fixed concentration of TS5-5 (30 nM) prior to (A, C, E) and after (B, D, F) correction for the inner-filter effect. The rate at each concentration is expressed in $\mu\text{M s}^{-1} \pm \text{SEM}$ for $n = 3$ wells from a single experiment performed in triplicate at 37 °C.

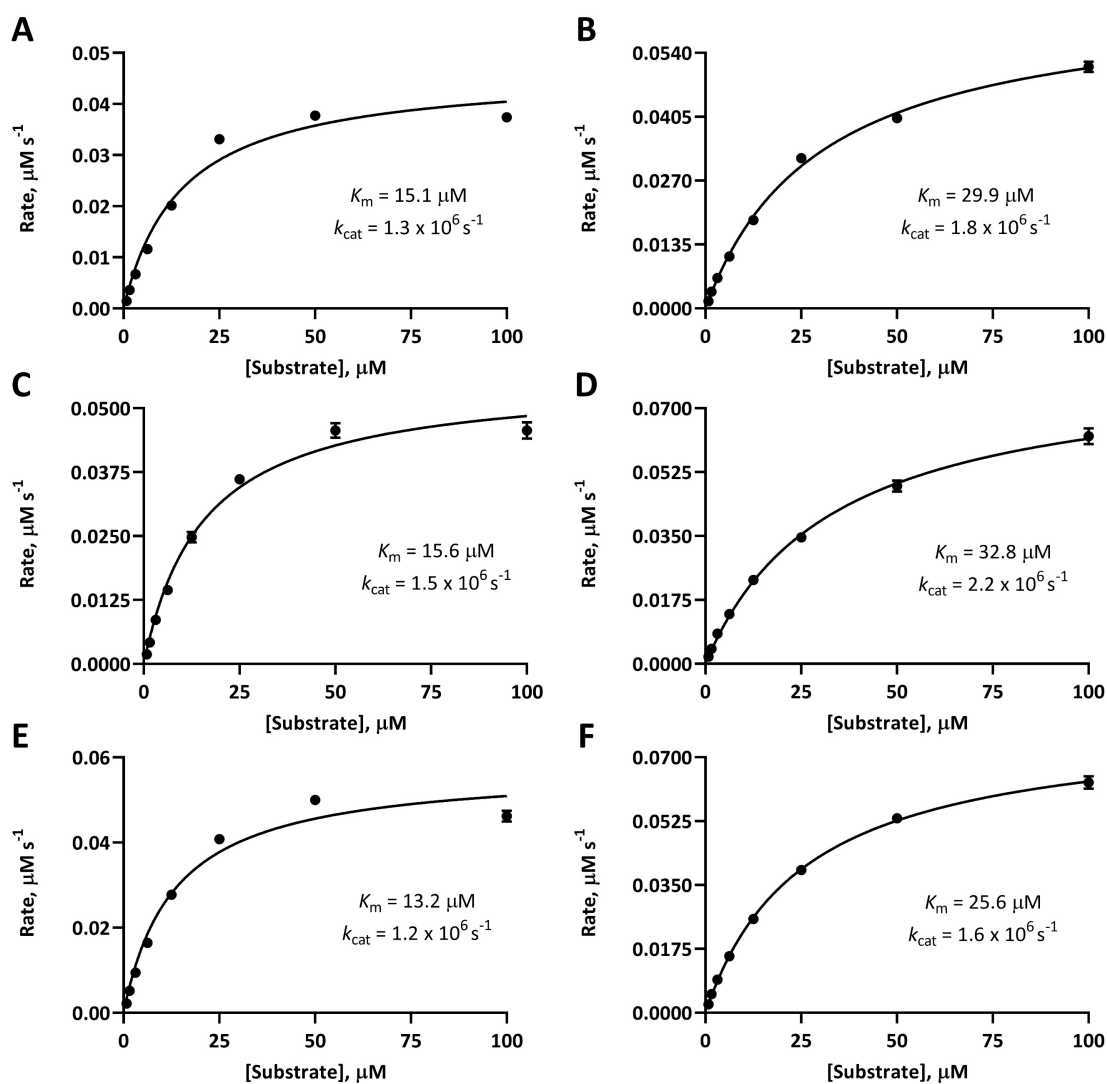


Figure 8.34. Michaelis-Menten saturation curves for cleavage of the library parent compound 5 by ADAMTS-5. Individual curves showing cleavage of KY(NO₂)TESER GK(Abz)IYYKKG at different concentrations by a fixed concentration of TS5-5 prior to (A, C, E) and after (B, D, F) correction for the inner-filter effect. The rate at each concentration is expressed in $\mu\text{M s}^{-1} \pm \text{SEM}$ for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. The concentration of TS5-5 for each experiment was 37 nM (A–D) or 50 nM (E, F).

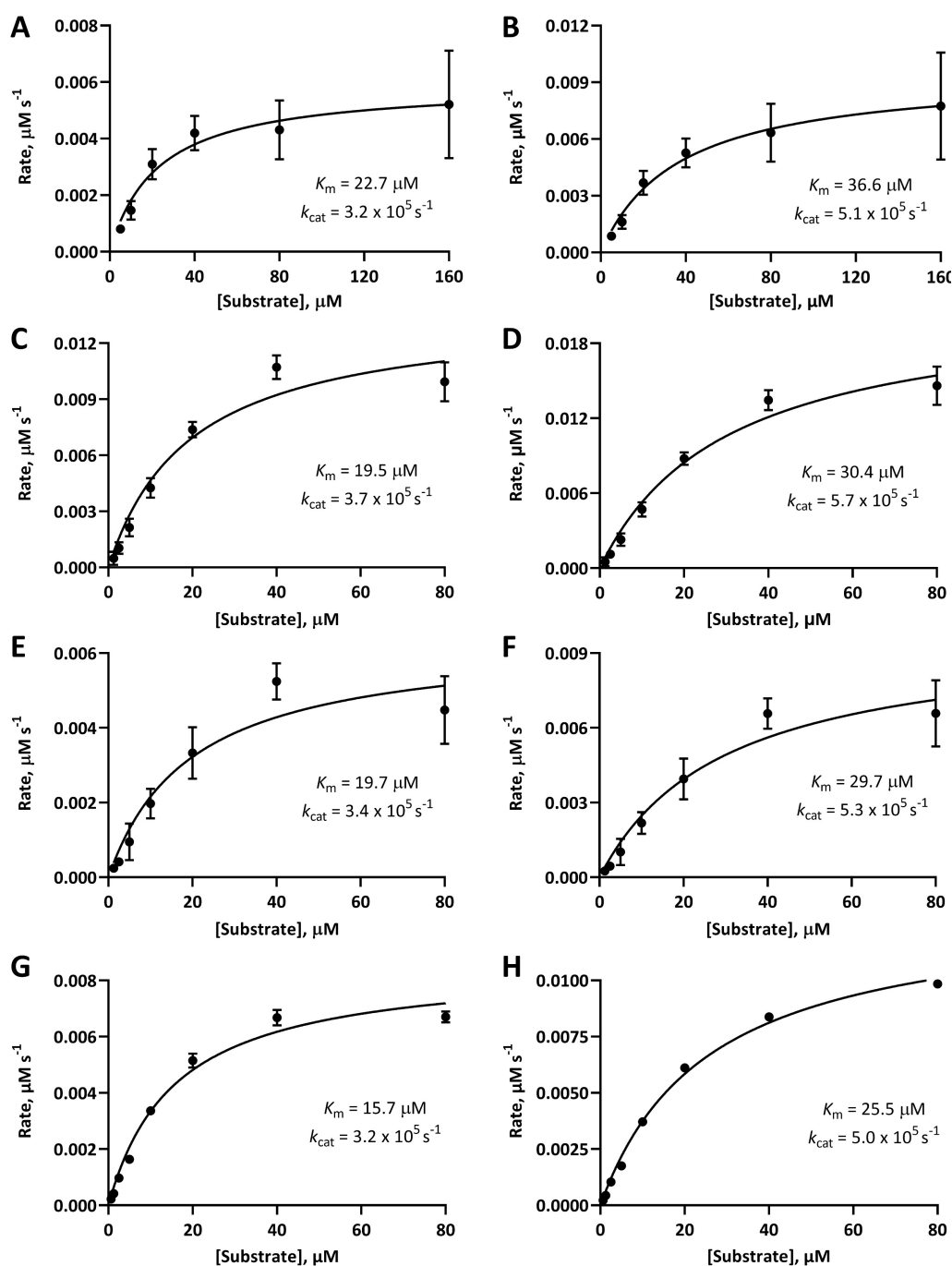


Figure 8.35. Michaelis-Menten saturation curves for cleavage of hit peptide 32 by ADAMTS-5. Individual curves showing cleavage of KY(NO₂)TENESRGK(Abz)IYYKKG at different concentrations by a fixed concentration of TS5-5 prior to (A, C, E, G) and after (B, D, F, H) correction for the inner-filter effect. The rate at each concentration is expressed in $\mu\text{M s}^{-1} \pm \text{SEM}$ for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. The concentration of TS5-5 for each experiment was 19 nM (A, B and E, F), 37 nM (C, D) or 27 nM (G, H).

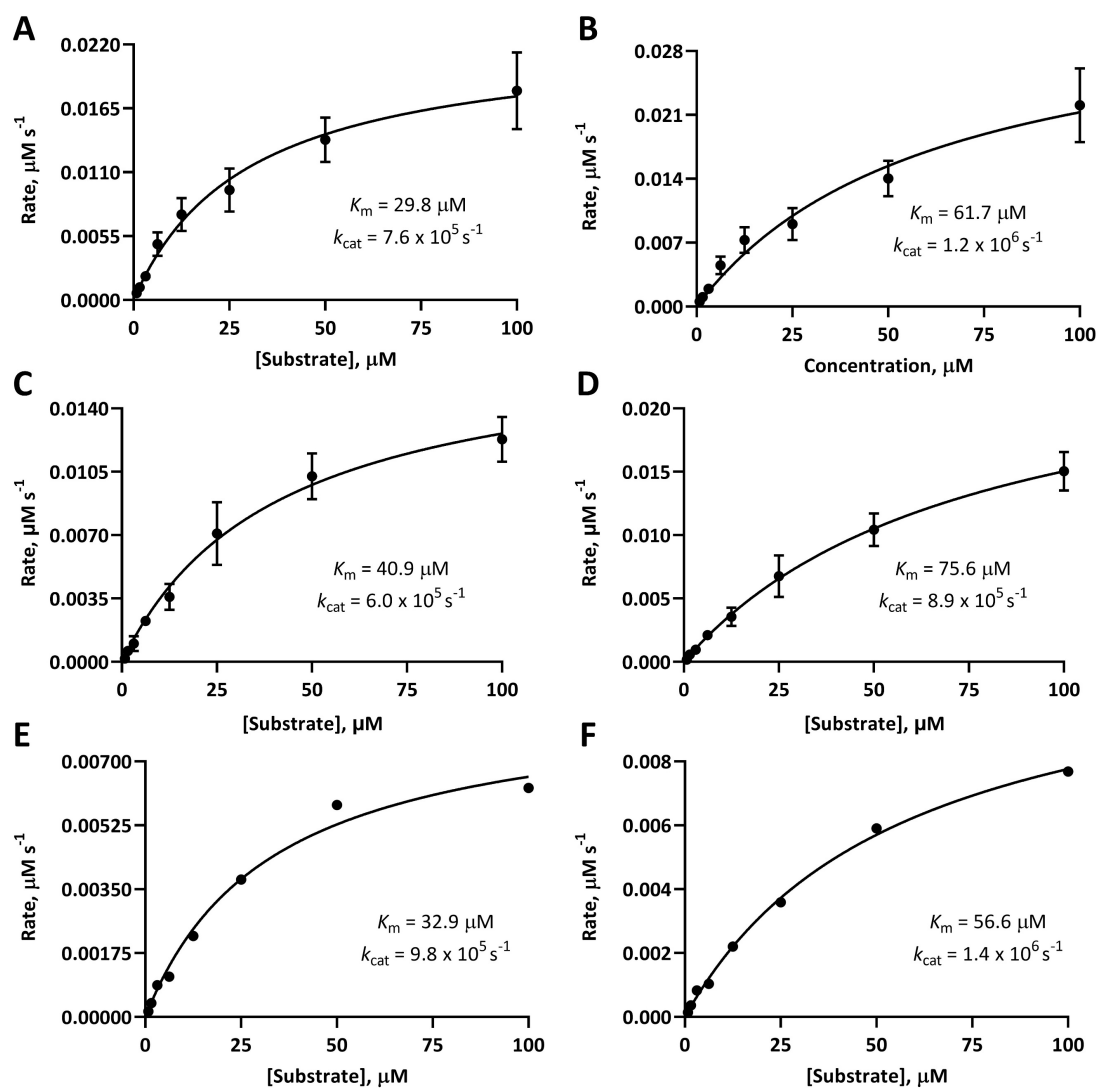


Figure 8.36. Michaelis-Menten saturation curves for cleavage of hit peptide 36 by ADAMTS-5. Individual curves showing cleavage of KY(NO₂)SESESRGK(Abz)IYYKKG at different concentrations by a fixed concentration of TS5-5 prior to (A, C, E) and after (B, D, F) correction for the inner-filter effect. The rate at each concentration is expressed in $\mu\text{M s}^{-1} \pm \text{SEM}$ for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. The concentration of TS5-5 for each experiment was 30 nM (A–D) or 90 nM (E, F).

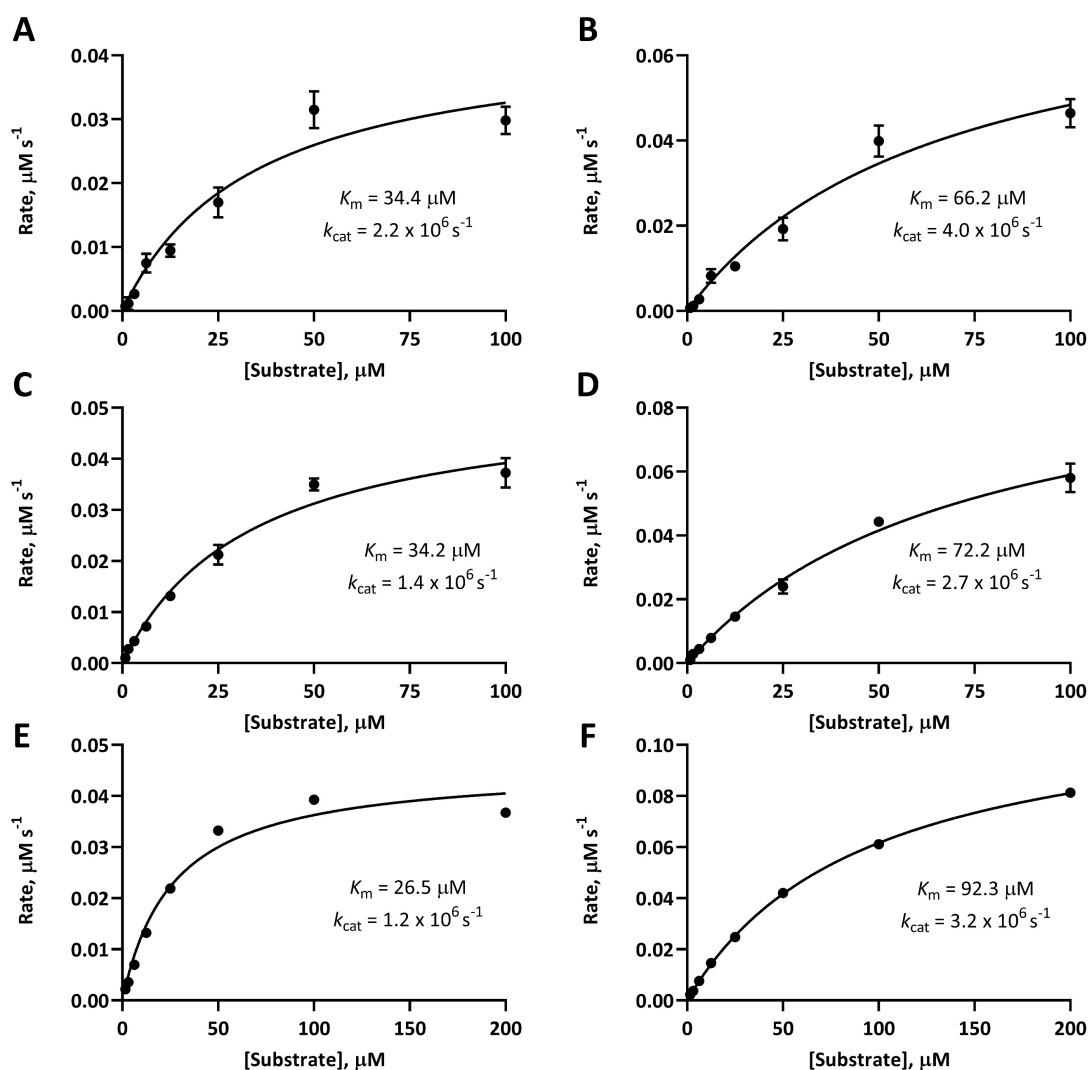


Figure 8.37. Michaelis-Menten saturation curves for cleavage of hit peptide 39 by ADAMTS-5.

Individual curves showing cleavage of KY(NO₂)SENESRGK(Abz)IYYKKG at different concentrations by a fixed concentration of TS5-5 (30 nM) prior to (A, C, E) and after (B, D, F) correction for the inner-filter effect. The rate at each concentration is expressed in $\mu\text{M s}^{-1} \pm \text{SEM}$ for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. The concentration of TS5-5 for each experiment was 20 nM (A, B) or 37 nM (C–F).

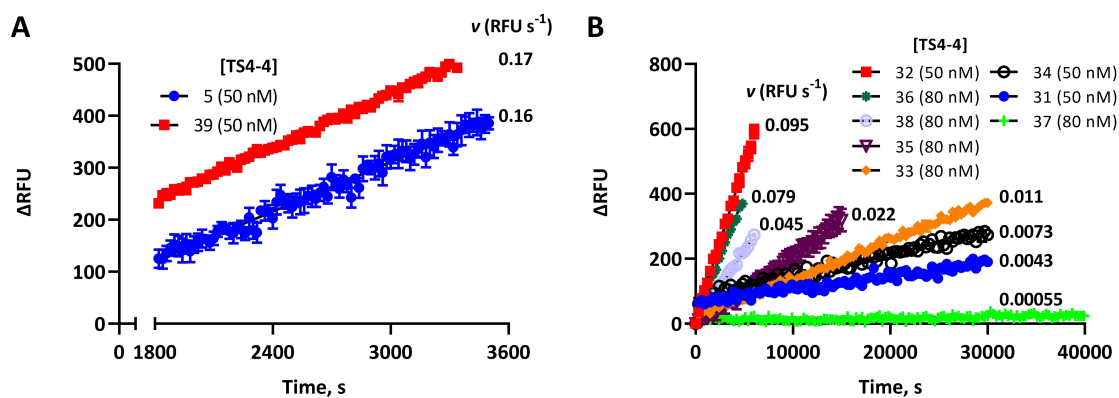


Figure 8.38. Cleavage of the library parent compound and peptide hits by ADAMTS-4. (A) Representative individual rate data for cleavage of the library parent compound 5 and peptide hit 39 (all 20 μM) by TS4-4 (50 nM); (B) representative individual rate data for cleavage of peptide hits 31–38 (all 20 μM) by TS4-4 (50 nM or 80 nM). All data was fitted by linear regression. The rate (v) is given in RFU s⁻¹. Each time point is expressed as ΔRFU (RFU at time t – RFU at time 0) $\pm$ SEM for $n = 3$ wells from a single experiment performed in triplicate at 37 °C. *Please note that data is shown for when the cleavage rate is linearly proportional to time.*

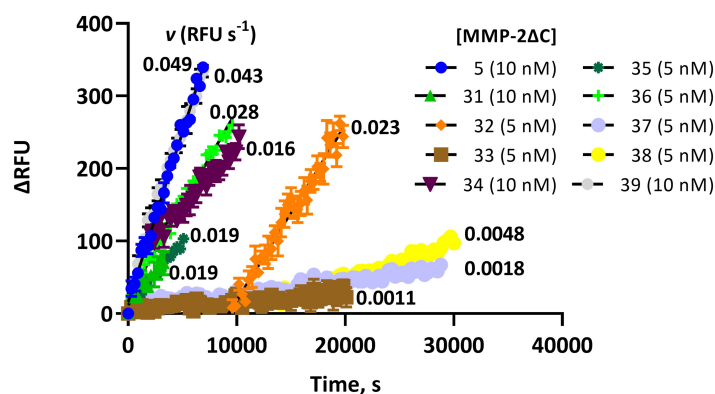


Figure 8.39. Cleavage of the library parent compound and peptide hits by MMP-2. Representative individual rate data for cleavage of the library parent compound 5 and peptide hits 31–39 (all 20 μM) by MMP-2 ΔC (5 nM or 10 nM) fitted by linear regression. The rate (v) is given in RFU s^{-1} . Each time point is expressed as ΔRFU ($\text{RFU at time } t - \text{RFU at time } 0$) $\pm$ SEM for $n = 3$ wells from a single experiment performed in triplicate at 37 $^{\circ}\text{C}$. *Please note that data is shown for when the cleavage rate is linearly proportional to time.*

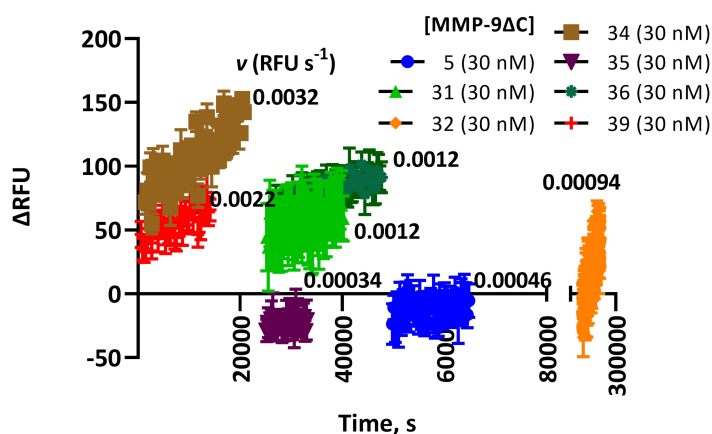


Figure 8.40. Cleavage of the library parent compound and peptide hits by MMP-9. Representative individual rate data for cleavage of the library parent compound 5 and peptide hits 31–32, 34–36 and 39 (all 20 μM) by MMP-9 ΔC (30 nM) fitted by linear regression. The rate (v) is given in RFU s^{-1} . Each time point is expressed as ΔRFU ($\text{RFU at time } t - \text{RFU at time } 0$) $\pm$ SEM for $n = 3$ wells from a single experiment performed in triplicate at 37 $^{\circ}\text{C}$. Please note that data is shown for when the cleavage rate is linearly proportional to time. The RFU values below zero indicate a very low level of cleavage below background fluorescence intensity and are not to be confused with an absence of cleavage. These values could not be improved through multiple repeats of the same experiment with longer incubation times or the highest proteinase concentration available.