The ubiquitin pathway in host-parasite interactions during infection by *Trichinella spiralis*

Rhiannon Rose White

Imperial College London
Infection and Immunity
Department of Life Sciences

Thesis submitted for the degree of Doctorate of Philosophy, 2014
Declaration of originality

I declare that all of the work presented in this thesis is my own, and that all else, figures, images, ideas, quotations, data, results, published or unpublished, have been acknowledged and referenced.

This thesis contains some modified material from my PhD transfer report.

Rhiannon Rose White
February 2014

Copyright declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Publications arising from this thesis

Morrow et al. 2013, Biochemistry [3]
Abstract

These studies present the identification of the first parasite secreted ubiquitin pathway enzyme, and report a system for the characterisation of the role of *Trichinella spiralis* secreted proteins in the context of mammalian muscle cell re-programming.

*Trichinella spiralis* invade terminally differentiated myofibres, secreting proteins (SP) into the host cell and inducing dedifferentiation and cell cycle re-entry. Since myogenic differentiation and the cell cycle are heavily influenced by the ubiquitin pathway, this study used *T. spiralis* as a model to investigate the extent to which parasites specifically manipulate the ubiquitin pathway during infection.

Ubiquitin is ligated to protein substrates by E1, E2 and E3 enzymes, and removed by deubiquitinating enzymes (DUBs). This ubiquitin 'tag' regulates the fate and function of the substrate protein. A DUB expressed by *T. spiralis*, TsUCH37, was characterised as a conserved proteasome interaction partner, however no evidence was found of the presence of this, or any other *T. spiralis* DUB in the muscle larvae SP. Upon further investigation an E2 enzyme, TsUBE2L3, was identified. The ubiquitin conjugation activity of *Trichinella spiralis* SP was confirmed, and was only possible in the presence of human E1 and E3 enzyme partners. The effect of TsUBE2L3 on mammalian muscle cells was investigated by expressing the *T. spiralis* protein in a mouse muscle cell line. Although no significant effect on the cell cycle or differentiation state of the muscle cells was observed, TsUBE2L3 expression led to a significant reduction of the tumour suppressor protein, p53 that was confirmed to occur at the protein level.

*Trichinella spiralis* strikes a delicate balance between host cell modulation and host protection. This thesis builds on the proposal that the host-targeted muscle cell modulators of *T. spiralis* may inspire the development of parasite-derived therapeutics for the treatment of disease.
Acknowledgements

Without the following people, this work would not have been possible.

Massive thanks to my supervisor Katerina, for support, guidance, inspiration, patience and understanding, of both the scientific and personal kind.

Sachiko Miyata for expertise, training and friendship.

My co-supervisor Kleoniki Gounaris and Murray Selkirk for expertise, advice and reagents.

The community of the 6th floor (and beyond) of the Sir Alexander Fleming Building; the Christophides, Dallman and McMullen groups, Marco Mol and especially the Sinden group: Andrew Blagborough, Sara Marques, Ursula Straschil, Chandra Ramakrishnan, Leanna Upton, Andrea Ruecker, Michael Delves, Mark Tunnicliff and Bob Sinden, with whom almost everything was shared, from technical problems and reagents, to drinks and laughs.


My family and friends, who have supported and sustained me throughout, and most of all the Aguirres the Holmes and the Whites.

And a most special thank you to Sam, for unconditional care, support, encouragement, inspiration and perspective.

Funding

My PhD was funded by a Medical Research Council Imperial College London Doctoral Training Grant award.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADRM1</td>
<td>Adhesion Regulating Molecule 1</td>
</tr>
<tr>
<td>AMC</td>
<td>7-amido-4-methylcoumarin</td>
</tr>
<tr>
<td>APC</td>
<td>Antigen Presenting Cell</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium Persulphate</td>
</tr>
<tr>
<td>AU</td>
<td>Absorbance Units</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchorinic Acid</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Assignment Search Tool</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>CIP</td>
<td>Calf Intestinal Phosphatase</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CPM</td>
<td>Counts Per Minute</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s Modified Eagle Medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl Sulphoxide</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>DUB</td>
<td>Deubiquitinating enzyme</td>
</tr>
<tr>
<td>E6-AP</td>
<td>E6-Associated Protein</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein Barr Virus</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ES</td>
<td>Excretory-Secretory</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed Sequence Tag</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence Activated Cell Sorting</td>
</tr>
<tr>
<td>FBS</td>
<td>Foetal Bovine Serum</td>
</tr>
<tr>
<td>FDR</td>
<td>False Discovery Rate</td>
</tr>
<tr>
<td>GAG</td>
<td>Group Antigens</td>
</tr>
<tr>
<td>GSFP</td>
<td>Gene Specific Forward Primer</td>
</tr>
<tr>
<td>GSRP</td>
<td>Gene Specific Reverse Primer</td>
</tr>
<tr>
<td>HA</td>
<td>Haemagglutinin</td>
</tr>
<tr>
<td>HAUSP</td>
<td>Herpesvirus-Associated USP</td>
</tr>
<tr>
<td>HECT</td>
<td>Homologous to E6-AP Carboxy Terminus</td>
</tr>
<tr>
<td>HEK</td>
<td>Human Embryonic Kidney</td>
</tr>
<tr>
<td>HEPES</td>
<td>2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency Virus</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish Peroxidase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>HSV</td>
<td>Herpes Simplex Virus</td>
</tr>
<tr>
<td>IB</td>
<td>Immuno-blot</td>
</tr>
<tr>
<td>IFA</td>
<td>Immuno-fluorescence Assay</td>
</tr>
<tr>
<td>IL-10</td>
<td>Interleukin 10</td>
</tr>
<tr>
<td>IP</td>
<td>Immuno-precipitation</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-Thiogalactopyranoside</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal Ribosome Entry Site</td>
</tr>
<tr>
<td>JAMM</td>
<td>JAB1/MPN/Mov34 domain</td>
</tr>
<tr>
<td>Lac</td>
<td>alpha-amino-beta-lactone</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani</td>
</tr>
<tr>
<td>M-MLV</td>
<td>Moloney Murine Leukemia Virus</td>
</tr>
<tr>
<td>MES</td>
<td>2-(N-morpholino)ethanesulfonic acid</td>
</tr>
<tr>
<td>MESNa</td>
<td>sodium 2-sulfanylethanesulfonate</td>
</tr>
<tr>
<td>MHC</td>
<td>Major Histocompatibility Complex</td>
</tr>
<tr>
<td>MIR</td>
<td>Mannosyltransferase Inositol 1,4,5-trisphosphate receptor Ryanodine receptor</td>
</tr>
<tr>
<td>MJD</td>
<td>Machado-Josephin domain</td>
</tr>
<tr>
<td>MTT</td>
<td>3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular Weight</td>
</tr>
<tr>
<td>MWCO</td>
<td>Molecular Weight Cut Off</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>NEDD4-L</td>
<td>Neural Precursor Cell Expressed Developmentally Down-regulated 4-like</td>
</tr>
<tr>
<td>Nedd8</td>
<td>Neural Precursor Cell Expressed, Developmentally Down-regulated 8</td>
</tr>
<tr>
<td>NEDP1</td>
<td>NEDD8-specific Protease 1</td>
</tr>
<tr>
<td>NEM</td>
<td>N-ethylmaleimide</td>
</tr>
<tr>
<td>ni-NTA</td>
<td>nickel-Nitrilotriacetic Acid</td>
</tr>
<tr>
<td>NP-40</td>
<td>Nonident P-40</td>
</tr>
<tr>
<td>nr</td>
<td>non-redundant</td>
</tr>
<tr>
<td>OTU</td>
<td>Ovarian Tumour Domain</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PBST</td>
<td>Phosphate Buffered Saline with 0.1% Tween</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulphonyl Fluoride</td>
</tr>
<tr>
<td>PPO</td>
<td>Polyphenol Oxidase</td>
</tr>
<tr>
<td>PRP38</td>
<td>Pre-mRNA Processing factor 38</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene Fluoride</td>
</tr>
<tr>
<td>qRT</td>
<td>quantitative Real-Time</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid Amplification of cDNA Ends</td>
</tr>
<tr>
<td>RFU</td>
<td>Relative Fluorescence Units</td>
</tr>
<tr>
<td>RING</td>
<td>Really Interesting New Gene</td>
</tr>
<tr>
<td>RIPA</td>
<td>Radio-Immuno-Precipitation assay buffer</td>
</tr>
<tr>
<td>RLM</td>
<td>RNA Ligase-Mediated</td>
</tr>
<tr>
<td>RPMI</td>
<td>Roswell park memorial institution medium</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse Transcriptase</td>
</tr>
<tr>
<td>RWD</td>
<td>RING finger and WD repeat domain</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>SMART</td>
<td>Simple Modular Architecture Research Tool</td>
</tr>
<tr>
<td>SP</td>
<td>Secreted Proteins</td>
</tr>
<tr>
<td>SUMO</td>
<td>Small Ubiquitin-like Modifier</td>
</tr>
<tr>
<td>SV40</td>
<td>Simian Virus 40</td>
</tr>
<tr>
<td>TAP</td>
<td>Tobacco Acid Phosphatase</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylethylenediamine</td>
</tr>
<tr>
<td>TF3BOK</td>
<td>Trifluoromethylbenzyloxymethyl-ketone</td>
</tr>
<tr>
<td>TGF</td>
<td>Transforming Growth Factor</td>
</tr>
<tr>
<td>TNF</td>
<td>Tumour Necrosis Factor</td>
</tr>
<tr>
<td>TRE</td>
<td>Tetracycline Response Element</td>
</tr>
<tr>
<td>TUBE</td>
<td>Tandem Ubiquitin Binding Entity</td>
</tr>
<tr>
<td>Ub</td>
<td>Ubiquitin</td>
</tr>
<tr>
<td>UBC</td>
<td>Ubiquitin Conjugating</td>
</tr>
<tr>
<td>UCH</td>
<td>Ubiquitin C-terminal Hydrolase</td>
</tr>
<tr>
<td>USP</td>
<td>Ubiquitin Specific Peptidase</td>
</tr>
<tr>
<td>VME</td>
<td>Vinylmethylester</td>
</tr>
<tr>
<td>VS</td>
<td>Vinylmethylsulphone</td>
</tr>
<tr>
<td>VSV-G</td>
<td>Vesicular Stomatitis Indiana Virus Glycoprotein</td>
</tr>
<tr>
<td>WDR</td>
<td>WD Repeat domain</td>
</tr>
</tbody>
</table>
Table of Contents

Declaration of originality ................................................................. 2
Copyright declaration ................................................................. 2
Publications arising from this thesis .................................................. 2
Abstract ....................................................................................... 3
Acknowledgements ........................................................................... 4
Funding ......................................................................................... 4
Abbreviations .................................................................................. 5
Table of Contents ........................................................................... 8

List of Figures .................................................................................. 11
Chapter 1 ......................................................................................... 11
Chapter 3 ......................................................................................... 11
Chapter 4 ......................................................................................... 12
Chapter 5 ......................................................................................... 13

List of Tables ..................................................................................... 14
Chapter 2 ......................................................................................... 14
Chapter 3 ......................................................................................... 14
Chapter 4 ......................................................................................... 14

List of appendices ................................................................................... 15

Chapter 1: Introduction ....................................................................... 17
  1.1. The ubiquitin pathway ............................................................... 17
       1.2. Ubiquitin conjugation ............................................................ 18
       1.3. The 26S proteasome .............................................................. 22
       1.4. Ubiquitin hydrolysis by DUBs .............................................. 24
  1.5. The ubiquitin pathway in disease .............................................. 27
       1.6. Muscle wasting disease ....................................................... 27
       1.7. Neurodegenerative disease .................................................. 27
       1.8. Cancer ................................................................................. 28
  1.9. The ubiquitin pathway in infectious disease ................................ 30
       1.10. The ubiquitin pathway in infection by parasites .................. 32
  1.11. Parasitic worms and their secreted proteins ............................... 33
       Trichinella spiralis ................................................................. 35
       1.12. History and epidemiology .................................................... 35
       1.13. The life cycle of Trichinella spiralis ..................................... 37
       1.14. Muscle cell differentiation, and T. spiralis induced dedifferentiation ................................................. 39
       1.15. Trichinella spiralis and the nurse cell .................................. 42
       1.16. Treatment of trichinellosis .................................................. 46
       1.17. The secreted proteins of T. spiralis muscle larvae .......................... 47
       1.18. T. spiralis as a model for investigating the ubiquitin pathway in host-parasite interactions ............... 50
  1.19 Aims of the project .................................................................. 51

Chapter 2: Materials and methods .......................................................... 53
  2.1. Bioinformatics analyses ............................................................ 53
  2.2. Parasite isolation, culture, re-infection and collection of secreted proteins ................................................. 53
       Mammalian cell culture ............................................................ 54
  2.3. HEK 293T cells (adherent) ....................................................... 54
  2.4. C2C12 myoblasts (adherent) ...................................................... 54
  2.5. C2C12 differentiation (adherent) ................................................ 54
Chapter 3: Characterisation of TsUCH37, a deubiquitinating enzyme expressed by T. spiralis muscle larvae

Chapter 3 Results:

3.1. Finding the true open reading frame for a T. spiralis DUB orthologue
3.2. Cloning of the confirmed open reading frame of TsUCH37 and expression and purification of recombinant 6His-TsUCH37
3.3. Validation of the ubiquitin hydrolysis activity of recombinant TsUCH37
3.4. Analysis of the Nedd8 hydrolysis activity of recombinant TsUCH37 and the exploration of dual Ub/Nedd8 hydrolysis by UCH domain DUBs
3.5. Analysis of proteins that associate with TsUCH37
3.6. Investigation of the potential of TsUCH37 as a drug target

Summary
Discussion .................................................................................................................. 122

Chapter 4: ........................................................................................................... 128
Looking for DUBs and ubiquitin conjugation enzymes in the secreted proteins of T. spiralis muscle larvae ................................................................. 128

Background ........................................................................................................... 128

Chapter 4 Results: ............................................................................................. 134
Looking for DUB and deNeddylation activity in T. spiralis secreted proteins ... 134
4.1. Production of the ubiquitin-based inhibitor probe HA-Ub-VME ................. 134
4.2. Reaction of ubiquitin-based inhibitor probes with T. spiralis secreted proteins ............................................................... 138
4.3. Verification of the absence of DUB and deNeddylation activity in the T. spiralis secreted proteins ............................................................. 149
Looking for ubiquitin activating (E1), conjugating (E2) and ligating (E3) activity in T. spiralis secreted proteins ........................................................................ 151
4.4. Production of modified ubiquitin-based inhibitor probes ....................... 151
4.5. An in vitro ubiquitin conjugation assay of T. spiralis secreted proteins .............................................................. 155
Identification of the proteins responsible for the ubiquitin conjugating (E2) and ligating (E3) activity in T. spiralis secreted proteins ................................ 158
4.6. Bioinformatics-based prediction of the T. spiralis secretome ................. 158
4.7. Further analysis of 2 putative T. spiralis E3 proteins predicted to be secreted by bioinformatics .............................................................................. 160

Summary ............................................................................................................... 172

Discussion ........................................................................................................... 173

Chapter 5: ........................................................................................................... 182
Characterisation of ubiquitin pathway enzymes secreted by T. spiralis muscle larvae ........................................................................................................ 182

Background ......................................................................................................... 182

Chapter 5 Results: ............................................................................................. 185
5.1. Validation of the secretion of TsUBE2L3 (E2 enzyme) and TsRNF13 (E3 enzyme) by T. spiralis muscle larvae .............................................................. 185
5.2. Confirmation of the full coding sequence of TsUBE2L3 by RACE-PCR ... 187
5.3. Differentiation of the C2C12 mouse skeletal muscle cell line from myoblasts into myotubes ............................................................... 193
5.4. Cloning and expression of TsUBE2L3-HA in C2C12 skeletal muscle cells ................................................................................................. 195
Characterisation of the function of the secreted T. spiralis E2 enzyme, UBE2L3: ........................................................................................................... 203
5.5. The effect of TsUBE2L3-HA on C2C12 skeletal muscle cell morphology. 203
5.6. The effect of TsUBE2L3-HA on C2C12 skeletal muscle cell differentiation factors ........................................................... 207
5.7. The effect of TsUBE2L3-HA on protein synthesis in C2C12 skeletal muscle cells .............................................................. 210
5.8. The effect of TsUBE2L3-HA on C2C12 skeletal muscle cell cycle factors 212
5.9. The effect of TsUBE2L3-HA on C2C12 skeletal muscle proliferation ....... 214
5.10. The effect of TsUBE2L3-HA on p53 in C2C12 skeletal muscle cells ....... 221

Summary of results ............................................................................................. 228

Discussion ........................................................................................................... 229

Chapter 6: ........................................................................................................... 236
Final Discussion ................................................................................................... 236
Significance of study ............................................................................................ 236
Key findings and future work ............................................................................. 238
Implications and future work ............................................................................ 243

Bibliography ...................................................................................................... 246
Appendices ........................................................................................................................................... 264
  Appendix 1. Co-precipitation of 6His-TsUCH37 with T. spiralis lysate ........................................... 264
  Appendix 2. Bioinformatics prediction of T. spiralis secreted proteins ................................................. 265
  Appendix 3. Proteomic analysis of anti-NEDD4-L reactive T. spiralis SP ........................................... 268
  Appendix 4. Analysis of the effect of MG132 on TsUBE2L3-HA expression in C2C12 myotubes ......................... 269
  Appendix 5. Analysis of UBE2L3 expression in C2C12 myotube empty vector and TsUBE2L3-HA cell lines ................................................................. 270
  Appendix 6. Paraffin sections of T. spiralis infected rat muscle tissue ................................................. 271
  Appendix 7. Structural analysis of the catalytic domain of TsUCH37 ................................................... 272

List of Figures

Chapter 1

  Figure 1. 1 The ubiquitin pathway ........................................................................................................... 19
  Figure 1. 2 Different ubiquitin tags ........................................................................................................ 21
  Figure 1. 3 The structure of the 26S proteasome ...................................................................................... 23
  Figure 1. 4 The mechanism of DUB-catalysed hydrolysis ..................................................................... 25
  Figure 1. 5 T. spiralis infection of skeletal muscle tissues ....................................................................... 36
  Figure 1. 6 The life cycle of T. spiralis .................................................................................................... 38
  Figure 1. 7 Schematic of myogenesis .................................................................................................... 41
  Figure 1. 8 The formation of the T. spiralis nurse cell ......................................................................... 44
  Figure 1. 9 The T. spiralis nurse cell complex ....................................................................................... 45
  Figure 1. 10 The morphology of a T. spiralis muscle larva .................................................................... 49

Chapter 3

  Figure 3. 1 Alignment of orthologues and gene predictions of TsUCH37 ........................................... 81
  Figure 3. 2 Gene predictions for TsUCH37 ........................................................................................... 82
  Figure 3. 3 RNA Ligase Mediated (RLM) RACE-PCR schematic ....................................................... 83
  Figure 3. 4 RACE-PCR of the TsUCH37 gene ...................................................................................... 87
  Figure 3. 5 Alignment of orthologues and RACE-PCR-confirmed coding sequence of TsUCH37 ........................................................................................................... 88
  Figure 3. 6 Analysis of the TsUCH37 coding sequence ................................................................. 91
  Figure 3. 7 Expression and purification of TsUCH37 ............................................................................ 92
Figure 3. 8 Ub-AMC hydrolysis by TsUCH37 .................................................................94
Figure 3. 9 Nedd8-AMC hydrolysis by TsUCH37 and orthologues .........................96
Figure 3. 10 Alignment of UCH37 and UCHL3 orthologues .................................100
Figure 3. 11 Model of Nedd8 interaction with PfUCHL3 ................................104
Figure 3. 12 Mutagenesis and Ub/Nedd8 activity profiling of TsUCH37 ...........106
Figure 3. 13 Analysis of TsUCH37 co-precipitation with ADMR1 ......................109
Figure 3. 14 Analysis of TsUCH37 co-precipitation with the *T. spiralis* proteasome113
Figure 3. 15 The effect of WP1130 on TsUCH37 Ub hydrolysis activity ..........115
Figure 3. 16 The effect of LDN-57444 on TsUCH37 Ub hydrolysis activity .........117
Figure 3. 17 The effect of LDN-57444 on *T. spiralis* viability .................................119

**Chapter 4**

Figure 4. 1 The mechanism of the HA-Ub-VME probe ........................................133
Figure 4. 2 Production of the HA-Ub-VME probe ..............................................136
Figure 4. 3 Purification and testing of the HA-Ub-VME probe ..........................137
Figure 4. 4 Reaction of *T. spiralis* secreted proteins with the HA-Ub-VME probe ...140
Figure 4. 5 Proteomic analysis of the HA-Ub-VME reaction with *T. spiralis* secreted proteins ........................................................................................................141
Figure 4. 6 Anti-ubiquitin immuno-blot of *T. spiralis* proteins ......................144
Figure 4. 7 Analysis of TsUBE2N .............................................................................148
Figure 4. 8 Testing for Ub and Nedd8 hydrolysis activity of *T. spiralis* secreted proteins ........................................................................................................150
Figure 4. 9 Production, purification and testing of the HA-Ub TF$_3$BOK probe ......153
Figure 4. 10 Production, purification and testing of the HA-Ub-Lac probe .........154
Figure 4. 11 *In vitro* ubiquitin conjugation assay of *T. spiralis* secreted proteins ....157
Figure 4. 12 Analysis of the secretion of NEDD4-L and WDR48 by *T. spiralis* ....165
Figure 4. 13 Proteomic analysis of *T. spiralis* secreted proteins .........................168
Figure 4. 14 The degradation of warheads used to make Ub-based inhibitor probes ........................................................................................................174
Chapter 5

Figure 5. 1 Analysis of the ubiquitin-related enzymes identified by the proteomic analysis of *T. spiralis* secreted proteins.................................................................186

Figure 5. 2 RACE-PCR analysis of the 5’ end of *T. spiralis* UBE2L3 cDNA...........190

Figure 5. 3 RACE-PCR analysis of the 3’ end of *T. spiralis* UBE2L3 cDNA ...........192

Figure 5. 4 Culture and differentiation of the C2C12 myoblast cell line ..................194

Figure 5. 5 Cloning and expression of TsUBE2L3 in C2C12 myoblasts using a lentiviral system.............................................................................................................197

Figure 5. 6 Cloning and expression of TsUBE2L3 in C2C12 myoblasts using an inducible lentiviral system .........................................................................................199

Figure 5. 7 Immuno-fluorescence analysis of inducible TsUBE2L3-HA expression in C2C12 myoblast cell lines.................................................................200

Figure 5. 8 Immuno-fluorescence analysis of inducible TsUBE2L3-HA expression in C2C12 myotube cell lines ...............................................................202

Figure 5. 9 A Analysis of the effect of TsUBE2L3-HA on C2C12 morphology...........204

Figure 5. 10 Analysis of the subcellular localisation of TsUBE2L3-HA in C2C12 myoblasts and myotubes .................................................................................206

Figure 5. 11 Analysis of the effect of TsUBE2L3-HA on muscle differentiation markers in C2C12 myoblasts and myotubes .....................................................209

Figure 5. 12 Analysis of the effect of TsUBE2L3-HA on the rate of protein synthesis in C2C12 myotubes .................................................................................211

Figure 5. 13 Analysis of the effect of TsUBE2L3-HA on cell cycle factors in C2C12 myoblasts and myotubes ...............................................................213

Figure 5. 14 A. Immuno-fluorescence analysis of the effect of TsUBE2L3-HA on DNA replication in C2C12 myoblasts and myotubes.................................217

Figure 5. 15 Flow cytometry analysis of the effect of TsUBE3L3-HA on DNA replication in C2C12 myoblasts .................................................................................220

Figure 5. 16 Analysis of the effect of TsUBE2L3-HA on p53 in C2C12 myoblasts and myotubes ..............................................................224

Figure 5. 17 Analysis of the effect of TsUBE2L3-HA on the transcription of p53 in C2C12 myotubes .................................................................................227
List of Tables

Chapter 2

Table 2.1 DNA mixture for transfection for lentivirus production ..............................55
Table 2.2 Table of SDS-PAGE gel reagent composition ........................................72
Table 2.3 Antibodies and probes used throughout the study ..................................73
Table 2.4 Plasmid vectors used throughout the study ...........................................74

Chapter 3

Table 3.1 Ub and Nedd8 activity profiles of UCH37 and UCHL3 orthologues ..........98
Table 3.2 UCH37 and UCHL3 residues that cluster with Ub/Nedd8 activity profile..99
Table 3.3 UCH37 and UCHL3 residues predicted to be important for Ub contact .102
Table 3.4 UCH37 and UCHL3 residues predicted to be important for Nedd8 contact ..........................................................................................................................103

Chapter 4

Table 4.1 Pathogen-derived ubiquitin pathway enzymes .......................................130
Table 4.2 Ubiquitin-related proteins identified by proteomic analysis of the HA-Ub-VME reaction with T. spiralis proteins .................................................................143
Table 4.3 Bioinformatics-based prediction of T. spiralis secreted proteins ..........159
Table 4.4 Ubiquitin-related T. spiralis proteins predicted to be secreted by wolfPSORT ..................................................................................................................163
Table 4.5 Antibodies to ubiquitin-related proteins predicted to be secreted by T. spiralis ................................................................................................................164
Table 4.6 Ubiquitin-related proteins identified by the proteomic analysis of T. spiralis secreted proteins ................................................................................169
List of appendices

Appendix 1. Co-precipitation of 6His-TsUCH37 with T. spiralis lysate ..................264
Appendix 2. Bioinformatics prediction of T. spiralis secreted proteins ..................267
Appendix 3. Proteomic analysis of anti-NEDD4-L reactive T. spiralis SP ..............268
Appendix 4. Analysis of the effect of MG132 on TsUBE2L3-HA expression in C2C12 myotubes ..........................................................269
Appendix 5. Analysis of UBE2L3 expression in C2C12 myotube empty vector and TsUBE2L3-HA cell lines .................................................................270
Appendix 6. Paraffin sections of T. spiralis infected rat muscle tissue ..............271
Appendix 7. Structural analysis of the catalytic domain of TsUCH37 ..............272
Chapter 1: Introduction

The purpose of the PhD project was to investigate the role of the ubiquitin pathway in host-parasite interactions. This investigation was carried out using *Trichinella spiralis* infection of mammalian muscle cells as a model system.

1.1. The ubiquitin pathway

The ubiquitin pathway is functional in all eukaryotic organisms and, as the name suggests, the focal point of the pathway is a protein that is expressed ubiquitously in all cell types, ubiquitin. It is a highly conserved biochemical pathway that modifies and regulates proteins at the post-translational level. The ubiquitin pathway was first discovered by Aaron Ciechanover, Avram Hershko, and Irwin Rose in the 1980s, who went on to characterise the role of the ubiquitin pathway in the regulation of protein degradation [4]. This major function of the ubiquitin pathway is to regulate the abundance of intracellular proteins by controlling their degradation via the direct communication with the proteolytic complex, the 26S proteasome. We now know that the ubiquitin pathway plays a vital role in the regulation of a great number of tightly regulated biological processes, from transcription and the cell cycle to trafficking and apoptosis [5]. As a testament to the importance of this pathway in cell biology, the work of Ciechanover, Hershko and Rose was awarded the Nobel Prize in Chemistry in 2004.

Ubiquitin (Ub) is a 76 amino acid protein (8.5 kDa) that is known as a small modifier [6]. This is because protein substrates of the ubiquitin pathway become modified with (or conjugated to) a ubiquitin ‘tag’. This tag serves as a signal that directs the substrate protein towards a particular fate depending on the timely requirements of the cell during the maintenance of healthy homeostasis. Other small modifier proteins similar to Ub have also been discovered and this thesis will discuss one other, Nedd8 (Neural Precursor Cell Expressed, Developmentally Down-Regulated 8). There is no evidence that Nedd8 is directly involved in protein degradation. It does however play an important role in processes such as the regulation of the cell cycle [7]. In humans Nedd8 is 53.4% identical in amino acid sequence and shares a common tertiary structure to ubiquitin [8]. Nedd8 and ubiquitin are processed in a similar manner, resulting in the Neddylation or ubiquitination of substrate proteins.
Chapter 1: Introduction

1.2. Ubiquitin conjugation

Ubiquitin is transcribed as a polyubiquitin precursor. The first step in the process of ubiquitination is therefore the proteolytic cleavage of polyubiquitin to form monoubiquitin. This is now ready to be conjugated to substrate proteins. The event of conjugation is catalysed by 3 ubiquitin-specific enzymes that function with one-another in a cascade of activity [9] (Figure 1.1). The first enzyme is a ubiquitin activating enzyme, called E1. This 110 kDa enzyme forms an E1-S-ubiquitin intermediate complex upon the formation of an ATP-dependent thioester linkage between the sulfhydryl group of the active site cysteine of the E1 and the C-terminal glycine carboxyl group of the ubiquitin. The E1 enzyme then passes the highly activated thiol-linked ubiquitin onto the active site cysteine of a ubiquitin conjugating enzyme (E2), forming an E2-S-ubiquitin intermediate. The E2 enzyme then specifically recognises the final protein in the enzymatic cascade, the ubiquitin ligase (E3). The specific E3, and sometimes the E2-E3 pair, is responsible for recognising a specific protein substrate. The E3 then mediates the ligation of the ubiquitin to the amino group of a lysine (K) residue of the protein substrate via a covalent isopeptide bond with the C-terminal glycine of the ubiquitin (Figure 1.1).
Figure 1.1 The ubiquitin pathway

Schematic showing the cleavage of the ubiquitin (Ub, lilac) precursor by deubiquitinating enzymes (DUBs). Followed by ubiquitin conjugation to a substrate protein (S) as catalysed by the E1, E2 and E3 enzyme cascade. The active-site cysteine of the E1 enzyme activates Ub by forming an ATP-dependent E1-S-Ub intermediate with the C-terminal glycine of Ub. Activated Ub is then passed onto the active-site cysteine of an E2 enzyme, which also forms a thioester bond with the C-terminal glycine of Ub. The E2 passes Ub onto an E3 ligase. RING E3 ligases bring the E2-S-Ub intermediate into contact with the substrate protein, allowing the substrate lysine amino group to attack the C-terminal glycine of Ub forming an isopeptide bond. E2 enzymes pass Ub onto the active site cysteine of HECT E3 ligases first, and then the E3 interacts with the substrate protein, which forms an isopeptide bond with Ub. Ub conjugation is reversed by DUBs to recycle ubiquitin and deubiquitinate the substrate protein. DUBs attack the C-terminal glycine of Ub to hydrolyse the isopeptide bond. Figure modified from Kerscher et al. 2006, Annu. Rev. Cell. Dev. Biol. [7]
Chapter 1: Introduction

Eukaryotes express only one E1 enzyme but express a large number of E2s (10s in humans) that contain the ubiquitin conjugating domain (UBCc) and an even larger number of E3s (100s in humans) [10]. There are a number of different types of E3 domains, the best-described being the HECT domains (Homologous with the E6-AP C-Terminus) and RING (Really Interesting New Gene) finger domains [11,12]. During ligation, ubiquitin is transferred from an E2 to the active-site cysteine of a HECT domain E3, forming an E3-S-ubiquitin intermediate before the ligation to the substrate takes place. In contrast, RING finger E3s catalyse the transfer of ubiquitin directly from the E2 to the protein substrate, thus associating with the substrate and the E2 simultaneously (Figure 1.1). Together, the specific E2-E3 partnership and the recognition of particular protein substrates by the E3, confer an extremely high level of specificity in the pathway.

In addition to lysine residues of substrate proteins, ubiquitin can be ligated to 7 possible lysine residues (K6, K11, K27, K29, K33, K48 and K63) on other ubiquitin proteins, forming chains of polyubiquitin [13]. These polyubiquitin chains are also found conjugated to substrate proteins. There is a large variation in the possible length of the chain, and the residue via which each ubiquitin is linked to another, providing many possible different ubiquitin tags (Figure 1.2). Proteins can also be monoubiquitinatated at multiple lysine residues (multiubiquitinated) and polyubiquitin chains can also be branched. Each particular ubiquitin tag serves as a signal to direct the substrate protein towards a different fate. For example, monoubiquitin often leads to changes in the sub-cellular localisation of the substrate protein, multiubiquitin has been shown to affect protein interactions and associations, whereas K48-linked polyubiquitin chains generally target proteins to the 26S proteasome for their degradation [7,14,15].
Figure 1.2 Different ubiquitin tags

Ubiquitin has 7 lysine residues that can be linked to other ubiquitin modifiers. Each type of ubiquitin tag serves as a signal for a different biological fate of the substrate protein. Figure modified from Wang et al. 2012, Front. Oncol. [16].
1.3. The 26S proteasome

The 26S proteasome is a protein complex located in the cytoplasm and the nucleus. It is composed of 2 multi-protein subunits, the 20S catalytic core and the 19S regulatory particle. The assembly of the 20S and the 19S subunits is reversible. Disassembly is induced by various stresses such as oxidative stress or starvation [17,18]. The 19S regulatory particle can associate with either or both ends of the 20S core and is composed of a lid structure and a base structure (Figure 1.3). The 19S lid comprises proteins Rpn3, 5-9, 11, 12 15 plus Rpn10, Rpn1, Rpn2 and Rpn13, which is also known as ADRM1. These proteins are responsible for the recognition and binding of the polyubiquitin tag of a substrate protein. Once the substrate protein is associated with the 19S lid, it begins to enter the 19S base, which links the 20S core to the 19S lid. This is composed of a ring of Rpt proteins (1-6). These proteins are ATPases that possess chaperone-like activity and are thought to be responsible for unfolding the substrate protein, allowing its passage into the 20S core [19]. The 20S subunit is cylindrical and is made up of alpha and beta subunits arranged in hexameric ring structures forming a hollow core. These proteins possess protease activity (caspase-like, chymotrypsin-like and trypsin-like) and catalyse the degradation of the unfolded protein as it passes through the proteolytic core [20]. The substrate protein is processed into short peptides and amino acids that pass out of the other end of the proteasome. Ubiquitin does not pass through the proteasome. This is cleaved from the substrate upon binding and entry of the substrate to the 19S regulatory particle. The cleavage and subsequent recycling of ubiquitin from protein substrates is mediated by proteasome-associated ubiquitin hydrolase (or deubiquitinating-DUB) enzymes.
Chapter 1: Introduction

Figure 1.3 The structure of the 26S proteasome

The 26S proteasome is composed of a 19S subunit, also known as the regulatory particle, and the 20S subunit, also known as the catalytic core. The 19S subunit is made up of a lid of Rpn proteins and a base Rpt protein ring. This regulates substrate protein association, deubiquitination and unfolding. The 20S catalytic core is made up of alpha and beta subunits arranged in hexameric rings forming a hollow core through which the unfolded protein passes. The 20S core possesses ATP-dependent protease activity that degrades the protein, releasing short peptides and amino acids. Image modified from Weissman et al. Nat. Rev. Mol. Cell. Biol. 2011 [21].
1.4. Ubiquitin hydrolysis by DUBs

Enzymes responsible for the cleavage of the ubiquitin precursor, the recycling of monoubiquitin from un-associated polyubiquitin chains and the removal of ubiquitin from substrate proteins are called deubiquitinating enzymes (DUBs) [22]. The enzymes that hydrolyse Nedd8, deNeddlating enzymes, function in a similar manner. It is thought that there may be in excess of 100 different mammalian DUBs, each of which usually recognises highly specific protein substrates. Many have been characterised and for some the crystal structures have been resolved [23-25]. DUB enzymes can be divided into 5 families based on their functional conserved domains: the Ub C-terminal hydrolase domain (UCH), Ub specific protease domain (USP), ovarian tumour domain (OTU), Machado-Josephin domain (MJD) and the JAB1/MPN/Mov34 domain (JAMM). The UCH, USP, OTU, and MJD enzymes are cysteine proteases and the JAMM family comprises metalloprotease enzymes [26]. The catalytic domains of the cysteine protease DUBs contain a highly conserved active site consisting of 4 amino acid residues essential for the cleavage of a specific substrate. For example, the catalytic site of the UCH domain enzymes comprise a glutamine, cysteine, a positively charged histidine and a negatively charged aspartic acid [25]. During ubiquitin hydrolysis, the active site cysteine thiol (sulfhydryl group) forms a thioester bond with the carbonyl of the C-terminal carbon of ubiquitin. The isopeptide bond between ubiquitin and the substrate is subsequently hydrolysed, releasing ubiquitin, the deubiquitinated substrate and the regenerated enzyme (Figure 1.4).
**Figure 1.4 The mechanism of DUB-catalysed hydrolysis**

A DUB (beige) recognises a specific protein substrate (dark green) and catalyses the hydrolysis of the isopeptide bond between the substrate and ubiquitin. The active-site cleft contains a histidine (His), an asparagine (Asn), an aspartic acid (Asp) and the catalytic cysteine (Cys). The histidine maintains the nucleophilicity of the cysteine, which attacks and hydrolysates the isopeptide bond of ubiquitinated substrate. Kessler, 2006, [27].
In humans there are 3 DUBs shown to specifically interact with the proteasome. The first, Rpn11 is an integral part of the 19S lid structure of the regulatory particle. Rpn11 is a JAMM domain metalloprotease that removes ubiquitin from protein substrates as they pass into the proteasome. The removal of ubiquitin, which otherwise sterically hinders the entry of the substrate into the proteasome, promotes the translocation of the substrate into the 20S core for degradation [28,29]. The second proteasome-associated DUB is the ubiquitin-specific protease domain enzyme, USP14 (Ubp6 in yeast). Unlike Rpn11, this protein is not always associated with the proteasome and association is reversible. USP14 associates with the 19S lid via Rpn1 and Rpn2 and catalyses the cleavage of ubiquitin from proteins as they bind to the proteasome [30]. This inhibits the activity of the regulatory particle, suggesting that the cleavage of the tag frees the substrate from the original fate of proteasomal degradation [31]. The final proteasome-associated DUB, whose function is a focal point of this thesis, is the ubiquitin-C-terminal hydrolase domain enzyme, UCH37 (or UCH-L5). UCH37 is nuclear and cytoplasmic and its interaction with the proteasome is also reversible. The binding of UCH37 to the proteasome significantly enhances its ubiquitin hydrolysis activity by relieving UCH37 of the auto-inhibitory effect of its C-terminal tail [32,33]. UCH37 binds to Rpn13/ADRM1, which in turn binds to Rpn2, thus competing with USP14 for proteasome association. UCH37 cleaves ubiquitin from the distal end of polyubiquitin chains of protein substrates [34,35]. Koulich et al. showed that RNAi of UCH37 speeds up protein degradation in HeLa cells suggesting that UCH37 also acts as a checkpoint for the fate of the substrate protein. Interestingly, UCH37 knockout in mice caused embryonic defects and prenatal death illustrating that the regulation of proteasomal protein degradation by UCH37 is essential [36].

In summary, ubiquitin conjugation is highly regulated by the balance of the abundance and activity of ubiquitin conjugation machinery; E1, E2 and E3 enzymes. Likewise ubiquitin hydrolysis is highly regulated by the balance of the abundance and activity of DUBs. Although in mammals many of these proteins have now been characterised, in almost all cases their specific protein substrates remain to be determined. The transient nature of the interaction of these enzymes with their substrates makes this a challenging task. However in order to further characterise the specific pathways and biological processes in which they function, this is an important angle of investigation.
1.5. The ubiquitin pathway in disease

Considering the important role of the ubiquitin pathway in maintaining the healthy homeostasis of a cell, and therefore of the healthy physiology of the organism, it is not surprising that the disruption of the ubiquitin pathway is directly implicated in a variety of diseases. These include cystic fibrosis, Liddle’s syndrome, Angelman syndrome, muscle wasting disease and many neurodegenerative disorders and cancers [37-40].

1.6. Muscle wasting disease

The terminal differentiation of skeletal muscle cells, from immature myoblasts to mature muscle fibres, will be discussed in more detail later (section 1.14). The processes of muscle differentiation, growth and homeostasis require the function of the ubiquitin pathway [41]. The ubiquitin pathway is also implicated in the disruption of these processes that can lead to muscle wasting diseases caused by skeletal muscle atrophy (the degeneration of skeletal muscle tissues). This can occur as a result of other diseases such as cancer, chronic obstructive pulmonary disease (COPD) and sepsis. During these disease states, the ubiquitin system in muscle cells is highly activated by a lack of physical activity, a loss of nervous communication to the muscle tissues and by malnutrition [40]. This leads to the rapid ubiquitin-mediated degradation of muscle proteins and thus the loss of the mass and function of the muscle tissue. This effect has been linked to the up-regulation of 2 muscle-specific atrophy-related (atrogen) genes that encode the MuRF1 (Muscle RING finger 1) protein that contains an E3 RING finger domain, and the MAFbx (Muscle Atrophy F-box) protein that contains an E3 RING finger domain similar to that of the cell-cycle specific E3, the SCF complex [42]. These muscle specific E3 ligase enzymes ubiquitinate and target myofibre proteins for proteasomal degradation [43,44]. MuRF1 and MAFbx have been shown to be activated in many different animal models of muscle wasting disease, suggesting that a common ubiquitin-related pathway controls muscle atrophy in various disorders.

1.7. Neurodegenerative disease

Neurodegenerative diseases such as Parkinson’s disease, Alzheimer’s disease and Huntington’s disease all involve the build up of toxic aggregates of abnormal proteins in neuronal cells. As well as regulating the degradation of specific proteins for
specific requirements, the ubiquitin-proteasome system is also responsible for the
degradation of misfolded or abnormal proteins [45]. It is thought that the disruption of
this process may lead to their build up and play a role in the formation of these
aggregates [46]. Not only have these aggregates have been found to contain
ubiquitin but mutations in various genes encoding ubiquitin pathway proteins have
been linked to the development of these diseases [47-49]. For example mutations in
the PARK2 gene that encodes the E3 ligase parkin, is found in autosomal recessive
juvenile Parkinsonism, and mutations in the ubiquitin gene itself are associated with
Alzheimer’s disease [50] [51].

1.8. Cancer

The regulation of the abundance or location of a cell cycle protein or an oncoprotein
often relies on proteasomal degradation [52,53]. The disruption of this process can
therefore lead to a disruption of the cell cycle and tumourigenesis. These disruptions
can arise from within the cell, for example by mutagenesis of cell cycle related genes.
Degradation of the G1 phase-specific cyclin-dependent kinase inhibitor, p27\textsuperscript{kip1} is
highly regulated by ubiquitination. The increased proteasomal degradation and
therefore down-regulation of this protein leads to cell cycle progression. The down-
regulation of p27\textsuperscript{kip1} has been found in many cancers such as breast and prostate,
implicating a role for the ubiquitin pathway in the development of these malignancies
[54]. In addition, the cell cycle regulator proteins, cyclins, are also highly regulated by
the ubiquitin pathway [55,56]. The over-expression of cyclin B, cyclin E and cyclin D1
has been observed in many different types of tumours [57-59]. Whether or not the
up-regulation of these proteins in tumour cells is a result of the inhibition of their
proteasomal degradation remains to be determined, but mutations in cyclin-specific
ubiquitin conjugation machinery have been found in several cancers [60]. SCF (Skp,
Cullin, F-box containing) is a cell cycle-specific complex that controls the transition of
the G1/S phase and the G2/M phase of the cell cycle. This is achieved through the
E3 ligase activity of SCF which ubiquitinates specific cell cycle factors. Similarly, the
anaphase promoting complex (APC) is an E3 ligase that ubiquitinates specific cell
cycle factors to initiate the progression of metaphase and anaphase of the M phase
of the cell cycle [61,62]. The regulation of cell cycle factor ubiquitination by SCF and
APC is therefore important for maintaining a healthy level of proliferation, and
alterations in the genes encoding these complexes have been found to correlate with
a number of cancers [63].
Cancer development may also be initiated by the disruption of the ubiquitin pathway and subsequent disruption of the cell cycle by viruses. High-risk strains of the human papillomavirus (HPV) express an E6 oncoprotein that associates with the human HECT domain E3, E6-AP (E6 Associated Protein, also known as UBE3A) in infected cervical epithelial cells [64]. Through an interaction with the host E2, UBE2L3 (UbcH7), this leads to the specific ubiquitination and subsequent degradation of the tumour suppressor protein p53 [65-67]. The suppression of p53 promotes the propagation of the virus via the uncontrolled proliferation of these cells, leading to the development of cervical cancer [68,69]. This is also an example of how the disruption of the ubiquitin pathway is implicated in infectious disease, which will be discussed in more detail in the next section.

Due to the discovery of the direct link between the ubiquitin pathway and disease, the proteasome has been recognised as a potential drug target for disease therapy for some time [70-72]. Some proteasome inhibitors are in clinical trials for human use including bortezomib, which in 2006 was approved by the Food and Drug Administration to treat multiple myeloma and mantle cell lymphoma [73-76]. However, because the proteasome is so highly expressed in all tissues, many proteasome inhibitors are non-selective and show toxicity. In order to develop inhibitors with greater selectivity, there has been an expansion in the interest in targeting other members of the ubiquitin pathway that are differentially expressed or activated, such as ubiquitin conjugation enzymes or DUBs [77]. Although as yet none have entered clinical trials, there are some recent examples of small molecule DUB inhibitors that show therapeutic benefits in cancer [78]. For example, Chauhan et al. showed that a specific inhibitor of USP7/HAUSP (herpesvirus-associated USP) induces apoptosis of multiple myeloma cells and inhibits tumour growth in mice [79]. Other DUBs, such as the proteasome-associated UCH37, have been directly linked to the deubiquitination and stabilisation of oncogenic proteins leading to the proliferation of cancer cells. UCH37 stabilises transforming growth factor-β (TGF-β) by targeting Smad transcription factors for proteasomal degradation via ubiquitination. Inhibiting UCH37, reverses this effect, and inhibits cancer cell migration [80,81]. In addition, a drug that specifically inhibits UCH37 was observed to induce apoptosis of a number of different types of tumour cells in mice [82]. Amongst other ubiquitin pathway enzymes, these DUBs may therefore prove as promising new targets for the development therapeutics for cancer, and ubiquitin-related diseases.
1.9. The ubiquitin pathway in infectious disease

In addition to the disorders discussed, the ubiquitin pathway is also implicated in infectious disease. Immunity to infection is highly dependent on the function of antigen presentation as carried out by antigen presenting cells (APCs, predominantly dendritic cells and macrophages, but also B lymphocytes). This is the process by which pathogen-derived antigens are degraded into short peptides and presented on the surface of APCs in complex with the major histocompatibility complex, MHC. These APCs then present the antigen-associated MHC complexes to T lymphocyte cells. T cells become activated to respond appropriately to the infection, either by killing the pathogen or mounting an adaptive immune response. The ubiquitin-proteasome system is involved in antigen uptake by B lymphocyte cells, proteasomal-mediated antigen degradation for peptide presentation by APCs, and the proteasomal regulation of levels of MHC protein by proteasomal degradation [83,84]. A functional ubiquitin-proteasome system is therefore crucial for managing immunity to infection.

Many pathogens specifically target the ubiquitin pathway in order to evade an immune response that threatens their survival [85,86]. For example, the lymphoma associated Epstein Barr virus (EBV) expresses a protein EBVNA1 (EBV Nuclear Antigen 1) that evades proteasomal degradation thus evading antigen presentation and a subsequent immune response. Similarly, the cytomegalovirus (CMV) proteins US2 and US11 evade antigenic presentation by preventing the translocation of the MHC complex to the cell surface. They do so by directing the MHC complex to the proteasome for degradation [87]. The disruption of antigen presentation during infection is not exclusive to viruses. Pathogenic bacteria have also developed mechanisms of manipulating ubiquitin pathway-mediated regulation of antigen presentation to evade immunity. For example Salmonella bacteria inhibit the translocation of MHC to the surface of APCs thus preventing the activation of a T cell mediated response [88]. The modulation of the host ubiquitin pathway by pathogens is not restricted to antigen presentation. Many pathogens target the ubiquitin pathway in order to enhance their virulence by disrupting other host cell functions such as the cell cycle. This can involve the specific targeting of deubiquitinating enzymes by pathogens [89]. For example, the Epstein Barr virus EBNA1 protein directly interacts with and inhibits the human (tumour suppressor) p53-specific DUB, USP7/HAUSP. This leads to the destabilisation of p53 and the disruption of the host cell cycle [90,91]. The widespread bacterial pathogen, Pseudomonas aeruginosa, chronically
infects the lung mucosal cells of patients with diseases such as COPD, cystic fibrosis and pneumonia. *P. aeruginosa* secretes an effector protein Cif, which targets and inhibits the host DUB USP10. This inhibits the removal of ubiquitin by USP10 from the cystic fibrosis transmembrane conductance regulator (CFTR) causing an increase in the proteasomal degradation of the CFTR. CFTR is an ion channel that transports chloride across the cell membrane. This internalization and degradation of CFTR inhibits the transport of chloride across the cell membranes of infected mucosal cells, causing a build up of thick mucous outside the cell. This prevents the clearance of the pathogen by the mucosal cell surface cilia, thus aggravating the infection [92].

Interestingly, some of the viral and bacterial proteins that hijack the host ubiquitin pathway are ubiquitin pathway enzymes themselves. These proteins can be completely pathogen specific (novel), or they can mimic host ubiquitin enzymes. This was a surprising discovery, since prokaryotes, viruses and bacteria, do not have a functional ubiquitin proteasome pathway of their own. These proteins have therefore evolved to act like host ubiquitin enzymes to specifically target the host ubiquitin pathway to their advantage. An example of a novel pathogen ubiquitin enzyme is the herpes simplex virus 1 (HSV-1) deubiquitinating enzyme encoded by the UL36 gene. This protein shows little homology to mammalian DUBs but can hydrolyse human ubiquitin in vitro [93]. Homologues of this protein are also expressed by murine CMV (cytomegalovirus) and gammaherpesvirus 68 [94,95]. Although little is known about the function of these UL36 homologues, for other pathogen-derived DUBs, their effects on host cells have been characterised. The pathogenic bacterium *Chlamydia trachomatis*, that causes one of the most common sexually transmitted diseases in the UK, expresses 2 DUBs, ChlaDUB1 and ChlaDUB2. Both can hydrolyse human ubiquitin and human Nedd8 in vitro. ChlaDUB1 has been shown to inhibit ubiquitin-mediated degradation of IκBα leading to a down-regulation of the pro-inflammatory protein, NFκB. A function that may lead to the suppression of an immune response required to clear the pathogen [96]. The mechanisms that underlie the effects of the ChlaDUBs are yet to be determined, however for other pathogen-derived ubiquitin enzymes, the direct interaction with the host ubiquitin pathway has been demonstrated. For example, *Salmonella* express an effector protein that mimics host HECT E3 ligases. This E3 enzyme, SopA, specifically collaborates with the host ubiquitin conjugation (E2) enzymes UbcH5 and UBE2L3 (UbcH7) in HeLa cells [97]. It is thought that this also leads to the ubiquitination and subsequent degradation of proteins involved in infection-induced inflammation. *Salmonella* also express a
deubiquitinating enzyme, AvrA, that is able to deubiquitinate host IκBα and β-catenin, leading to an inhibition of the host inflammatory response to infection, a common target of immune-evasive pathogens [98].

Targeting the ubiquitin pathway to treat non-infectious diseases such as cancer has emerged as a viable strategy. Therefore, drugs that target pathogen-derived ubiquitin pathway enzymes may also be useful for combatting infectious diseases. Efforts are being made to characterise for example, pathogen derived DUBs, for this very reason. SARS coronavirus causes severe acute respiratory syndrome. The drug GRL0617 targets the SARS coronavirus DUB, Papain-like protease (Plpro). This inhibits virus replication, thus illustrating the potential of DUB inhibitors in the treatment of infection [99-101].

This leads us to the first point in question: If there is potential in viral and bacterial ubiquitin pathway components as drug targets in infection, is there also potential for the development of parasite therapeutics that target the parasite ubiquitin pathway during infection?

1.10. The ubiquitin pathway in infection by parasites

Parasites are most often eukaryotes and therefore express their own ubiquitin pathway. Considering the importance of this pathway in eukaryotic cell biology, and its conservation throughout evolution, the first hypothesis was proposed that the ubiquitin pathway at least plays a role in host-parasite interactions during infection by parasites.

Ubiquitin pathway enzymes have now been characterised in various parasites. For example DUBs are expressed by the human apicomplexan parasites P. falciparum and T. gondii and the mammalian parasitic worm Trichinella spiralis [1,102-104]. These DUBs, UCHL3 orthologues in T. gondii and P. falciparum and UCH37 orthologues in P. falciparum and T. spiralis, were discovered using reagents based on human ubiquitin. Due to the high level of conservation of ubiquitin, even between mammals and parasites, it is unsurprising and yet interesting that these parasite enzymes can hydrolyse human ubiquitin. Furthermore, drugs that specifically target the human ubiquitin system, the majority of which are proteasome inhibitors, have also been shown to kill parasites. For example, proteasome inhibitors have been shown to kill the human malaria parasite Plasmodium falciparum and the zoonotic parasite, Toxoplasma gondii [105-110]. Studies have shown that P. falciparum
express a functional proteasome that is essential for the survival of both the infective and transmission stages of the parasite (reviewed by Aminake et al. [111]). Due to the demand of new drug targets for the treatment of human malaria, and the importance of the proteasome in \textit{P. falciparum} cell biology, the ubiquitin pathway of \textit{P. falciparum} is now being considered by as a potential drug target. To this end, efforts are being made to further characterise the \textit{P. falciparum} ubiquitin proteasome system [112-114].

Although this shows promise in targeting parasite ubiquitin pathways to kill parasites, in the interest of developing more specific drugs, and vaccines to prevent infection, the main perpetrators of host-parasite interactions during infection must be identified. This prompted the second question: in addition to viruses and bacteria, can parasites specifically target the host ubiquitin pathway to their advantage during infection? To date, no parasite ubiquitin pathway enzymes have been shown to be specifically host targeted. This is not necessarily because they do not exist, but due to the fact that separating the proteins that are specifically host-targeted from those that are parasite’s ‘own’ is a major challenge. This is especially difficult for intracellular parasites such as \textit{P. falciparum}. For example, the blood stage \textit{P. falciparum}, a major cause of the pathogenesis of malaria, cannot be cultured without the host blood cells. It is therefore extremely difficult to differentiate host proteins from parasite proteins and parasite-derived host-targeted proteins.

Parasitic nematode worms (also known as helminths) however, provide somewhat of a solution to this problem.

\textbf{1.11. Parasitic worms and their secreted proteins}

This solution lies within the study of the secretions of parasitic worms. Parasite-derived secretions are often referred to as the excretory-secretory fraction (ES). Parasite ES is released into the host during most life cycle stages of all helminths. Helminth ES contains a mixture of proteins, lipids and polysaccharides [115]. The term ‘ES’ does not distinguish between molecules that are actively secreted and those might be released as a consequence of biological processes such as digestion, egg laying or moulting. Because this study focuses on the actively secreted protein component of ES, this thesis continues to refer to the ‘secreted proteins’ (SP) of helminths. The composition of the helminth SP adapts to the changing external environment and the developments in the life cycle stage of the parasite. Along with the parasite surface proteins, the secreted proteins are the main mediators of host-
parasite interaction, thus communicating with the host in an attempt to ensure optimal immune evasion and survival.

For some helminths, characterising the secreted proteins and the host-parasite interactions in which they are involved is a major challenge, as the natural life cycle of the parasite cannot be successfully modelled in vitro. Studies on these parasites therefore rely on mouse models and ex vivo studies. In some cases, it is impossible to access certain life cycle stages of the parasite whereas other stages prove accessible [2]. Pure SP can only be collected from parasites that have been removed from their natural niche and cultured in optimised conditions. Therefore parasites in culture exist in a suspended phase of development void of the biochemical cues that would be present in the natural environment. In this case studying the SP is highly informative, but not necessarily determinative of the proteins that are secreted during natural infection. One such example is the food borne parasite Trichinella spiralis. Adult T. spiralis parasites can be extracted from the intestine of a host, as can L1 larvae that have been enzymatically digested out of the muscle tissue of a host animal. The muscle larvae can be cultured for 4-6 days. Once in culture, the secreted proteins can be collected from the culture supernatants without host protein contamination, thus allowing the study of actively secreted and presumably host-targeted proteins [116]. Furthermore, T. spiralis is an interesting helminth in the context of the role of the ubiquitin pathway in host-pathogen interactions. T. spiralis is unique in that during the chronic phase of infection the parasite exists inside skeletal muscle cells. Proteins expressed by the muscle larvae are secreted directly into the host cell. This does not kill the cell, rather mediates the complete re-programming and transformation of the cell cycle and differentiation state of the muscle cell. Because T. spiralis muscle larvae are intracellular, they have direct access to the host ubiquitin pathway and many of the host cell processes that are disrupted during this stage ordinarily rely on the ubiquitin pathway for their regulation (discussed in more detail in sections 1.14 and 1.15). Furthermore, T. spiralis infects all mammals indiscriminately, passing through the same life cycle stages and eliciting similar pathogenicity. The infection is therefore analogous in humans and rodents, allowing us to study mouse and rat models with confident extrapolation of what might be happening in the human host.

For these reasons, the infection of skeletal muscle cells by T. spiralis was chosen as a model system for investigating whether or not parasites can deliberately target the host ubiquitin pathway during infection.
Chapter 1: Introduction

**Trichinella spiralis**

1.12. History and epidemiology

*Trichinella* is a promiscuous parasitic roundworm (nematode) that infects mammals, birds and reptiles causing trichinellosis (also known as trichinosis). Species of *Trichinella* are differentiated between those that become encapsulated in host muscle tissue during infection: *T. britovi*, *T. murrelli*, *T. nativa*, *T. nelsoni*, and *T. spiralis*, and those that do not: *T. papuae*, *T. pseudospiralis* and *T. zimbabwensis* [117]. Trichinellosis infection in humans has been found on all continents except Antarctica making it the most widespread helminth in the world [118]. The most common perpetrator of human infection is the species *Trichinella spiralis*. *T. spiralis* was named after the characteristic spiral structure that the worm adopts inside muscle tissue (Figure 1.5). These microscopic worms were first described in 1835 in London by Jim Paget, who observed coiled larvae in the diaphragm of an autopsy subject. The discovery was later reported and claimed by Richard Owen [119].

Infection in the UK and the US was common at the time. Between 1860 and 1877, 150 trichinellosis epidemics were reported [120]. It was later discovered that *Trichinella* is food borne, and advances in farming standards and in food hygiene preparation led to the near eradication of the disease in humans in the UK and the US. However significant outbreaks still occur today. The most recent have affected parts of Eastern Europe, Russia, China and south East Asia [121-123].

In many cases the parasite is transmitted via the consumption of undercooked, cured or raw meat, which is frequently of a wild origin rather than commercially farmed. Muscle stage *Trichinella* can even survive temperatures below freezing, especially in the tissues of arctic mammals that demonstrate resistance to freezing temperatures, such as polar bears [124]. In fact this is one theory for the cause of death of Salomon August Andrée and his artic explorers in 1897. During an attempt to be the first discoverers of the North Pole, the sudden demise of the group is thought to have followed their consumption of raw high-arctic bear meat. The only way to ensure the destruction of *T. spiralis* muscle larvae is to heat the contaminated meat throughout, above 70°C [118]. Although today trichinellosis in humans still contributes to helminth-related morbidity and disability adjusted life years, perhaps more importantly it is considered a serious agricultural problem due to the transmission from wild animals to livestock, and is therefore still considered a significant yet possibly neglected tropical disease [125].
Figure 1.5

Figure 1.5 *T. spiralis* infection of skeletal muscle tissues

Diagram shows *T. spiralis* muscle larvae coiled up inside skeletal muscle tissue, as would have been observed by Jim Paget in 1835. Photograph taken by S. J. Upton of Kansas State University Biology Division.
1.13. The life cycle of *Trichinella spiralis*

Although many other helminths pass through a free-living stage, or have 2 hosts, the transmission of *Trichinella* occurs directly from host to host. The life cycle of encapsulating species such as *T. spiralis* is even more unique in that it exists in both extracellular and chronic intracellular stages in the same host. During the chronic intracellular stage *T. spiralis* become encapsulated in muscle tissues, where they can persist for decades.

Encapsulated *T. spiralis* L1 larvae are ingested in contaminated meat and released from infected muscle tissue following digestion of the meat by gastric enzymes, to which the parasite itself is resistant thanks to its strong outer cuticle (Figure 1.6). L1 larvae invade the epithelia of the small intestine, inducing inflammation and causing symptoms that are similar to those experienced during various intestinal infections such as vomiting and diarrhoea. For this reason the condition is often misdiagnosed [126]. L1 larvae moult through 4 larval stages before developing into adults approximately 30-34 hours after ingestion [127]. Adults reproduce and after 5-7 days females release newborn larvae. After this the adult parasites are cleared from the gut by the host immune system. The newborn larvae cross the intestinal epithelia, entering the lymphatic and vascular circulation, within which they are transported around the body. The number or dose of newborn larvae and their path taken around the body determines the potential for acute illness. Infection of the eyes, heart or CNS may result in oedema, vasculitis, myalgia, myocarditis and encephalitis, which can lead to fatality [122]. At approximately 28 days after ingestion the newborn larvae establish chronic infection in their final target tissue, skeletal muscle. Here they invade fully differentiated skeletal muscle cells (myofibres). A single larva usually invades a single muscle fibre, although in acute infection as many as 4 larvae have been observed to invade a single myofibre. It is thought a terminally differentiated skeletal myofibre may be the only cell in the animal large enough to accommodate the parasite, and hence the specificity of the target tissue. Here, the phenotype of the host cell is modified completely. As the *T. spiralis* newborn larvae develop into L1 larvae, the host cell develops into a unique complex that harbours the parasite. The worm can reside in this complex relatively undetected by host immunity and for this reason the complex has been named the nurse cell.
Figure 1.6 The life cycle of *T. spiralis*

Diagram showing the full life cycle of *T. spiralis* inside an animal (mammal/bird/reptile) host. Severe pathology can occur if newborn larvae infect the central nervous system (CNS) or the heart. In acute cases this can lead to death. Image taken from Worm Book, courtesy of John W. Karapelou [127].
Since the target tissue of *T. spiralis* is skeletal muscle, understanding the biology of this tissue of choice is important for understanding the development of the nurse cell. The next section therefore introduces the process of skeletal muscle formation and maintenance in vertebrates.

### 1.14. Muscle cell differentiation, and *T. spiralis* induced dedifferentiation

The development of skeletal muscle tissues (myogenesis) in vertebrates initiates in a division of the developing embryo called the somite [128,129]. Progenitor cells in a region of the somite respond to muscle specific signals and develop into myoblast cells. These signals include the myogenic regulatory factor (MRF) family of transcription factors, MyoD, myogenin, Myf5 and MRF4. In fact, in double knockout MyoD and Myf5 in mice, no skeletal muscle tissue forms, illustrating how key these proteins are to the process of myogenesis [130]. This family of proteins are also required for the next step in myogenesis, the differentiation of the myoblasts into the myofibre. Firstly, myoblast cells become programmed to develop into myocytes (Figure 1.7). This involves the activation of muscle specific genes that commit the myoblast to exit the cell cycle [131]. Dramatic changes in transcription and ubiquitination lead to the repression of cell cycle factors such as cyclin D1, E and A, the induction of cyclin-dependent kinase inhibitors such as p21, and the up-regulation of tumour suppressor proteins such as the retinoblastoma protein pRb and p53. These events lead to terminal cell cycle exit in the G0 phase of the cell cycle [132-134]. Multiple myocytes then fuse with one another and further differentiate, forming myofibres [135]. The up-regulation of muscle cell-specific genes then leads to the maturation of the skeletal muscle fibre [136]. Myofibres are long thin multinucleated cells that can no longer divide and have the ability to contract. Skeletal muscle tissue is composed of multiple myofibres that have lined up in orientation, forming contractile bundles [137].

As briefly mentioned, the ubiquitin pathway plays an important role in myogenesis. For example, the MRF MyoD (myoblast determination factor) is a short-lived protein required for the commitment of myoblasts to differentiation. Inhibition of MyoD (only) inhibits the differentiation of myoblasts [138,139]. Although MyoD, and the MyoD interaction partner E2A (another transcription factor), are detected at the protein level in immature myoblasts (and thus expressed), their activity is down-regulated by post-translational mechanisms including proteasomal degradation [140-142]. During the differentiation of myoblasts into mature muscle cells, ubiquitin-mediated degradation
of MyoD and E2A is reduced, leading to their stabilisation. In turn this leads to the transcriptional activation of target genes required for the development of the mature muscle identity. After the differentiation process has taken place, MyoD is again down-regulated by proteasomal degradation.

During embryogenesis and beyond, not all of the progenitor cells in the aforementioned region of the somite become differentiated. These cells that do not differentiate are referred to as satellite cells and remain amongst the surrounding muscle tissues, even when fully formed in adulthood. Satellite cells have the ability to either remain quiescent, self renew or proliferate and develop into myoblast cells \[143,144\]. These cells are responsible for the growth, repair and regeneration of muscle fibres. Satellite cells become activated to differentiate into myoblasts and myocytes in response to growth factors, stress, aging or injury \[145,146\]. These new myocytes fuse with existing myofibres, thus replenishing the muscle tissue. This process is crucial for the maintenance of skeletal muscle homeostasis, since there are very few examples in nature of terminally differentiated skeletal muscle cells that can dedifferentiate, or re-enter the cell cycle, and it has never been observed in to occur spontaneously in mammalian skeletal muscle cells \[147,148\].

\textit{In vitro} however, the mitotic reactivation of mammalian skeletal myofibres can be induced by the small molecule myoseverin, infection by oncogenic DNA viruses (SV-40 and Polyoma), the down-regulation of the MRF myogenin and by the adenoviral oncogene E1A \[149-153\]. These studies indicate that despite their terminal differentiation state, mammalian skeletal myofibres do possess the potential to re-enter the cell cycle. Indeed when \textit{T. spiralis} newborn larvae invade host myofibres, this is exactly what happens. \textit{T. spiralis} invades the muscle cell without killing it, and induces the formation of the nurse cell. The nurse cell is the result of the dedifferentiation and cell cycle re-entry of the terminally differentiated host muscle cell.
Figure 1.7 Schematic of myogenesis

The development of a myofibre from muscle satellite cells and the involvement of muscle regulatory factors (MRFs), Myf5, MyoD, myogenin and MRF4. Quiescent muscle satellite cell are activated by growth factors, stress, aging or injury to develop into cycling myoblasts. Myoblasts commit to differentiate into myocytes, which exit the cell cycle and fuse. The fusion of multiple myocytes forms an elongated multinucleated terminally differentiated myofibre. Image adapted from Le Grand and Rudniki, 2008, Curr. Opin. Cell. Biol.
1.15. *Trichinella spiralis* and the nurse cell

*T. spiralis* newborn larvae break into and enter host skeletal muscle cells. The development of the nurse cell then begins with a significant morphological change of the host cell, where the myofibre shortens and becomes rounded and enlarged, partitioning itself from adjacent muscle fibres (Figure 1.8). One of the first biochemical changes that is detected is host cell mitochondrial dysfunction that leads to hypoxia [154,155]. The parasite then induces a significant change in host cell gene transcription. A down-regulation of terminally differentiated muscle specific genes such as actin, myosin light and heavy chain and tropomyosin is observed, and by day 9 after invasion, no host muscle markers can be detected [156,157]. In addition an up-regulation in cell cycle-related genes is observed. This re-programming of host protein expression leads to the reversal of the differentiation state of the muscle cell [158-161]. Between day 2 and 8 after invasion, the host cell re-enters the cell cycle by replicating DNA, before becoming arrested in the G2/M phase where the enlarged nuclei remain as 4n thereafter [162,163]. Between day 10 and 12, the expression of host collagen is induced, which is required for the formation of a protective capsule that surrounds the nurse cell complex [162,164-166]. Around day 11-12, proteins that play a role in angiogenesis are up-regulated, such as vascular endothelial growth factor, leading to the formation of a network of circulatory vessels around the nurse cell [167]. It is believed that this circulation provides a mechanism for the delivery of nutrients and removal of waste products from the complex. These vessels are described as venous sinusoids, as they are larger than normal capillaries and carry only deoxygenated blood [167,168]. Since the muscle larvae does not require oxygen, and a hypoxic environment in the nurse cell is induced, the parasite it thought to be anaerobic, requiring the blood for other nutrients such as glucose. As well as the cell cycle re-entry of terminally differentiated skeletal muscle cells, the stimulation of sinusoid growth is rarely observed elsewhere in nature. [156,162,163,169,170]. Approximately 15 days after the invasion of the myofibre the nurse cell is fully formed (Figure 1.9) [171].

The specific characteristics of the nurse cell are not observed anywhere else; they are unique to *T. spiralis* infection. Although similarities are observed, the dedifferentiation profile of a nurse cell is not akin to that induced by SV40 and Polyoma infection [153,172]. The mechanisms underlying nurse cell formation still remain undetermined. It is proposed that the parasite (rather than the host cell, surrounding cells or infiltrating cells) is responsible for the changes that result in this
transformation and there have been many studies that support this hypothesis. Firstly, the changes in host muscle and cell cycle factors occur at the transcriptional level, indicating a direct effect on the host cell genes [156,157]. Secondly, the nuclei that replicate their DNA, becoming enlarged and chronically 4n, are derived from the myofibre itself, and not from infiltrating activated satellite cells [163]. Furthermore, *T. spiralis* antigens can be located in the nurse cell nuclei and cytoplasm [173] [174,175]. The secreted proteins are therefore the prime candidates for the mediators of the reprogramming of muscle cells during nurse cell formation.
Figure 1.8

*Figure 1.8 The formation of the *T. spiralis* nurse cell*

Terminally differentiated host skeletal myofibres are invaded by *T. spiralis* newborn larvae. This causes a change in morphology, a loss of host cell muscle specific proteins leading to dedifferentiation and cell cycle re-entry as characterised by DNA synthesis and nuclear enlargement. *T. spiralis* secreted proteins are detected inside the host cell, and host collagen expression is up-regulated. This forms a capsule around the parasite-harbouring cell that becomes segregated from the original muscle bundle. Angiogenesis is induced characterised by a formation of a network of circulatory vessels that surround the new complex, termed a nurse cell. The cell then remains in G2/M cell cycle arrest. Meanwhile, the *T. spiralis* develops from a newborn larva into an L1 stage muscle larva.
Figure 1.9 The *T. spiralis* nurse cell complex

A. Photograph of a nurse cell isolated from infected muscle tissues, by Eric Grave. B. Hematoxylin and eosin stained paraffin section of *T. spiralis* infected rat muscle tissue, 1 month post infection.
1.16. Treatment of trichinellosis

Current treatment of Trichinellosis largely involves the use of wide-spectrum anthelmintics such as albendazole, mebendazole and thiabendazole that are effective during the early intestinal stages of infection (5-7 days after ingestion) [176]. However, due to the frequent misdiagnosis of Trichinellosis, this window is often missed. After this, the newborn larvae enter the circulation where successful treatment now relies on intestinal absorption of the drug. General anthelmintics show poor solubility and therefore low efficacy in treatment of circulating stage *T. spiralis* [177,178]. Furthermore, due to the protective characteristics of the nurse cell, muscle stage larvae cell are extremely resistant to treatment [179].

During the initial intestinal phase of infection, *T. spiralis* induces a T helper 1 (T₃H₁) immune response [180]. This mediates the eventual clearance of the adult parasites, but not before reproduction has taken place and the newborn larvae are on their way to the muscle. In the muscle tissues, *T. spiralis* muscle larvae stimulate an overall T helper 2 (T₃H₂) immune response followed by an induction of regulatory cytokines such as TGF-β and IL-10 [181] [182-184]. This response is protective to both the parasite and the host as it prevents elevated and potentially pathogenic pro-inflammatory cytokine activation. Sustained T₃H₂ immunity and the induction of regulatory cytokines is characteristic of helminth infection, whereas T₃H₁ immune responses can be characteristic of allergy and autoimmune disease and T₃H₂-based disorders such as asthma, do not observe the induction of regulatory cytokines. Thus helminth infection, including *T. spiralis*, can in some cases provide relief from chronic autoimmune and inflammatory diseases. Using *T. spiralis*, this has been demonstrated in animal models of colitis, multiple sclerosis (experimental autoimmune encephalomyelitis), and asthma (experimental airway allergic inflammation) [185-187].

Despite the overall T₃H₂ response, during the invasion of muscle tissues by *T. spiralis* newborn larvae a T-cell-mediated inflammatory response is observed, illustrated by a failure of immune cell recruitment to nurse cells of T cell deficient mice [188]. The invasion event by newborn larvae is physically damaging to the muscle cells. Together with newborn larvae antigens, this stimulates inflammation with the infiltration of innate immune cells, primarily macrophages and eosinophils. Together the invasion and the subsequent inflammation (myositis) can cause muscle damage. The collagen capsule of the nurse cell eventually protects the parasite from this
response but the surrounding host tissue is susceptible. Symptoms of nurse-cell induced myositis include cachexia, edema, and dehydration presenting as muscle pain, fatigue, weakness and loss. The immune-related symptoms are often effectively treated using corticosteroids, however these drugs do not target the parasite [176]. It is therefore of importance to investigate new *T. spiralis*-specific targets for the development of drugs or vaccines that show improved efficacy in the treatment of the muscle phase of infection. The identification of new targets will undoubtedly involve the characterisation of the proteins directly involved in host-parasite interactions: the secreted proteins.

### 1.17. The secreted proteins of *T. spiralis* muscle larvae

*Trichinella* muscle larvae are approximately 1mm in length. Adult females grow to between 1.4 and 4 mm and adult males to between 1.4 and 1.8 mm in length [189]. They can just about be seen with the naked eye as tiny flecks. Using techniques such as histology and electron microscopy, Despommier and later Takahashi described an organ in *T. spiralis* called the stichosome [190-192]. This organ is found in the upper part of the body of all nematodes, both parasitic and free living, and is composed of stichocyte cells (Figure 1.10). These cells contain secretory granules and are analogous to mammalian granulocytes. The *T. spiralis* secreted proteins are stored in these granules before being released into the host. This is illustrated by the binding of antibodies from *T. spiralis* infected animals to stichocyte cells within the muscle larvae [193,194]. There are thought to be in excess of 100 proteins secreted from the stichocyte cells by *T. spiralis* muscle larvae. These proteins can be detected in both the cytoplasm and nucleus of the nurse cell using antibodies against an immunogenic *T. spiralis*-specific glycan tyvelose, which decorates the secreted glycoproteins [175].

Many of the changes in host cell transcription of muscle cells during nurse cell formation have now been reported using microarray [158,159]. We therefore have a good idea of the molecular basis of the changes that occur to form nurse cells from the point of view of the muscle cell, but very little idea of the molecular role of the parasite. In 1993, Jasmer *et al.* attempted to determine the source of the induction of muscle cell transformation. They did so using a technique that had been previously shown successful for creating a synchronous *T. spiralis* infection. This involves the collection of newborn larvae released by pregnant adult *T. spiralis* in culture and the injection of the newborn larvae directly into the muscle tissue [195]. Before injection,
Jasmer et al. irradiated the newborn larvae, causing an inhibition of their growth and an inhibition of the development of the α and β stichocyte cells (Figure 1.10). The irradiated newborn larvae were able to invade the host muscle cells, but their cell cycle re-entry was markedly delayed. In addition, once the nurse cell did eventually form, it was significantly smaller than those formed by healthy T. spiralis. Results showed that the healthy development of some of the secretory cells was, in part, required for normal nurse cell formation. These experiments therefore suggest that the secretion of proteins by T. spiralis newborn larvae, underlie at least some of the mechanisms involved in nurse cell formation as it is observed during infection. In order to maintain the chronic status of the nurse cell for prolonged periods of time (years), it is plausible that the parasite continues to secrete proteins into the nurse cell, long into its development into L1 (muscle) stage larvae. Indeed, Guiliano et al. discovered novel T. spiralis secreted proteins in the nurse cell nuclei of rat muscle tissue 2 months after infection, where the parasite has long existed as an L1 stage larvae. It is therefore strongly suggestive that the direct communication between host-targeted T. spiralis secreted proteins and the host cell stimulates the transformation of skeletal muscle cells.

To investigate how the secreted proteins achieve this, they must be characterised and functions must be assigned. To date, only a small proportion of T. spiralis secreted proteins have been identified or characterised [196,197] [198-200]. These proteins include proteases, nucleotidases, antigens and novel T. spiralis proteins. Some of those that have been characterised have also been found located inside the nurse cell and inside the nurse cell nuclei [201-204]. However the role of these proteins in the host-parasite interactions that lead to nurse cell formation remains to be determined [205].
**Figure 1.10 The morphology of a *T. spiralis* muscle larva**

1.18. *T. spiralis* as a model for investigating the ubiquitin pathway in host-parasite interactions

The ubiquitin pathway is heavily implicated in cell cycle control, myogenesis, the terminal differentiation of skeletal muscle cells, and when it is disrupted can lead to muscle wasting disorders [40,42]. Using models of muscle cell reprogramming such as *T. spiralis* infection, and by studying the host-targeted secreted proteins of *T. spiralis*, this project aimed to investigate whether or not parasites target the ubiquitin pathway during infection. The purpose of this was two-fold: firstly to help us to understand muscle physiology, and secondly to highlight potential vaccine or drug targets for the development of novel anthelmintics.

By characterising the proteins that are expressed and secreted by *T. spiralis* and, most importantly, investigating the biological systems in which they function, this investigation aimed to shed light on the complex processes of skeletal muscle deregulation. This may help us understand the molecular basis of muscle differentiation, dedifferentiation, cell cycle regulation and in turn, muscle tissue disorders, muscle cell cycle-related diseases and the process of induced angiogenesis. From here, ideas may arise of how to find ways to deliberately manipulate these processes for therapeutic purposes. This information may contribute to efforts that are being made to understand the process of muscle cell dedifferentiation in order to explore the possibilities of induced muscle tissue repair and regeneration [206,207].

During my MRes project that preceded the PhD investigation, I identified a number of putative DUBs expressed by *T. spiralis* muscle larvae [1]. One of these DUBs, TsUCH37, is the orthologue of human UCH37 (UCH-L5) a proteasome-associated DUB that, as mentioned previously, is currently receiving attention for its potential as a drug target for the treatment of cancer. Being the only UCH domain proteasome-associated DUB in mammals, UCH37 plays a critical role in protein degradation and is essential for survival [36]. The characterisation of TsUCH37 therefore provided an interesting starting point for the investigation of the ubiquitin pathway in *T. spiralis* muscle larvae, its evolution and potential as a drug target.
1.19 Aims of the project

Main aim:

To investigate the role of the ubiquitin pathway during the infection of host muscle tissues by *Trichinella spiralis*.

**Aim 1.** To characterise TsUCH37, a DUB expressed by *T. spiralis* muscle larvae

A. To characterise the function of TsUCH37

B. To investigate the role of TsUCH37 as a drug target

**Aim 2.** To investigate whether parasites specifically target the host Ub pathway during infection using *Trichinella spiralis* as a model system

A. To look for TsUCH37, and other *T. spiralis* ubiquitin pathway enzymes, in the secreted proteins of the muscle larvae

B. To characterise the function of secreted *T. spiralis* ubiquitin pathway enzymes in mammalian skeletal muscle cells
Chapter 2: Materials and methods
Chapter 2: Materials and methods

2.1. Bioinformatics analyses

The *T. spiralis* contig database was downloaded (August 2010) from the University of Washington Genome Centre [208]. Contig 1.2, between bases numbered 30000 to 40000, was analysed using the following programmes: For alignments: Geneious software using either the MUSCLE or the Geneious alignment algorithm [209]. For homology and conserved domain identification: NCBI BLAST and SMART (Simple Modular Architecture Research Tool, [210,211]). For Gene prediction: AUGUSTUS [212,213], FGENESH [214], SNAP [215] and GenemarkHmm [216]. Gene predictions were carried out in collaboration with Eliseo Papa (Harvard/MIT). Initial mass bioinformatics analyses of *T. spiralis* annotated proteins (RefSeq) were carried out by Derek Huntley, Bioinformatics Support, Imperial College London. Programmes were used to identify protein localisation motifs and signal sequences using wolfPSORT [217], iPSORT [218], and signalP [219]. Sequences were further analysed using SMART and InterPro to identify conserved domains.

2.2. Parasite isolation, culture, re-infection and collection of secreted proteins

Ethics statement: All procedures involving care and maintenance of animals were approved by the Imperial College Ethical Review Committee and performed under license from the UK Home Office.

*T. spiralis* parasites were maintained in female Sprague-Dawley rats and L1 stage muscle larvae were isolated from infected tissue by digestion of skeletal muscle with acidified pepsin as described by Arden et al. [116]. Briefly, muscle larvae were isolated from infected rats by digestion of skeletal muscle with 1% v/v HCL and 1% w/v pepsin in tap water at 37ºC for 1.5 hours shaking before being filtered through muslin cloth. Muscle larvae were pelleted by gravity, and pellets were washed in sterile 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4.2H2O, 2 mM KH2PO4 pH 7.4) and cultured in sterile serum-free RPMI (Gibco) supplemented with 1% w/v glucose, 100 U/ml penicillin, 100 µg/ml gentamycin, 20 U/ml nystatin and 2 mM glutamine at 37ºC and 5% CO2 with a daily change of medium for a maximum of 6 days. Culture supernatants were collected and filtered through a 0.2 µm membrane. The secreted proteins were then dialysed into 25 mM Hepes buffer pH 7.4 and
concentrated using 10 kDa molecular weight cut-off (MWCO) vivaspin columns (Sartorius Stedim). Protein concentrations were measured using the BCA assay (Pierce). For re-infections, 3500 muscle larvae in 300 µl of PBS were used to infect each rat orally by gavage feeding.

**Mammalian cell culture**

### 2.3. HEK 293T cells (adherent)

HEK 293T (Human Embryonic Kidney 293T cells) cells were used for human cell lysate protein control samples and for lentivirus production. Cells were cultured in DMEM supplemented with 10% v/v foetal bovine serum (Gibco), 100 U/ml penicillin 100 µg/ml streptomycin and 4 mM L-glutamine (Sigma-Aldrich) at 37°C and 5% CO₂. Cells were grown in vacuum gas plasma treated flasks or petri dishes and routinely split 1:20 every 3 days using 0.05% w/v trypsin-EDTA. For experiments were harvested by directly lysing in NP-40 lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, and 1% NP-40 v/v). Cell stocks were frozen in liquid nitrogen in growth medium supplemented with 10% v/v DMSO.

### 2.4. C2C12 myoblasts (adherent)

Mouse C2C12 myoblasts were originally established by Yaffe and Saxel and produced by Blau, *et al.* [220,221]. Cells were cultured in high-glucose DMEM (Life Technologies Cat # 41965, without pyruvate, without HEPES) supplemented with 20% v/v foetal bovine serum (Gibco), 100 U/ml penicillin 100 µg/ml streptomycin and 4 mM L-glutamine (Sigma-Aldrich) at 37°C and 5% CO₂. Cells were grown in vacuum gas plasma treated flasks or petri dishes and split 1:10 every 2 days using 0.5% v/v trypsin-EDTA and never allowed to grow beyond 60-70% confluency. For experiments (apart from immuno-fluorescence and flow cytometry), cells were harvested in cold 1x PBS and pelleted by centrifugation for 10 minutes at 5,000xg. Cell stocks were frozen in liquid nitrogen in growth medium supplemented with 40% FBS v/v and 10% v/v DMSO.

### 2.5. C2C12 differentiation (adherent)

C2C12 myoblast cells were grown to 90% confluency, washed in sterile 1x PBS and media was changed to starvation/differentiation media (high-glucose DMEM (Life Technologies Cat # 41965, without pyruvate, without HEPES) supplemented with 2%
v/v horse serum (Gibco), 100 U/ml penicillin 100 µg/ml streptomycin and 4 mM L-glutamine. Media was changed every day and cells were differentiated for 3-4 days (or when fully matured myotubes were observed) before using for subsequent experiments. Cells were harvested in cold 1x PBS and pelleted by centrifugation for 10 minutes at 5,000xg.

2.6. Lentivirus production

HEK 293T cells have been modified to express the simian virus 40 (SV40) large T antigen. This allows the replication of episomal (non-integrated) plasmids that contain the SV40 origin of replication (oriC). Lentiviral accessory plasmids were kindly donated by Hidde Ploegh. HEK 293T cells were seeded into 15 cm petri dishes 24 hours before transfection. Fresh HEK 293T growth media was added. A mixture of the mammalian (lentiviral) expression vector (either pCSGW tdTomato TsUBE2L3-HA, pLVX Tet On Advanced, pLVX TsUBE2L3 or pLVX HA-Ub) and the lentiviral accessory plasmids was prepared, (pCMV-VSVG, pCMV-Gag/Pol, pCMV-TAT and pCMV-REV) using the relative amounts listed in Table 2.1. Cells were transfected with the plasmid mixture using TransIT-293 Transfection Reagent (Mirus Bio) according to the manufacturer’s instructions, using Opti-MEM I reduced serum media (Gibco). Growth media (supernatant containing lentivirus particles) was collected every 24 hours for 72 hours and fresh growth media was added. Supernatants were pooled and sterile-filtered (0.45 µm pore size) before concentration of lentivirus by centrifugation at 20,000xg for 90 minutes. Lentivirus was resuspended in fresh C2C12 growth media and used directly for transduction.

<table>
<thead>
<tr>
<th>Amount of DNA (µg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammalian expression vector</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>30.6</td>
</tr>
</tbody>
</table>

Table 2.1 DNA mixture for transfection for lentivirus production

2.7. Lentivirus transduction of C2C12 myotubes

For C2C12 myotubes, myoblasts cell lines were counted, seeded into a 6-well plate and differentiated. Three days after the addition of differentiation media, cells were
transduced. Lentivirus (2 ml per well) was supplemented with Polybrene (Sigma-Aldrich, 8 µg/ml) and added to each well. Cells were centrifuged with the lentivirus for 90 minutes at 2000xg for “spinfection”. Fresh differentiation media was then added to the cells 6 hours post-transduction. Cells transduced with pCSGW tdTomato lentivirus were examined 48 and 72 hours post-transduction using a Leica DMR HC fluorescence microscope (Leica Microsystems, Germany) to visualise expression of the tdTomato fluorescent protein.

2.8. Lentivirus transduction of C2C12 myoblasts and generation of stable lines (inducible expression system)

For C2C12 myoblasts and generation of stable cell lines, cells were counted and seeded into a 6-well plate 24 hours before transduction to ensure 70% confluency for transduction. The control vector (pLVX Tet On) lentivirus and the expression vector (pLVX Tight Puro) lentivirus were added to the cells (with 8 µg/ml Polybrene) at the same time for simultaneous transduction (1 ml each per well). Transductions were carried out as described above and 48 hours later, geneticin (G418, 500 µg/ml) and puromycin (4 µg/ml) were added to the growth media for selection. Control cells were completely dead after 5 days of drug selection. Cells were maintained in growth media supplemented with 250 µg/ml G418 and 4 µg/ml puromycin. Stable C2C12 cell line protein (of interest) expression was induced using 2 µg/ml doxycycline.

2.9. RNA preparation for RACE and q-RT-PCR

For RNA preparation from T. spiralis pelleted T. spiralis L1 larvae were frozen in liquid nitrogen before being shattered in a percussive disruptor. For C2C12 myotubes, cell pellets were harvested as described above (section 2.5). RNA was then extracted using Trizol reagent (Invitrogen) using the manufacturer’s protocol. All procedures involving Trizol and phenol/chloroform were carried out in a fume hood. Briefly, 1 ml of Trizol was added to 100 mg of shattered muscle larvae and pipetted and vortexed for lysis. Lysate was then cleared by centrifugation at 12,000xg for 10 minutes. The supernatant containing was transferred to a fresh eppendorf and 0.2 ml chloroform was added. Samples were vortexed and centrifuged at 12,000xg for 15 minutes at 4°C. The upper aqueous phase was transferred to a new eppendorf, 0.5 ml 100% v/v isopropanol was added and incubated at room temperature for 10 minutes before centrifugation at 12,000xg at 4°C for 10 minutes. The pellet was then washed 3 times in 1 ml 75% v/v ethanol (using centrifugation steps) before being
solubilised in 50 µl of RNase free water. To remove genomic DNA, RNA was treated with Turbo DNase using the Turbo DNA-free kit (Ambion). For RACE, total RNA was used directly after extraction.

2.10. RACE (Rapid Amplification of cDNA Ends)

Primers based on the AUGUSTUS algorithm predicted 931 bp coding sequence of TsUCH37 were used as a positive control and to test GSRPs. Two RACE kits were used (according to manufacturer's instructions), to process RNA as explained in results section 3.1. The FirstChoice RNA Ligase Mediated Rapid Amplification of cDNA Ends (RLM-RACE) kit from Ambion and the GeneRacer kit from Invitrogen. Nested PCR was then carried out using gene specific primers with the RACE primers supplied by the RACE kits.

Gene specific RACE primers:

**GSRP TsUCH37 (FGENESH), primer 1**
CACCGGCACCAACGCCTCGCAACGCCACCACACATT

**GSRP TsUCH37 (AUGUSTUS), primer 3**
CGTCGGTCCCAAATCTACGTCCGGGTGGTCGCAGTT

**GSFP TsUCH37 (same for AUGUSTUS and FGENESH) primer 4**
CGGGCGGCTCCACTGCGTCTCGGCACTGTT

**GSRP TsUBE2L3**
CCAATTCTGTACGAAGCGAATGTTCC

**GSFP TsUBE2L3**
ACGGAGCACATTATTGATTCGTTGATC

RACE-PCR amplified products were ligated into pGEMTeasy using TA cloning (described in the results section 3.1) and transformed into *E. coli* XL10-Gold competent cells (Agilent Technologies) according to the manufacturer’s instructions. Bacteria were plated onto ampicillin and XGal (80 µg/ml) and IPTG 20 mM agar plates, incubated overnight at 37°C and blue/white screening was used to select positive colonies. To verify the presence of the correct insert, clonal cultures were grown and plasmids were purified using the Qiaprep mini-prep kit (Qiagen) according to the manufacturer’s instructions. The insert was sequenced in pGEMTeasy (Beckman Coulter Genomics) and diagnostic restriction digests using BamHI and NotI were carried out. Insert sequences were verified by alignment.
2.11. qRT-PCR

For qRT-PCR total RNA (extracted and purified using Trizol/phenol/chloroform, see above) was first passed through an RNeasy column using the RNeasy kit (Qiagen). Total RNA was used to synthesise cDNA with both random primers and oligo-dT(12-18) primers using the Takara PrimeScript First Strand cDNA Synthesis kit (Clontech) according to the manufacturer’s instructions. QRT-PCR was carried out using TaqMan Gene Expression Assays (p53: Mm01731287_m1, MHC: Mm01332489_m1, and GAPDH: Mm99999915_g1) and TaqMan Fast Universal PCR Master Mix (2x) in a MicroAmp Fast Optical 96-Well Reaction Plate (all Life Technologies. All qRT-PCR samples were analysed in triplicate, using 1 μl of 1/50 dilution of cDNA per reaction. PCR was carried out using the ABI Prism 7500 Fast Real-Time PCR System (Applied Biosystems) according to the manufacturer’s instructions. The GAPDH TaqMan assay was used as “housekeeping” reference to normalise the samples according to the amount of total cDNA used. Gene expression in C2C12 myotubes expressing TsBE2L3 was quantitated relative to C2C12 myotubes carrying the empty vector DNA. The amplification efficiency of the MHC and P53 target genes was approximately equal to the amplification efficiency of GAPDH and so Comparative CT (ΔΔCT) Method was used (see Applied Biosystems User Bulletin #2).

Cloning

2.12. TsUCH37

Pelleted *T. spiralis* L1 larvae were frozen in liquid nitrogen before being shattered in a percussive disruptor. Samples were thawed and total RNA was extracted using the RNeasy kit (Qiagen). The RACE-confirmed sequence was amplified using the One Step RT-PCR kit (Qiagen).

Primers:
Forward
TGCAGGATCCatggctgaaggaacttggttttaa
Reverse
ACAGTGGCGGCGGtttaagcaagaaatcatgtgcaaa

Primers were designed to include BamHI and NotI linkers (capitals) against the 5’ and 3’ regions (in capitals, respectively). RT-PCR amplification of a 313 bp fragment
(121-434) of *T. spiralis* GM2 activator protein was carried out as a control [199]. The PCR product was ligated into the multiple cloning site of pGEMTeasy (Promega) using TA cloning (described in results section 3.1) and transformed into *E. coli* XL10-Gold competent cells (Agilent Technologies) according to the manufacturer’s instructions. Bacteria were plated onto ampicillin and XGal/IPTG agar plates, incubated overnight at 37°C and blue/white screening was used to select positive colonies. To verify the presence of the correct insert, clonal cultures were grown and plasmids were purified using the Qiaprep mini-prep spin kit (Qiagen) according to the manufacturer’s instructions. Diagnostic restriction digests using BamHI and NolI were carried out and the inserts were sequenced in pGEMTeasy (Beckman Coulter Genomics). DNA fragments were separated by agarose (0.8% w/v) gel electrophoresis, stained with ethidium bromide and visualised using UV light. The correct insert was then gel-purified using the QIAquick spin gel purification kit (Qiagen) according to the manufacturer’s protocol (Qiagen), digested using BamHI and NolI enzymes (Promega), purified and cloned into pPET28a(+), containing an upstream His-tag sequence. Plasmids were transformed into *E. coli* XL10-Gold competent cells (Agilent Technologies) according to the manufacturer’s instructions and selected for kanamycin resistance. Plasmids were purified using the QIAprep spin mini or midi prep kits (according to the manufacturer’s protocol, Qiagen).

### 2.13. Site-directed mutagenesis

The mutant 6His-TsUCH37 D12N was generated by mutating the Aspartic acid residue at position 12 to an Asparagine residue. This corresponded to codon GAT, at nucleotides 34 to 36 where G was mutated to A (small letters, bold). Mutagenesis was carried out using the QuickChange Site-directed Mutagenesis kit according to the manufacturer’s instructions (Stratagene).

Primers for mutagenesis:

**Fw TsUCH37 G34A**
AATGGTTTTAAATAGAAAGTAATCCTGGAATATTTACTGAAATGATTC

**TsUCH37 wild type**
AATGGTTTTAAATAGAAAGTGATCCTGGAATATTTACTGAAATGATTC

**Rev TsUCH37 G34A**
GAATCATTTTCAGTAATTTATCTCCAGGATACTTTCTATTTAACACCAATT
Diagnostic digests were carried out using Dpn I and inserts were sequenced to confirm the presence of the nucleotide substitution (Beckman Coulter Genomics). DNA fragments were separated by agarose gel electrophoresis and visualised by ethidium bromide staining and UV light. 6His-TsUCH37 D12N was amplified by PCR using the primers originally designed to clone the wild-type 6His-TsUCH37. The PCR product was cloned into pGEMTeasy (as described above) before being cloned into pPET28a+ (as described above). The same expression and Ni-NTA-mediated Histagged purification methods were used for 6His-TsUCH37 D12N as were used for 6His-TsUCH37 wild type (described above).

2.14. TsUBE2L3

Pelleted T. spiralis L1 larvae were frozen in liquid nitrogen before being shattered in a percussive disruptor. Samples were thawed and total RNA was extracted using the RNeasy kit (Qiagen). Total RNA was used to synthesise cDNA using oligo-dT_{12-18} primers (M-MLV Reverse Transcriptase, Life Technologies). The RACE-confirmed coding sequence for TsUBE2L3 was amplified by PCR using the following primers:

Forward
CCGGATCCatgactgcgagtcgtagacttcaaaaagag
Reverse
GCGCGGCGCGttaGTCGAGTGCAGTAGTCTGGTACGTCAACATCCTGACTTTTCCGCGTATT

The forward primer was designed to contain a 5’ BamHI linker sequence (bold) and the reverse primer was designed to contain a 3’ NotI linker sequence (bold) a stop codon (small letters) and a HA-tag (capitals). Both coding sequence portions are in small letters. The PCR product was ligated into the multiple cloning site of pGEMTeasy (Promega) using TA cloning (described in the results section 3.1) and plasmids were amplified in bacterial cultures, and purified as described for TsUCH37 above. Diagnostic restriction digests using BamHI and NotI were carried out and the inserts were sequenced in pGEMTeasy (Beckman Coulter Genomics). Plasmids containing the correct insert were then digested using BamHI and NotI enzymes (Promega), the inserts were gel-purified using the QIAquick spin gel purification kit (Qiagen) according to the manufacturer’s protocol (Qiagen) and cloned into either pCSGW tdTomato or pLVX Tight Puro (Table 2.3). Plasmids were transformed into E. coli XL10-Gold competent cells (Agilent Technologies) according to the manufacturer’s instructions and selected for ampicillin resistance. Plasmids were
purified using the QIAprep spin mini or midi prep kits (according to the manufacturer's protocol, Qiagen).

2.15. Recombinant protein expression and purification

Purified pPET28a(+) 6His-TsUCH37 plasmids were transformed into *E. coli* Rosetta-2 (BL21 derivatives, Novagen), plated onto agar plates and clones were selected for kanamycin resistance. These were then grown overnight in 5 ml Luria-Bertani Broth (LB) kanamycin (50 ug/ml) media before inoculating 1 litre LB. At an optical density of 0.6 (600 nm) bacterial cultures were induced with 0.5 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) and grown for a further 3 hours at 37°C before being harvested. Cell pellets were lysed either using BugBuster™ protein extraction reagent (Novagen) supplemented with lysozyme and DNase I or by sonication for 4 x 10 minutes on ice in lysis buffer (50 mM NaH₂PO₄ pH 8.0, 300 mM NaCl) supplemented with lysozyme and DNase I. Lysates were cleared by centrifugation for 20 minutes at 16,000 xg at 4°C. Samples of the soluble and insoluble fractions were separated and analysed by SDS-PAGE. Inclusion bodies were re-solubilised in 50 mM NaH₂PO₄ pH 8.0, 300 mM NaCl, 8M Urea at 4°C. Recombinant protein was purified from re-solubilised inclusion bodies using Ni-NTA resin (Qiagen) under denaturing conditions according to the manufacturer’s protocol. Recombinant 6His-TsUCH37 was then slowly dialysed using SnakeSkin (Pierce) out of denaturing buffers into native protein buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl) at 4°C overnight with a change of buffer followed by a further 4 hours.

2.16. AMC activity assays

Recombinant purified proteins or cell or parasite lysates were used at the concentrations indicated in the results. Ub C-terminal 7-amido-4-methylcoumarin (Ub-AMC) was used at 250 nM and Nedd8-AMC was used at 500 nM (both from Boston Biochem) diluted in reaction buffer (150 mM NaCl, 50 mM Tris/HCl pH 7.5, 2 mM EDTA, 2 mM DTT supplemented with 1 mg/ml bovine serum albumin). All assays were carried out at room temperature. Negative control samples were incubated with 2 mM NEM for 20 minutes before addition of the AMC substrate. AMC cleavage was measured by fluorescence at 368 nm excitation and 467 emission wavelengths on a FLUOstar microplate reader (BMG LABTECH). All measurements were made in a 384 well plate (Nunc, black, Thermo-scientific) in a total volume of 20 µl, in triplicate. All protein concentrations were measured using the BCA assay
(Pierce). For inhibition assays, protein samples were incubated with either LDN-57444 or WP1130 for 1 hour at 37°C prior to the addition of Ub-AMC.

2.17. Co-precipitation with 6His-TsUCH37

For co-precipitation analyses *T. spiralis* muscle larvae or HEK 293T cells were lysed using lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, and 1% NP-40 v/v) and dounce homogenisation. Lysates were cleared by centrifugation at 16,000 xg and the cleared supernatant lysate protein concentration was determined by BCA assay (Pierce). Lysates (3 mg per sample) were pre-cleared using wash buffer (Tris-HCl pH 7.4, 150 mM NaCl, and 0.1% NP-40 v/v) equilibrated-Protein G sepharose beads (Sigma-Aldrich) at 4°C for 2 hours. Recombinant purified 6His-TsUCH37 (500 µg) was bound to Ni-NTA agarose beads (Qiagen) in binding buffer (1x PBS pH 8.0, 300 mM NaCl, 0.1% NP40 v/v, 50 mM imidazole and 2 mM β-mercaptoethanol) for 2 hours at 4°C and beads were then washed 3 times in binding buffer. Pre-cleared *T. spiralis* lysates were incubated with either 6His-TsUCH37-bound Ni-NTA resin or native Ni-NTA (no protein bound control) overnight at 4°C. Resin was washed 3 times in wash buffer before boiling in 2x protein (SDS) loading buffer for SDS-PAGE analysis of co-precipitated proteins. Proteins were visualized either by immuno-blot analyses or by staining with colloidal coomassie G-250 for LC/MS/MS.

2.18. PNGase treatment

The removal of *N*- and *O*-linked glycans from *T. spiralis* secreted proteins was carried out using Peptide-*N*-Glycosidase F (PNGase F, New England Biolabs) according to the manufacturer’s instructions. Briefly, *T. spiralis* secreted proteins were heated at 94°C with glycoprotein denaturing buffer (0.5% w/v SDS, 40 mM DTT, 1 µl of 10x buffer per 20 µg protein) before adding G7 reaction buffer (500 mM NaH₂PO₄ pH 7.5) and NP-40 (1% v/v) and 2 µl PNGase F and incubating at 37°C for 3 hours.

2.19. Tandem mass spectrometry (LC/MS/MS)

Proteins for analysis by tandem mass spectrometry were separated by SDS-PAGE and manually extracted from the gel. All samples analysed in collaboration with Eric Spooner (Whitehead Institute for Biomedical Research) were processed as follows: protein samples were digested by trypsinisation as described by Kinter et al. [222], and analysed by MS/MS by as
described by Borodovsky et al. [223]. Briefly, samples were separated using a nanoflow liquid chromatography system (Waters Cap LC). The LC system was directly coupled to a tandem mass spectrometer (Q-TOF micro, Micromass). Analysis was performed in survey scan mode and parent ions with intensities greater than 6 were sequenced in MS/MS mode using MassLynx 3.5 Software (Micromass). LC/MS/MS data were then searched against the *T. spiralis* protein database (NCBI RefSeq) using two independent algorithms, Mascot [224] and SEQUEST [225] as stated in the results.

All samples analysed in collaboration with Michael Weekes and Steve Gygi (Harvard Medical School) were processed as follows: proteins were digested with trypsin using standard protocols and peptides were analysed on an Orbitrap XL2 mass spectrometer equipped with an Agilent 1100 binary pump and a Famos microautosampler. Peptides were separated using a gradient of 6 to 28% v/v acetonitrile in 0.125% v/v formic acid over 90 minutes. Peptides were detected in the Orbitrap by means of a data-dependent top 10 method. Each full scan was followed by the selection of the most intense ions, up to 10, for collision-induced dissociation (CID) in the linear ion trap. MS2 spectra were searched using SEQUEST v.28 against a composite database derived from the UniProt *Trichinella spiralis* proteome, its reversed complement and known contaminants. Peptide spectral matches were filtered to a 1% false discovery rate (FDR) using the target-decoy strategy combined with linear discriminant analysis. Peptides from all fractions in each experiment were combined and assembled into proteins. Protein scores were sorted by rank, and filtered to 1.7% FDR.

**2.20. UCH inhibitor compounds**

WP1130 was kindly provided by Dr William Bornmann (University of Texas, M.D. Anderson Cancer Center, Houston, TX). The compound was solubilised in DMSO and used at concentrations between 5 µM and 500 µM as indicated in the results. LDN-57444 (Sigma Aldrich) was solubilised in DMSO and used at concentrations between 10 µM and 1 mM as indicated in the results. DMSO alone was used as a negative control for inhibition assays. HsUCHL1 and HsUCH37 were provided by Katerina Artavanis-Tsakonas and used at 100 nM and 500 nM respectively as positive controls for inhibition.
2.21. MTT viability assay

Parasite viability was measured by a quantitative colorimetric assay with the tetrazolium salt 3-[4,5-diethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide, MTT (Sigma-Aldrich). Following treatment in culture, 1000 T. spiralis muscle larvae were incubated in 5 mg/ml MTT in phenol red free RPMI 1640 (Sigma-Aldrich) for 4 hours at 37°C. Formazan crystals were solubilised by shaking larvae in 200 µl of 100 % v/v DMSO for 1 hour at room temperature. Parasites were removed by centrifugation and the absorbance of the supernatant at 575 nm was determined. Parasites that had been killed by heat treatment (65°C for 10 minutes) were used as a positive control for drug inhibition. All measurements were made in triplicate. Data was statistically analysed by Students t test.

2.22. Crystallisation and structure determination of TsUCH37

Crystallography and structure determination was carried out by Marie Morrow at Purdue University. The materials and methods for protein are reported by Morrow et al., 2013, Biochemistry [3].

2.23. Ubiquitin-based inhibitor probe production:

HA-Ub-VME, HA-Ub-TF₃BOK and HA-Ub-Lac

Human Ub with an upstream HA tag cloned into pTYB1 was kindly donated by Hidde Ploegh (Whitehead Institute for Biomedical Research). PTYB1 contains a C-terminal intein/chitin-binding domain. HA-Ub-intein/chitin fusion protein was expressed in BL21 Rosetta 2 E.coli (Novagen) with a 2-hour induction using 0.5 IPTG at 30°C. Bacteria were harvested, centrifuged at 5,000xg for 20 minutes and pellets were resuspended in 25 mM MES (2-(N-morpholino)ethanesulfonic acid) pH 6.0, 100 mM NaOAc supplemented with protease inhibitor cocktail (Roche) and 1 mM MgCL₂, 10 µg DNase I, 1 mM DTT and lysozyme, 10 µg /ml. Bacteria were lysed by sonication for 4 x 10 minutes on ice. Lysate was cleared by centrifugation at 4°C for 30 minutes at 15,000xg. The cleared lysate was then purified using a chitin affinity column (New England Biolabs) and a protocol modified from Borodovsky et al. [223]. The lysate was loaded onto a chitin column (15 ml packed resin per 50 ml of lysate). The column was washed using lysis buffer supplemented with 0.5M NaCl (200 ml) and then using lysis buffer (400 ml). HA-Ub-MESNa was then eluted by incubating the
column in lysis buffer supplemented with 50 mM β-mercaptoethanesulfonic acid sodium salt (Mesa) overnight at 37°C. The HA-Ub-MESNa thioester was concentrated using 3 kDa molecular weight cut-off (MWCO) vivaspin columns (Sartorius Stedim). HA-Ub-MESNa was converted to HA-Ub-VME/ TF₃BOK/Lac probe using either glycine-VME that was kindly donated by Hidde Ploegh (generated as described by Borodovsky et al., supplemental data) or glycine 2,6-trifluoromethylbenzoxymethyl-ketone (TF₃BOK) or alpha-amino-beta-lactone (Lac) (both produced in collaboration with Jennifer Ward as described in her MRes research manuscript 2012). The probe (500 µl of 2 mg/ml HA-Ub-MESNa) was mixed with 300 µl of 2M NHS (N-hydroxysuccinimide), 1M NaHCO₃ pH 8 and 800 µl of 0.25M of either warhead (glycine-VME, glycine 2,6-trifluoromethylbenzoxymethyl-ketone (TF₃BOK) and alpha-amino-beta-lactone (Lac), 1M NaHCO₃ pH 8 and incubated at 37°C overnight. The mixture was then diluted out into 50 mM NaOAc pH 4.5 (buffer A), and loaded onto a Superloop MonoS column. Probe was further purified by ion exchange chromatography using buffer B (50 mM NaOAc pH 4.5 + 1M NaCl, 0.5 ml/min flowpaths, pressure 2.5 mPa) with the help of Jan Silhan and Frank Zhao. Fractions were concentrated using 3 kDa MWCO vivaspin columns (Sartorius Stedim) and tested for activity by diluting 1:3 into 50 mM Tris pH 7.4 and 150 mM NaCl and incubating for 1.5 hours at room temperature with recombinant purified 6His-TsUCH37 (1 µM). Control 6His-TsUCH37 samples were pre-incubated with 2 mM N-ethylmaleimide (NEM, an alkylating agent that irreversibly inhibits active-site cysteine activity of enzymes, Sigma Aldrich) for 20 minutes at room temperature. Samples were then analysed by SDS-PAGE and visualised by coomassie staining or immuno-blot analysis.

### 2.24. Ubiquitin-based inhibitor probe reactions

**HA-Ub-VME, HA-Ub-TF₃BOK and HA-Ub-Lac**

The secreted proteins of *T. spiralis* muscle larvae were collected as described in materials and methods section 2.2. Protein concentration was determined by BCA assay (Pierce). For probe reactions, the reducing agent dithiothreitol (DTT) was added (1 mM) to the secreted proteins. The Ub-based inhibitor probes HA-Ub-VME/TF₃BOK/Lac (generated as described above) and FLAG-Nedd8-VS (generated by Hemelaar et al. [226] and kindly donated by Hidde Ploegh) were added to the secreted proteins of *T. spiralis* muscle larvae. For all reactions 0.3 µg probe was used per 20 µg parasite protein. For immuno-blot analyses, 300 µg of *T. spiralis* samples were then analysed by SDS-PAGE and visualised by coomassie staining or immuno-blot analysis.
secreted protein was reacted with the probes. For the scaled up reaction for HA-immuno-precipitation and LC/MS/MS, 3 mg of *T. spiralis* secreted protein was reacted with the probes. *T. spiralis* secreted proteins were reacted with the probe for 1.5 hours at room temperature. NEM was added (2 mM) to control protein samples and incubated at room temperature for 20 minutes before reaction with the probe. For immuno-blot analyses, samples were boiled in 4x protein (SDS) loading buffer (200 mM Tris-HCl pH 6.8, 8% w/v SDS, 0.4% w/v bromophenol blue, 40% v/v glycerol, 200 mM DTT) and separated by SDS-PAGE before immuno-blot analyses using anti-HA or anti-FLAG antibodies (as described in materials and methods section 2.36). For the scale-up reaction for LC/MS/MS, samples were used for anti-HA immuno-precipitation experiments (see below).

### 2.25. HA-Ub-VME HA immuno-precipitations

For anti-HA immuno-precipitation experiments SDS was added (0.4% w/v) to each sample before diluting out the SDS to 0.1% w/v with wash buffer containing 0.1% v/v NP-40, 50 mM Tris-HCl pH 7.4, 150 mM NaCl. Protein G sepharose resin (Sigma-Aldrich) was used to pre-clear samples at 4°C for 2 hours. Anti-HA affinity matrix (Roche) was then added to each sample (25 µl packed resin per ml of protein sample) and incubated at 4°C overnight. For eluting immuno-precipitated proteins, resin was washed 3 times in wash buffer before being boiled in an equal volume of 2x protein loading (SDS loading) buffer (100 mM Tris-HCl pH 6.8, 4% w/v SDS, 0.2% w/v bromophenol blue, 20% v/v glycerol, 100 mM DTT). Samples were separated from the resin by filtration, resolved using SDS-PAGE and visualised followed by colloidal coomassie G-250 staining. Protein bands were then manually excised for analyses by LC/MS/MS (see materials and methods section 2.19).

### 2.26. Ubiquitin conjugation assay

All reactions were carried out in 25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer pH 7.4 supplemented with protease inhibitors (1 mM phenylmethanesulfonylfluoride (PMSF), 0.3 µg/ml trypsin inhibitor and 0.025 TIU/ml aprotinin). Ubiquitin conjugation reactions were carried out using the parkin auto-ubiquitination kit according to the manufacturer’s instructions (Boston Biochem). Reactions were incubated at 37°C for 1 hour. Proteins were separated by SDS-PAGE and analysed by binding streptavidin-HRP at 1:20000 (Pierce). Horseradish peroxidase (HRP) is a plant-derived enzyme that can oxidise chemiluminescent
substrates. When enhanced chemiluminescent substrate, ECL is oxidised by HRP, light is produced as a by-product.

2.27. Gradient ammonium sulphate precipitation of ES protein

Ammonium Sulphate (4 M saturated (NH₄)₂SO₄) pH 7.4 was added at room temperature to 100 µg of ES (in 25 mM Hepes buffer pH 7.4 supplemented with protease inhibitor cocktail (Roche) to achieve incremental increases in salt concentration. Precipitated proteins were centrifuged at 13,000 rpm and the supernatant was transferred to a new eppendorf containing a specified increased amount of ammonium sulphate. This process was repeated until 90% of salt saturation (4 M) was achieved. Protein precipitates were resuspended in 25 mM Hepes buffer pH 7.4 with protease inhibitors. Proteins were then transferred to PVDF membrane and analysed by immuno-blot using anti-NEDD4-L. In a scale up reaction, 500 µg of ES was fractionated at salt concentrations of 34, 44, 54, 59-70 (1% increase each time) and 80% saturation. Protein pellets were resuspended in 80 µl of 1x SDS loading dye. 6.25 µl was analysed by SDS-PAGE and immuno-blot using anti-NEDD4-L to locate fractions containing the reactive protein. Fractions containing proteins of interest were then separated by SDS-PAGE (95%) and visualised by silver staining using the SilverQuest Silver Staining kit (Life Technologies according to the manufacturer's instructions). These were separated alongside 5% of the same samples that were analysed by immuno-blot using anti-NEDD4-L. Proteins of interest were manually excised, using the immuno-blot signal as a guide. These were analysed by LC/MS/MS as previously described in section 2.19.

2.28. Immuno-fluorescence

C2C12 myoblasts were seeded into µ-Slide 8 well, ibiTreat, tissue culture treated plates (ibidi) and differentiated as described in section 2.5. After induction of TsUBE2L3 expression, cells were washed in ice cold 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄.2H₂O, 2 mM KH₂PO₄ pH 7.4) and then processed for immuno-fluorescence using a series of treatments: 20 minutes fixation at room temperature in 4% paraformaldehyde in 1x PBS, 3 x washes in 1x PBS, 10 minutes permeabilisation at room temperature in 0.2% v/v Triton-X and 3 x washes in 1x PBS. Samples were then blocked overnight at 4°C in 3% BSA w/v, 5% goat serum in 1x PBS. Samples were then incubated in primary antibodies diluted in antibody buffer (3% BSA w/v in 1x PBS 0.05% Tween-20 v/v) for 1 hour at room temperature, followed by secondary
antibodies diluted in antibody buffer for 1 hour at room temperature. Samples were washed 3 times in 1x PBS before incubation with Hoechst 33342 nucleic acid stain (Invitrogen) at a 1:20,000 dilution for 15 minutes at room temperature. A small volume of 1x PBS was added to each well of the plate and samples were visualised using a Leica SP5 MP/FLIM inverted confocal microscope fitted with water immersion lenses.

2.29. $^{35}$S Radiolabelling of C2C12 proteins

C2C12 myotubes were induced using doxycycline. At 20 hours post-induction growth media without cysteine or methionine was added for 1 hour at 37°C (Gibco). The proteasome inhibitor MG132 (25 µM, Z-Leu-Leu-Leu-al, Sigma-Aldrich) and $^{35}$S labelled cysteine and methionine amino acids (100uCi/ml PerkinElmer) were added to the growth media for 3 hours at 37°C. Cells were harvested and lysed and radiolabelled proteins were spotted onto Whatman filter paper that was then submerged in liquid scintillation cocktail (PerkinElmer) inside glass scintillation vials. The amount of $^{35}$S radiolabelled protein was measured for 30 seconds as counts per minute (CPM) using a Wallac 1400 DSA scintillation counter. Radiolabelled proteins were then used for radio-immuno-precipitation experiments as described below. After immuno-precipitation experiments, radioactive proteins were separated by SDS-PAGE. Gels were processed for autoradiography (at room temperature) by incubating in 100% v/v DMSO for 1 hour, fresh 100% v/v DMSO for 30 minutes, followed by 20% v/v PPO (polyphenyl ether) in DMSO for 30 minutes, then fresh 20% v/v PPO in DMSO overnight. Gels were then washed multiple times in water before being dried onto Whatman filter paper using a Bio-Rad Gel Dryer 583. Proteins were visualised by autoradiography.

2.30. Immuno-precipitations

Cells were lysed in RIPA buffer (1% v/v Triton X-100, 1% w/v sodium deoxycholate, 0.1% w/v SDS, 50 mM Tris pH 7.4, 150 mM NaCl and protease inhibitor cocktail (Roche). For immuno-precipitations protein concentrations were normalised based on a BCA assay (Pierce). For radio-immuno-precipitations protein amounts were normalised based on radioactivity (CPM). Protein G sepharose resin (Sigma-Aldrich) was used to pre-clear 300-500 µg of proteins in RIPA lysis buffer at 4°C for 2 hours. Protein G was removed and samples were diluted out to a volume of 1 ml using wash buffer (RIPA buffer without deoxycholate). Antibodies were added to Protein G
(25 µl packed resin per ml proteins) and incubated for 45 minutes at 4°C. Protein G/antibody mixture was added to proteins and samples were incubated for 45 minutes at 4°C. Resin containing protein complexes were washed 4 times in wash buffer and re-suspended in 25 µl of 2x protein (SDS) loading buffer (100 mM Tris-HCl pH 6.8, 4% w/v SDS, 0.2% w/v bromophenol blue, 40% v/v glycerol, 20% v/v, 100 mM DTT). Proteins were eluted by boiling samples for 5 minutes at 97°C. Proteins were separated by SDS-PAGE followed by either immuno-blot analysis or autoradiography.

2.31. HA immuno-precipitation

Cells were lysed in RIPA buffer supplemented with EDTA-free protease inhibitor cocktail (Roche). A BCA assay (Pierce) was used to normalise protein concentrations. Samples (300-500 µg) were pre-cleared using protein G agarose (Sigma) for 45 minutes at 4°C. Samples were then incubated with anti-HA high affinity matrix (25 µl packed resin per ml of protein sample, Roche) for 45 minutes at 4°C and the matrix was washed in NP-40 buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% NP-40 v/v) supplemented with protease inhibitors.

2.32. Re-immuno-precipitation

Proteins were eluted from the precipitation matrix by boiling in 1% SDS w/v, 1 mM DTT for 2 minutes at 97°C and resuspended in 1 ml of NP-40 buffer supplemented with protease inhibitors. Samples were then incubated with protein G agarose that had been pre-incubated with the antibody in NP-40 buffer for 45 minutes at 4°C. Protein G was then washed 4 times in NP-40 buffer, re-suspended and boiled at 97°C in protein (SDS) loading buffer for 5 minutes to elute proteins. Proteins were separated by SDS-PAGE and analysed by immuno-blot.

2.33. TUBE 2 assays

Enrichment of polyubiquitinated proteins was carried out using TUBE 2 agarose resin (tebu-bio) according to the manufacturer’s instructions. Proteins were then processed using the re-immuno-precipitation protocol above before being separated by SDS-PAGE and analysed by immuno-blot.
2.34. Click-IT EdU proliferation assays

EdU (5-ethyl-2'-deoxyuridine, 10 µM) was added to C2C12 cell growth media for the indicated amount of time and incubated at 37°C. For immunofluorescence, cells were grown and processed directly in ibidi u-Slide 8-well plates (ibidiTreat, tissue culture treated). For flow cytometry myoblast cells were trypsinised and washed in 1x PBS before being transferred to a 96-well v bottom plate where after each incubation and wash step plates were centrifuged at 500xg for 3 minutes. All samples were processed using immuno-fluorescence fixation and permeabilisation techniques (section 2.28 and 2.35). The Click-iT EdU Alexa Flour 647 Imaging kit (Life Technologies) was then used according to the manufacturer’s instructions. Samples were either visualised by confocal microscopy or analysed by flow cytometry.

2.35. Flow cytometry

Cell samples were trypsinised and then transferred to a 96-well v bottom plate. Firstly cells were fixed in 4% paraformaldehyde in 1x PBS (in suspension). All following treatments were carried out in suspension where after each incubation and wash step, plates were centrifuged at 500xg for 3 minutes. Cells were permeabilised at room temperature in 0.2% v/v Triton-X and washed 3 times in 1x PBS. Samples were then processed using the Click-iT EdU Alexa Flour 647 kit for Flow Cytometry (Life Technologies) according to the manufacturer’s instructions. Finally cells were washed twice in 0.5 ml FACS buffer (1x PBS, 1% w/v BSA, 0.05% v/v sodium azide, 2 mM EDTA) using centrifugation at 900 xg at 4°C for 3 minutes between washes. They were then resuspended in FACS buffer and transferred to FACS tubes. Samples were analysed using a BD Fortessa Analyser with lasers and filters 633-670/14 for detection of Alexa Flour 647. Data was analysed using FlowJo software (TreeStar).

2.36. Immuno-blot analyses

Proteins were transferred onto PVDF membrane using semi-dry transfer electrophoresis in transfer buffer (25 mM Tris-HCL pH 7.6, 192 mM glycine, 20% methanol). The PVDF membrane was blocked for 1 hour at room temperature in 5% w/v non-fat milk with PBS-Tween (0.1% v/v, PBST). Membranes were incubated in primary antibodies in 2% w/v non-fat milk with PBST for 90 minutes at room temperature or 4°C overnight. Membranes were washed in PBST 3 x 15 minutes at
room temperature before being incubated in secondary antibodies in 2% w/v non-fat milk with PBST for 90 minutes at room temperature. Membranes were washed in PBST 3 x 15 minutes at room temperature. Proteins were visualised using enhanced chemiluminescence substrate (ECL, PerkinElmer) on a Bio-Rad ChemiDoc XRS Gel Photo Documentation System. Analysis of the presence of biotinylated proteins was carried out by binding streptavidin-HRP to proteins transferred onto PVDF membranes (as described above for a regular immuno-blot, except membranes were blocked with 5% w/v BSA in PBST and probed with streptavidin-HRP diluted in 5% w/v BSA in PBST). For the conjugation of horseradish peroxidase (HRP) to antibodies, the Lightening Link Horseradish Peroxidase kit (Innova Biosciences) was used according to the manufacturer’s instructions.

2.37. Statistics

Unless otherwise stated data values are expressed as the mean with the standard error of the mean (SEM) represented by error bars. Unless otherwise stated the Student’s t-test was used to assess statistical significance of data. Results were taken to be statistically significant where p<0.05 unless otherwise stated.

2.38. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Resolving gels (8-12% acrylamide) were prepared using a mixture of the following reagents: 1.5M Tris-HCl pH 8.8, 0.4% w/v SDS, 30% acrylamide, 10% w/v ammonium persulphate (APS) and TEMED (tetramethylethlenediamine). Stacking gels (4% acrylamide) were prepared using the following reagents: 0.5M Tris-HCl pH 6.8, 0.4% w/v SDS, 30% acrylamide, 10% w/v ammonium persulphate (APS) and TEMED. The volume of each reagent varied depending on the percentage of acrylamide in the mixture as presented in Table 2.2 below (shows amounts required for 1x 1.0 mm Novex cassette). Once prepared, gels were either poured into Novex 1.0 mm empty gel cassettes (Life Technologies) or, for midi gels, between large glass plates. If gradient gels were required, pre-cast 4-12% gradient gels were purchased from Life Technologies (NuPAGE 4-12% Bis-Tris Mini Gels). Prepared gels were submerged in SDS-PAGE running buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1% SDS) in an SDS-PAGE tank. Protein samples were boiled in 1x protein (SDS) loading buffer (200 mM Tris-HCl pH 6.8, 8% w/v SDS, 0.4% w/v bromophenol blue, 40% v/v glycerol, 200 mM DTT) for 6 min at 94°C and loaded into the wells of
the gel. PageRuler Plus pre-stained protein ladder (Thermo Scientific) was loaded alongside the sample proteins to serve as a protein size marker. Gels were subjected to electrophoresis using a constant voltage (V) of 90V, which was then increased to 120 V after stacking was complete. Gels were then either stained using coomassie or silver-stain or transferred onto PVDF membranes for immuno-blots (section 2.36).

Table for SDS-PAGE gel preparation

<table>
<thead>
<tr>
<th>acrylamide</th>
<th>8%</th>
<th>10%</th>
<th>12%</th>
<th>4% (stacking)</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>2.7 ml</td>
<td>2.3 ml</td>
<td>1.9 ml</td>
<td>1.3 ml</td>
</tr>
<tr>
<td>acrylamide (30%)</td>
<td>1.6 ml</td>
<td>2 ml</td>
<td>2.4 ml</td>
<td>0.265 ml</td>
</tr>
<tr>
<td>1.5M Tris-HCl pH 8.8 with 0.4% SDS</td>
<td>1.6 ml</td>
<td>1.6 ml</td>
<td>1.6 ml</td>
<td>0.4 ml</td>
</tr>
<tr>
<td>10% APS</td>
<td>30 µl</td>
<td>30 µl</td>
<td>30 µl</td>
<td>30 µl</td>
</tr>
<tr>
<td>TEMED</td>
<td>6 µl</td>
<td>6 µl</td>
<td>6 µl</td>
<td>3 µl</td>
</tr>
</tbody>
</table>

Table 2. 2 Table of SDS-PAGE gel reagent composition

Table shows the volumes required of each reagent to make one resolving gel (with the indicated acrylamide percentage) using Novex 1.0 mm cassettes (Life Technologies) plus the 4% stacking gel.
### 2.39. Antibodies

<table>
<thead>
<tr>
<th>Protein/epitope</th>
<th>Host animal</th>
<th>Clonality</th>
<th>Working dilution (1::)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>strepavidin-HRP</td>
<td>n/a</td>
<td>n/a</td>
<td>10000</td>
<td>Pierce</td>
</tr>
<tr>
<td>His-HRP</td>
<td>n/a</td>
<td>n/a</td>
<td>5000</td>
<td>Pierce</td>
</tr>
<tr>
<td>Primary antibodies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLAG-HRP</td>
<td>mouse</td>
<td>monoclonal M2</td>
<td>1000</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>HA-HRP</td>
<td>rat</td>
<td>monoclonal, 3F10</td>
<td>1000</td>
<td>Roche</td>
</tr>
<tr>
<td>HA</td>
<td>rat</td>
<td>monoclonal, 3F10</td>
<td>100 (IFA)</td>
<td>Roche</td>
</tr>
<tr>
<td>alphatheta-Tubulin</td>
<td>mouse</td>
<td>monoclonal</td>
<td>1000 (IB) 500 (IFA)</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>anti-WDR48</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>1000</td>
<td>(Sigma Aldrich)</td>
</tr>
<tr>
<td>NEDD4-L</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>500</td>
<td>Protein Tech Group</td>
</tr>
<tr>
<td>5'NT</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>400</td>
<td>GI: 22656349, supplied by Kleoniki Gounaris</td>
</tr>
<tr>
<td>UBE2N</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>1000</td>
<td>LifeSpan Bioscience</td>
</tr>
<tr>
<td>ubiquitin</td>
<td>rabbit</td>
<td>monoclonal</td>
<td>1000</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>UBE2L3</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>500</td>
<td>Epitomics</td>
</tr>
<tr>
<td>p53</td>
<td>rabbit</td>
<td>monoclonal</td>
<td>500 (IB) 50 (IP)</td>
<td>Merck Millipore</td>
</tr>
<tr>
<td>cyclin D1</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>1000</td>
<td>Cell signalling</td>
</tr>
<tr>
<td>cyclin E</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>500</td>
<td>Abcam</td>
</tr>
<tr>
<td>cyclin B1</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>1000</td>
<td>Cell signalling</td>
</tr>
<tr>
<td>pH3</td>
<td>rabbit</td>
<td>polyclonal</td>
<td>500</td>
<td>Pierce</td>
</tr>
<tr>
<td>myogenin</td>
<td>mouse</td>
<td>monoclonal</td>
<td>1000</td>
<td>Merck Millipore</td>
</tr>
<tr>
<td>MHC</td>
<td>mouse</td>
<td>monoclonal</td>
<td>500</td>
<td>Merck Millipore</td>
</tr>
<tr>
<td>Secondary antibodies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse-HRP</td>
<td>Goat</td>
<td>polyclonal</td>
<td>10000</td>
<td>Thermo Fisher Scientific</td>
</tr>
<tr>
<td>Rabbit-HRP</td>
<td>Mouse</td>
<td>monoclonal</td>
<td>7500</td>
<td>Sigma Aldrich</td>
</tr>
</tbody>
</table>

Table 2.3 Antibodies and probes used throughout the study
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Use</th>
<th>Details</th>
<th>Antibiotic resistance cassette (bacterial/mammalian)</th>
<th>Source/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pPET28a+</td>
<td>Bacterial expression vector</td>
<td>T7 promoter cloning for N or C-terminal His-tag. Multiple cloning site (MCS).</td>
<td>kanamycin/none</td>
<td>Novagen</td>
</tr>
<tr>
<td>pGEMTeasy</td>
<td>Cloning vector</td>
<td>TA cloning and sequencing vector (T7/SP6 sequencing prime binding) Multiple cloning site (MCS).</td>
<td>ampicillin/none</td>
<td>Promega</td>
</tr>
<tr>
<td>pCMV-VSV-G</td>
<td>Lentivirus accessory plasmid for virus production</td>
<td>Encodes lentiviral protein VSV-G, lentivirus coat protein.</td>
<td>ampicillin/none</td>
<td>Hidde Ploegh</td>
</tr>
<tr>
<td>pCMV-Gag/Pol</td>
<td>Lentivirus accessory plasmid for virus production</td>
<td>Encodes lentiviral proteins Gag (lentivirus matrix and capsid proteins), Reverse Transcriptase, Integrate and the HIV Protease.</td>
<td>ampicillin/none</td>
<td>Hidde Ploegh</td>
</tr>
<tr>
<td>pCMV-Tat</td>
<td>Lentivirus accessory plasmid for virus production</td>
<td>Encodes lentiviral protein Tat.</td>
<td>ampicillin/none</td>
<td>Hidde Ploegh</td>
</tr>
<tr>
<td>pCMV-Rev</td>
<td>Lentivirus accessory plasmid for virus production</td>
<td>Encodes lentiviral protein Rev.</td>
<td>ampicillin/none</td>
<td>Hidde Ploegh</td>
</tr>
<tr>
<td>pCSGW-IRES-tdTomato</td>
<td>Mammalian expression vector</td>
<td>Has lentiviral packaging signal (ψ), IRES tdTomato, multiple cloning site (MCS), Rev Response Element (RRE), Central Polypurine Tract (cPPT), and HIV-1 LTR promoter.</td>
<td>ampicillin/none</td>
<td>Ilaria Nisoli, Hugh Brady Group</td>
</tr>
<tr>
<td>pLVX-Tight-Puro</td>
<td>Mammalian expression vector</td>
<td>Inducible system expression vector. Has multiple cloning site (MCS) under control of P&lt;sub&gt;P&lt;sub&gt;tight&lt;/sub&gt; promoter. Has ψ, RRE and cPPT.</td>
<td>ampicillin/puromycin</td>
<td>Clontech Lenti-X™ Tet-On Advanced</td>
</tr>
</tbody>
</table>
Chapter 3:
Characterisation of TsUCH37, a deubiquitinating enzyme expressed by *T. spiralis* muscle larvae
Chapter 3:

Characterisation of TsUCH37, a deubiquitinating enzyme expressed by *T. spiralis* muscle larvae

Some of the data discussed in this chapter was published in 2011. After publication the project was continued and the results presented here comprise an expanded version [1]. This work also led to a collaboration with a group at Purdue University, and some of the data from this was published in 2013 [3].

During my Masters of Research (MRes) degree at Imperial College London, I carried out a 3-month project under the supervision of my PhD supervisor to be, Katerina Artavanis-Tsakonas. The main project aim was to investigate the various roles of the ubiquitin pathway during the infection of muscle tissue by the mammalian parasitic nematode, *T. spiralis*. I initially set out to look at the ubiquitin pathway of the parasite itself. During the project, I found that *T. spiralis* muscle larvae express various putative *T. spiralis* deubiquitinating enzymes (DUBs). The most abundant *T. spiralis* DUB identified was named TsUCH37 after its orthology to mammalian UCH37 (also known as UCHL5). The PhD project followed on from this, and chapter 3 presents the subsequent characterisation of TsUCH37.

Background

The ubiquitin proteasome system is essential for the regulation of many important cellular processes [5]. It is therefore implicated in a great number of diseases, especially cell cycle disorders such as cancer [227-229]. Therapeutic drugs that target the ubiquitin proteasome system are therefore in demand and some are already in clinical trials for human use [73,74]. The ubiquitin proteasome system is also implicated during infection by pathogenic organisms, where the pathogen may disrupt the host ubiquitin pathway to its own advantage during infection. Despite the absence of their own ubiquitin proteasome system, some viruses and bacteria express components of the ubiquitin pathway that manipulate host proteins [93,97,98]. Because parasites are most often eukaryotic, they have their own Ub proteasome system, the function of which is essential for their homeostasis and survival, thus parasite-derived ubiquitin proteasome components potentially make good drug targets for infection [230,231]. Most of the examples of drugs that target the parasite ubiquitin system are proteasome inhibitors. However, because the structure of the proteasome is so highly conserved, many proteasome inhibitors are non-selective and potentially toxic. In order to develop
inhibitors with greater selectivity, interest in targeting ubiquitin conjugation and hydrolysis enzymes has expanded.

I was interested in exploring the role of the *T. spiralis* ubiquitin pathway during the muscle stages of trichinellosis in order to use *T. spiralis* as a model for further investigating the potential of parasite-derived ubiquitin pathway proteins as drug targets. I began by investigating whether or not *T. spiralis* muscle larvae express active ubiquitin pathway machinery. During the MRes project, I utilised a ubiquitin-based inhibitor probe, HA-Ub-VME, that identifies deubiquitinating enzymes (DUBs). A detailed explanation of the mechanism of the probe can be found in chapter 4, background. *T. spiralis* muscle larvae were reacted with HA-Ub-VME and probe-protein complexes were immuno-precipitated using anti-HA antibodies. These complexes were further analysed by tandem mass spectrometry (LC/MS/MS). This method identified 5 different putative *T. spiralis* DUBs (MRes report 2010, [1]). The most abundant DUB identified was a conserved ubiquitin C-terminal hydrolase domain protein, orthologous to the highly conserved 37 kDa DUB, UCH37 (also known in mammals as UCHL5). The orthologous *T. spiralis* sequence is annotated in the draft genome (NCBI Ref Seq) as a putative ubiquitin carboxyl- hydrolase ubh-4, GI:339238735 [232]. This sequence did not correlate with multiple orthologous sequences, and the accuracy of the annotation was disputed.

Mammalian UCH37 was first identified as proteasome associated DUB [34,35]. UCH37 can be nuclear or cytoplasmic, and is responsible for the hydrolysis of Lys-48-linked polyubiquitin [32,233]. UCH37 is the only UCH domain DUB that has been shown to associate with the 19S subunit of the 26S proteasome [33]. UCH37 interacts directly with the proteasome subunit ADRM1. In turn, ADRM1 binds to Rpn2, a protein that makes up the 19S base of the proteasome. In vitro, the ubiquitin hydrolysis activity of the full length human UCH37 is enhanced by the addition of the ADRM1 protein. This enhancement of UCH37 activity is not observed when ADRM1 is added to a truncated version of UCH37 that lacks the C-terminal portion of the protein that has been shown to directly bind to ADRM1 [32,33,233]. Human UCH37 and ADRM1 therefore both play a role in the proteosomal degradation of protein substrates. HsUCH37 is expressed in many different cell types and tissues and knockout of UCH37 in mice is lethal at the embryonic stage [36]. The orthologue of this protein in *Plasmodium falciparum* PfUCH54 (named because it weighs 54 kDa rather than 37) was also identified using the human Ub-based inhibitor probe, HA-Ub-VME [102].

Having identified TsUCH37, an abundant *T. spiralis* DUB for which a function has been assigned for numerous orthologues, it was of interest to characterise the function of the
Chapter 3: Results

*T. spiralis* protein, and ascertain whether or not its expression is essential for the survival of the parasite. Barring some pioneering studies in carried out in *Schistosoma mansoni*, the majority of parasitic nematodes, including *T. spiralis* have proven thus far largely resistant to common techniques of genetic manipulation [234,235]. As yet there is no obvious explanation for the resistance observed [236]. The methodological possibilities for studying the functions of *T. spiralis* proteins are therefore limited. In addition, the *T. spiralis* life cycle does not include a free-living stage. This means that the muscle larvae can only be cultured for a limited amount of time after isolation from muscle tissues before they die (in my experience 4-6 days). For these reasons combined, the endogenous expression and function of a *T. spiralis* protein cannot be studied using genetic manipulation. It was therefore decided that studying the function of the *T. spiralis* DUB, TsUCH37, would largely require the cloning of the TsUCH37 gene followed by recombinant protein expression and biochemical analyses.

This chapter presents the analysis of the coding sequence of TsUCH37, followed by the expression, purification and functional characterisation of the recombinant protein. The potential of TsUCH37 as a drug target was then investigated using drug inhibitor assays. Finally, results prompted a collaboration with a crystallography group at Purdue University (US) [3].
Chapter 3: Results

3.1. Finding the true open reading frame for a *T. spiralis* DUB orthologue

The putative coding sequence of for TsUCH37 translates into a protein sequence of 900 amino acids (GI: 339238735 - 2703 bp). This was aligned with UCH37 orthologues from 6 other species using a MUSCLE alignment algorithm (Figure 3.1) [237]. The putative UCH37 sequence from the filarial parasitic nematode *Brugia malayi* (GI: 158597396) showed the greatest identity to the *T. spiralis*. Also aligned were UCH37 orthologues of (in order of descending identity to the *T. spiralis* orthologue) the *M. musculus* (GI: 4878011), *H. sapiens* (GI: 4877999), *C. elegans* (GI: 71981272), *D. melanogaster* (also known as p37A, GI:17648095) and *S. cerevisiae* (YUH1, GI:151945164). All of which have been confirmed as functional ubiquitin C-terminal hydrolases [36,238-240]. Alignment revealed that the annotated putative *T. spiralis* sequence was considerably longer than its orthologues, with an N-terminal extension of almost 600 amino acids. When analysed by SMART software (Simple Modular Architecture Research Tool, [210,211]) 2 additional Pfam domains were identified: a SEC14 domain and a PRP38 domain, both located within the N-terminal extension. The peptides identified by LC/MS/MS only aligned within the C-terminal region containing the peptidase_C12 domain and not within the N-terminal extension (MRes report 2010 [1]). This information led into question the accuracy of the gene annotation. To be able to clone the gene for TsUCH37, the true open reading frame had to first be confirmed.

The draft annotation of the *T. spiralis* genome was carried out using the gene prediction algorithm FGENESH without using any prior knowledge of potential gene or transcript arrangement in *T. spiralis*, or ‘hints’ from the gene or transcript arrangement of similar organisms (FGENSH ‘de novo’ [214,232]). The *T. spiralis* genomic contig containing the TsUCH37 gene (The Genome Institute at Washington University, contig 1.2, bases 34000-37000) was therefore re-analysed using 6 additional gene prediction algorithms: FGENESHC using the original EST hit (identified by the original LC/MS/MS - GI:157958881, before the draft genome was available – [1]) as a hint, FGENESH+ using the C. elegans UBH4 as a hint, AUGUSTUS using the EST as a hint, AUGUSTUS using the EST and C. elegans UBH4 as hints, GenemarkHMM and SNAP (Figure 3.2). Gene prediction analyses were carried out in collaboration with Eliseo Papa of Harvard/MIT. Out of all 7 algorithms, 5 predicted the same start codon, which aligned well with all 6 orthologous UCH37 protein sequences when translated (Figure 3.5). This consensus start codon did not agree with the putative annotation in the database (GI:339238735, based on FGENSH ‘de novo’ prediction). The consensus
prediction did not include any of the upstream portion, which translates to contain the SEC14 and PRP38 domains. Within bases 34000-37000 of contig 1.2, AUGUSTUS predicted an additional gene upstream of the UCH37 gene, and GenemarkHMM and SNAP predicted 2 additional genes in the upstream portion of the contig, rather than predicting extra exons of the same gene as was the case for FGENESH de novo. The predicted end of the coding sequence varied slightly between each programme. Overall, when translated the AUGUSTUS sequence prediction of 930 bp, 7 exons and 6 introns, aligned most closely with the other UCH37 orthologues (Figure 3.5). This sequence was therefore hypothesised to be the true open reading frame of the *T. spiralis* UCH37.

In order to experimentally confirm the true open reading frame for the *T. spiralis* UCH37, rapid amplification of cDNA ends (RACE) experiments were carried out. This technique is used to confirm the full sequence of an mRNA transcript for a particular gene. Total RNA is isolated from a sample and, for 5' RACE, the 5' phosphate group is removed from degraded mRNA, rRNA, tRNA and DNA by calf intestinal phosphatase (CIP) (Figure 3.3). The CAP structure is then removed from intact mRNA using tobacco acid pyrophosphatase (TAP), leaving a 5' phosphate group on only the intact mRNA transcripts. RNA ligase is then used to ligate the 5' RACE adapter oligo onto the intact mRNA transcript. This is reverse-transcribed to cDNA for the amplification by PCR using a gene specific reverse primer and a primer that anneals to the 5' RACE oligo, thus amplifying the 5’ end, including the start codon, of the transcript. For 3' RACE, a 3' RACE adapter oligo is ligated onto the poly-adenosine tail at the 3’ end of an intact mRNA transcript. This is reverse-transcribed to cDNA for the amplification by PCR using a gene specific forward primer and a primer that anneals to the 3’ RACE oligo, thus amplifying the 3’ end, including the stop codon, of the transcript. These DNA fragments can then be sequenced to confirm the full coding sequence.
The NCBI annotated TsUCH37 (GI:339238735, FGENESH 2703 bp) sequence was aligned with UCH-L5/37 orthologous sequences from 6 other species using a Geneious alignment algorithm with Geneious software (Drummond, geneious.com). Sequences in order of descending identity to the *T. spiralis* orthologue: *M. musculus* (GI:229577283), *H. sapiens* (GI: 312922359), *C. elegans* (GI:71981272), *D. melanogaster* (GI:17648095) and *S. cerevisiae* (GI:151945164). Blocks indicate nucleotide bases and lines indicate gaps in alignment. The TsUCH37 sequence was analysed by SMART software identifying 3 Pfam domains: a SEC14 domain, a PRP38 domain (both located within the N-terminal extension) and a ubiquitin C-terminal hydrolase (UCH) peptidase_C12 domain.
Gene predictions for TsUCH37

The *T. spiralis* genomic contig (contig 1.2, bases 34000-37000, The Genome Institute at Washington University) was analysed using 6 gene prediction algorithms: FGENESHC using the EST GI:157958881 as a hint, FGENESH+ using the C. elegans UBH4 as a hint, AUGUSTUS using the EST as a hint, AUGUSTUS using the EST and UBH4 as hints, GenemarkHMM and SNAP. The FGENESHC predicted start codon agrees with all subsequent gene prediction programs employed, and the most common (consensus) gene prediction is highlighted.
Total RNA was isolated from *T. spiralis* muscle larvae. **A.** For 5’ RACE, the 5’ phosphate group was removed from degraded mRNA, rRNA, tRNA and DNA by calf intestinal phosphatase (CIP). The CAP structure was then removed from intact mRNA using tobacco acid pyrophosphatase (TAP), leaving a 5’ phosphate group on intact mRNA transcripts. The 5’ RACE adapter oligo was ligated onto the intact mRNA transcript using RNA ligase. This was reverse-transcribed to cDNA and amplified by PCR using a gene specific reverse primer and a primer that anneals to the 5’ RACE oligo, thus amplifying the 5’ end, including the start codon, of the transcript. **B.** For 3’ RACE, a 3’ RACE adapter oligo was ligated onto the poly-adenosine tail at the 3’ end of an intact mRNA transcript. This was reverse-transcribed to cDNA and amplified by PCR using a gene specific forward primer and a primer that anneals to the 3’ RACE oligo, thus amplifying the 3’ end, including the stop codon, of the transcript.
For RACE analysis of the *T. spiralis* UCH37 transcript, two gene-specific reverse primers (GSRP) were designed, one using the annotated FGENESH-derived sequence (primer 1) and the other using the AUGUSTUS-derived sequence (primer 3). All theoretical primer-binding sites are illustrated in Figure 3.4 A. Each was designed to anneal 332 bp and 321 bp (respectively) from the predicted start codon. The AUGUSTUS GSRP however, should also anneal to the FGENESH sequence albeit further downstream if the prediction is correct (2094 bp from the predicted start codon).

A gene-specific forward primer (GSFP) was also designed for amplification of the 3’ end (primer 4). This was designed to anneal to a sequence that is in agreement with both the FGENESH prediction and the AUGUSTUS prediction. This was predicted to amplify a fragment of the gene 396 bp from the AUGUSTUS predicted stop codon, and 425 bp from the FGENESH predicted stop codon. Forward and reverse (end-to-end) primers were also designed on the 930 bp AUGUSTUS predicted sequence (primers 2 and 5, Figure 3.4 A).

Total RNA was extracted and purified from *T. spiralis* muscle larvae. Control primers to amplify a 313 bp fragment of the transcript of the *T. spiralis* GM2-activator protein were kindly supplied by Murray Selkirk [199]. These were used in a PCR reaction to test the quality of the RNA. AUGUSTUS-based gene specific primers were also tested by PCR. These reactions were separated by agarose gel electrophoresis and visualised under UV light after incubation with ethidium bromide. (Figure 3.4 B). End-to-end primers 2 and 5 amplified a fragment of DNA of an AUGUSTUS-predicted length just under 1000 bp. However when using the AUGUSTUS GSRP (primer 3) with the forward primer (2), 3 bands between 300 and 500 bp were observed. These may have amplified as a result of gDNA contamination of the *T. spiralis* RNA. The fragment amplified by primers 4 and 5 also appeared slightly longer than the AUGUSTUS-predicted length of 396 bp.

RACE was then carried out using the FirstChoice RNA Ligase Mediated Rapid Amplification of cDNA Ends (RLM-RACE) kit from Ambion. *T. spiralis* RNA was processed as instructed by the manufacturer and reverse transcribed to cDNA (using the Superscript II Reverse Transcriptase kit from Invitrogen) before undergoing PCR followed by nested PCR. Nested PCR was carried out using the RACE 5’ primers that anneal to the 5’ RACE oligo, paired with primers 1, 3 and 5 (Figure 3.4 C). Reactions were separated by agarose gel electrophoresis and visualised under UV light after incubation with ethidium bromide. No band was observed when the FGENESH GSRP (primer 1) was paired with the RACE oligo 5’ primers. Again, 3 bands were observed when the RACE 5’ oligo primer was paired with the AUGUSTUS GSRP (primer 3), the smallest of which was most abundant and at the expected size of around 300 bp.
These PCR reactions were carried out using a 40 second polymerase extension time, since the predicted fragments were to be 332 bp and 321 bp. The RACE 5’ primers were then paired with the AUGUSTUS predicted end primer (primer 5) and again the AUGUSTUS predicted GSRP (primer 3). This time, 2 minutes was used for polymerase extension time. This was to allow for amplification of a potentially longer sequence predicted by FGENESH. Using primer 5 a band was observed that corresponded to the AUGUSTUS predicted full-length sequence of 930 bp, rather than the FGENESH predicted full length of 2703 bp. With primer 3 a band that corresponded to the AUGUSTUS predicted 321 bp was observed, rather than that of 2094 predicted by FGENESH. One additional band was also observed at approximately 700-800 bp that may have been a result of *T. spiralis* RNA or cDNA degradation. The fragment that was amplified using primer 3 with a 2-minute extension time was purified and ligated into pGEMTeasy using a restriction enzyme free ligation method known as TA cloning. Briefly, linearised pGEMTeasy is modified to contain a thymine (T) DNA base overhang on each 5’ end of the complementary vector strands. The polymerase used for the PCR, leaves adenine (A) overhangs on the 3’ end of the PCR product. Since adenine and thymine are complementary DNA bases, the T and A will ligate to one another. pGEMTeasy containing the purified insert was sequenced by Beckman Coulter Genomics. The sequence obtained was 100% identical to the AUGUSTUS prediction of the first 321 bp of the coding sequence for the *T. spiralis* UCH37. This start site aligned well with the other UCH37 orthologues and thus the start codon was confirmed (Figure 3.5).

In order to amplify the 3’ end of the transcript, the same gene specific forward primer (primer 4) was used for both the AUGUSTUS and FGENESH predictions because both gene predictions are in agreement with this part of the coding sequence (Figure 3.4 D). PCR reactions were carried out using RNA that had been reverse transcribed for 3’ RACE as shown in Figure 3.3 B. As a control, primer 4 was paired with primer 5 to amplify the AUGUSTUS predicted end of the transcript. Primer 4 was also combined with the RACE 3’ primer and RACE 3’ nested primer to obtain a fragment containing the true stop codon. DNA was amplified in each sample. Using primers 4 and 5, a band of an expected size larger than 300 bp was observed. The band amplified using primer 4 and the RACE 3’ primer was very slightly larger than this. This fragment was manually excised from the gel, purified and ligated into pGEMTeasy using TA cloning as described above. The sequence of the insert was obtained (Beckman Coulter Genomics) and confirmed to be 100% identical to the AUGUSTUS predicted coding sequence for the 3’ end of the TsUCH37 gene. It was therefore confirmed, that the AUGUSTUS prediction, represented the true coding sequence for the *T. spiralis* UCH37.
Figure 3.5 shows how this sequence aligned well with orthologous sequences for UCH37, with a good level of agreement at the beginning and end of the protein sequence.
Figure 3.4

A. Primers 1-5 were designed to bind to either the AUGUSTUS predicted (930 bp) or the FGENESH de novo predicted (2703 bp) coding sequences of TsUCH37. RACE adapter oligos were ligated onto the 5' end and the 3' end of the mRNA transcript. RACE 5' and 3' primers were provided with the FirstChoice RLM-RACE kit from Ambion. Primer binding sites are indicated. Sizes of the expected DNA fragments, based on the distances from the end of the predicted transcripts to the primer-binding site, are indicated. RACE-PCR was carried out. All PCR reactions were separated by agarose gel electrophoresis and visualised under UV light after incubation with ethidium bromide. B. Primers 2-5 were tested on mRNA from *T. spiralis* muscle larvae. Control primers to amplify a 313 bp fragment the transcript of the *T. spiralis* GM2-activator protein were used as a positive control. C. 5' RACE-PCR was carried out using the RACE 5' primer and gene prediction-specific reverse primers 1, 3 and 5. A DNA fragment at a size corresponding to the AUGUSTUS-predicted start of the TsUCH37 transcript was purified for cloning and sequencing (red box). D. 3' RACE-PCR was carried out using the RACE 3' primer and gene prediction-specific forward primer 4. The DNA fragment amplified by primer 4 + the RACE 3' primer was purified for cloning and sequencing (red box).
Figure 3.5 Alignment of orthologues and RACE-PCR-confirmed coding sequence of TsUCH37

The NCBI annotated TsUCH37 (GI:339238735, FEGENSH, 2703 bp) sequence was aligned with UCH-L5/37 orthologous sequences from 6 other species using a Geneious alignment algorithm with Geneious software (Drummond, geneious.com). Sequences in order of descending identity to the T. spiralis orthologue: M. musculus (GI:229577283), H. sapiens (GI:312922359), C. elegans (GI:71981272), D. melanogaster (GI:17648095) and S. cerevisiae (GI:151945164). This alignment was then re-aligned with the RACE-PCR-confirmed coding sequence for TsUCH37 (the same sequence as predicted by AUGUSTUS gene prediction software). Blocks indicate nucleotide bases and lines indicate gaps in alignment. The TsUCH37 was analysed by SMART software identifying 3 Pfam domains: a SEC14 domain and a PRP38 domain, both located within the N-terminal extension, and a ubiquitin C-terminal hydrolase-specific peptidase_C12 domain.
3.2. Cloning of the confirmed open reading frame of TsUCH37 and expression and purification of recombinant 6His-TsUCH37

With the true open reading frame of TsUCH37 confirmed it was possible to clone the coding sequence and express recombinant protein, which would allow further biochemical analyses of the function of the parasitic DUB.

The translated open reading frame of the *T. spiralis* UCH37 orthologue contains 309 amino acids, forming a protein with a predicted molecular weight of 35.2 kDa and a pI of 4.42 (Figure 3.6 A). Residues 5-209 span the peptidase_C12 domain and residues C85, H161 and D176 comprise the catalytic triad that is common to cysteine proteases. The positions of these catalytic residues are conserved amongst UCH domain DUBs. When aligned with the human UCH37, the sequences are 45.9% identical (Figure 3.6 B). Primers were designed to clone the confirmed coding sequence minus the start codon, with a 5’ BamHI restriction site and a 3’ NotI restriction site from *T. spiralis* cDNA (primer Table). This was digested and ligated into the pPET28a(+) bacterial expression vector. The complementary BamHI and NotI restriction sites of the multiple cloning site ensured that the sequence could be transcribed in frame with a start codon and an N-terminal His-tag (6 x histidine residues). The predicted molecular weight of 6His-TsUCH37 was 38.8 kDa (Figure 3.6 C). The confirmed coding sequence for TsUCH37 with an N-terminal His-tag was cloned for expression in a bacterial system, for purification and activity profiling.

Bacteria do not have a ubiquitin proteasome system, and the laboratory *E. coli* strain used for expression of TsUCH37 does not express DUBs. This allowed the purification of 6His-TsUCH37 without contaminating DUB proteins that would confuse activity profiling. The His-tag was added to the N-terminal end of TsUCH37, because studies have shown that the C-terminus of UCH37 is important for the association with the interaction partner ADRM1 and the proteasome. The plasmid was transformed into competent *E. coli* and bacteria were cultured and induced to express the 6His-TsUCH37 using an IPTG inducible system. Bacteria were then harvested and protein was purified from both soluble lysate and insoluble inclusion bodies using His-tag-binding Ni-NTA resin (Figure 3.7 A and B). Purification fractions were separated by SDS-PAGE for analysis, revealing that the protein (observed at the expected size) was present in a larger proportion in inclusion bodies than in soluble lysate. Purification from inclusion bodies yielded much purer protein than purification from soluble lysate, although a contaminant band at around 20 kDa was persistent in most batches. Recombinant protein was therefore purified from inclusion bodies under denaturing
conditions. Purification in denaturing buffers also ensures that, since all proteins are unfolded, contaminating protease enzymes cannot degrade the protein of interest. This is important when purifying protease enzymes such as DUBs, because many protease inhibitors cannot be included in purification buffers as they may inhibit the activity of the protease of interest itself.

The recombinant protein was then slowly refolded using dialysis into native buffers using SnakeSkin dialysis tubing (Pierce). Once concentrated, it was confirmed that 6His-TsUCH37 was purified under denaturing conditions to a greater level of purity than previous purifications from soluble lysate (data not shown). The presence of the smaller protein however was persistent (Figure 3.7 B, a and b). As expected, the recombinant 6His-TsUCH37 did react with anti-His-tag antibodies, further validating the expected size of the purified protein (Figure 3.7 C). The persistent smaller protein did not react with anti-His antibodies, indicating that this protein is a contaminant that has an affinity for Ni-NTA resin, rather than being a breakdown product of 6His-TsUCH37. For a protein to be active, it is essential that the native confirmation is adopted upon re-folding. It was therefore necessary to validate the ubiquitin hydrolysis activity of the recombinant re-folded 6His-TsUCH37 before further analyses were carried out.
Chapter 3: Results

Figure 3.6

A. The 930 bp RACE-confirmed open reading frame of TsUCH37 was translated into a protein sequence of 309 amino acids with a predicted molecular weight of 35.2 kDa and an isoelectric point (pI) of 4.42.

B. The translated protein sequence of TsUCH37 was aligned with the human UCH37 protein sequence using Geneious alignment software (Drummond 2012 Bioinformatics) showing 45.9% identity.

C. The TsUCH37 coding sequence was cloned into the expression vector pPET28a+ with an N-terminal 6 x Histidine tag (black box). The recombinant protein translation is 343 amino acids long and 38.8 kDa in molecular weight. A, B and C. The UCH peptidase_C12 domain is highlighted by a grey box and catalytic residues are indicated with star symbols.
Figure 3.7 Expression and purification of TsUCH37

A. 6His-TsUCH37 was expressed in E coli. Using Ni-NTA resin (Qiagen), 6His-TsUCH37 was purified from both the (S) soluble lysate (Ni-NTA resin washes and elutions: W and E) and the (I) inclusion bodies (Ni-NTA resin wash buffer at pH 6.3 and elution buffers at pH 5.9 and 4.5). Protein that did not bind to Ni-NTA resin is annotated as flow through (FT). B. His-tag purification under denaturing conditions was optimised for the purification of 6His-TsUCH37 from inclusion bodies. Purified protein was observed at the expected size of 38.8 kDa (a). Another band, representing either a degradation product or a contaminant was observed at approximately 20 kDa (b). C. Anti-His-tag antibodies were reacted with lysate of T. spiralis muscle larvae, purified 6His-TsUCH37 and HEK 293T cell lysate by immuno-blot (IB). The arrow indicates the 6His-TsUCH37-reactive band at the expected molecular weight of 38.8 kDa. All protein marker sizes are indicated in molecular weight (kDa).
3.3. Validation of the ubiquitin hydrolysis activity of recombinant TsUCH37

In order to verify that the recombinant purified 6His-TsUCH37 was an active DUB, a Ub-AMC assay was employed. Ub-AMC is human Ub that contains the C-terminal modification, 7-amido-4-methylcoumarin (AMC). Cleavage of the Ub-AMC substrate by DUB enzymes releases free fluorogenic AMC from the Ub (Figure 3.8 A). The fluorescence produced is cumulative, and can be recorded as a measure of Ub hydrolysis activity. A Ub-AMC hydrolysis assay provides a convenient and sensitive method of testing the activity of a DUB. It also allows for the measurement of real-time, physiologically relevant activity of an enzyme, rather than end-point activity. In other words, the reaction of a DUB with a reagent such as HA-Ub-VME (used originally to identify TsUCH37) demonstrates end-point activity. After a period of incubation time, a DUB is able to form a covalent bond with HA-Ub-VME. During this experiment, both the DUB and the HA-Ub-VME are provided in saturating conditions. These diagnostic results tell us that these two compounds interact, eventually. A Ub-AMC assay however, is carried out using a nM-µM range of concentrations of the Ub-AMC substrate and the DUB enzyme. Measurements are taken immediately after the mixture of the two components, and followed over time, representing a level of activity that will be more relevant to the physiological conditions of a cell.

A series of 6His-TsUCH37 dilutions were reacted with Ub-AMC. Fluorescence was measured at 1-minute intervals for 20 minutes (Figure 3.8 B). Purified recombinant Plasmodium falciparum PfUCHL3 was also reacted with Ub-AMC (at a saturating concentration) as a positive control for activity. The Ub hydrolysis activity of this conserved DUB enzyme has been previously reported [103]. NEM is a cysteine protease inhibitor that irreversibly binds to the active site cysteine of enzymes such as DUBs (see chapter 4 background section for a detailed explanation). For a negative control to ensure that activity was cysteine-based, proteins were pre-incubated with NEM. Ub-AMC hydrolysis activity by 6His-TsUCH37 was observed as quickly as 2 minutes after addition of the Ub-AMC substrate. The activity of 6His-TsUCH37 began to plateau at 20 minutes after Ub-AMC addition, indicating depletion of the substrate. The activities of both TsUCH37 and PfUCHL3 were silenced by pre-incubation with NEM confirming that the activity was dependent on an active site cysteine. Results validated that the recombinant 6His-TsUCH37 was refolded into an active confirmation and could therefore be used as a reliable reagent for the characterisation of the T. spiralis DUB.
Figure 3.8

A. Schematic of Ub-AMC hydrolysis assay. The active-site cysteine residue of a DUB mediates the cleavage of Ub from AMC. Unconjugated AMC fluoresces at 467 nm. The measurement of relative fluorescence units (RFU) over time corresponds to DUB activity. B. 6His-TsUCH37 (2 µM, 1 µM and 500 nM) was reacted with Ub-AMC for 20 minutes. RFU measurements at 467 nm were taken at minute intervals. The hydrolysis of Ub-AMC by each sample was measured in triplicate. Points show the mean RFU and the standard error is indicated as bars. PfUCHL3 was assayed as a positive control (74 nM). Each protein was pre-incubated with the cysteine protease inhibitor NEM as a negative control. Data are presented as the mean of triplicate values from a single experiment and the error bars represent SEM. Consistent data were obtained from similar experiments performed using different concentrations of protein and different protein batches.
3.4. Analysis of the Nedd8 hydrolysis activity of recombinant TsUCH37 and the exploration of dual Ub/Nedd8 hydrolysis by UCH domain DUBs

My supervisor, Katerina Artavanis-Tsakonas, has previously shown that the malaria parasite *Plasmodium falciparum* orthologue of TsUCH37, PfUCH54 (GI: 23496100), is able to hydrolyse both Ub-AMC and the AMC conjugate of the ubiquitin-like modifier Nedd8 (Nedd8-AMC, [102]). Recent data generated by an undergraduate student in our lab, Sara Aguilera, suggests that the primate malaria parasite, *Plasmodium knowlesi* UCH37 (PkUCH37) can also hydrolyse both Ub and Nedd8 (GI: 221056068, data not shown). UCH37 orthologues of higher eukaryotes such as the human UCH37 however, do not demonstrate dual Ub/Nedd8 hydrolysis activity and are specific for Ub. I was therefore interested in testing whether or not this dual function was conserved by the parasitic nematode *T. spiralis*.

Recombinant purified 6His-TsUCH37 was tested for Nedd8-AMC hydrolysis activity in parallel with 3 other recombinant purified proteins: the human Nedd8-specific protease 1 (NEDP1), the human UCH37 and the *P. falciparum* UCH54 (Figure 3.9 A, B, C, D). As a negative control, proteins were pre-incubated with NEM to confirm cysteine specific activity. A reaction with Ub-AMC was also included for NEDP1, HsUCH37 and TsUCH37 to confirm that the protein was not degraded or denatured. Fluorescence measurements were taken every minute for 20 minutes. HsNEDP1 showed very weak Ub-AMC hydrolysis compared to a good level of Nedd8-AMC hydrolysis. Conversely, HsUCH37 exhibited good Ub-AMC hydrolysis and very weak Nedd8 hydrolysis. PfUCH54 showed a high level of Nedd8-AMC hydrolysis as has been previously reported [102]. TsUCH37 showed no Nedd8 hydrolysis activity, but showed good Ub-AMC activity as demonstrated above in Figure 3.9. All activities were silenced by the pre-incubation of the proteins with NEM.
Figure 3.9 Nedd8-AMC hydrolysis by TsUCH37 and orthologues

*In vitro* Nedd8 and Ub-AMC hydrolysis assays of recombinant purified proteins. Fluorescence (representing the release of free AMC) was measured in relative fluorescence units (RFU) over time (minutes) for 20 minutes. **A.** Nedd8-AMC and Ub-AMC hydrolysis by the human Nedd8 specific protease 1 (NEDP1, 500 nM). **B.** Nedd8 and Ub-AMC hydrolysis activity of the human UCH37 (1 µM). **C.** Nedd8 hydrolysis by the *Plasmodium falciparum* UCH37. **D.** Nedd8 and Ub-AMC hydrolysis by 6His-TsUCH37. All enzymes were pre-incubated with the cysteine protease inhibitor NEM as a negative control for cysteine dependent activity. The hydrolysis of Nedd8-AMC or Ub-AMC by each sample was measured in triplicate. Data are presented as the mean of triplicate values from a single experiment and the error bars represent SEM. Consistent data were obtained from similar experiments performed using different concentrations of protein and different protein batches.
Chapter 3: Results

Given the high level of conservation of the ubiquitin pathway in eukaryotes, it is surprising that the *Plasmodium falciparum* and the *Plasmodium knowlesi* UCH37 orthologues have dual Nedd8 and Ub hydrolysis activity but that the mammalian and *T. spiralis* orthologues do not. Orthologues of another conserved UCH domain DUB, UCHL3, have also been shown to have dual activity in *Saccharomyces cerevisiae* (Sc) yeast (GI: 1015802), *Plasmodium falciparum* (Pf, GI: 282403542, *Drosophila melanogaster* (Dm) fruit fly (GI: 17136836), humans (Hs, GI: 5174741) and mice (Mm, GI: 7578956). Table 3.1 lists the orthologous UCH37 DUBs and the orthologous UCHL3 DUBs that have been tested for ubiquitin and Nedd8 hydrolysis activity. Tests for the deNeddylation activity of the *Plasmodium knowlesi, Drosophila melanogaster, human, and Trichinella spiralis* UCH37 proteins and the human and *Drosophila melanogaster* UCHL3 protein, were carried out in our laboratory during the course of the project. All other proteins were previously tested and evidence can be found in the literature (Table 3.1, [102,103,241,242]).

To compare the level of sequence conservation between the proteins listed in Table 3.1, all sequences were aligned using Geneious alignment parameters [237] (Figure 3.10). Patterns of clustering were observed, where some residues were conserved amongst the proteins that have dual deubiquitinating and deNeddylation activity only, and the same residue at this position differed in the proteins that only have deubiquitinating activity. In total, 5 residues were highlighted as being conserved in an activity dependent manner, 2 that were 100% conserved, and 3 that were conserved in hydrophobicity/hydrophilicity. These residues are listed in Table 3.2. For the first residue, an asparagine (N) was found in all the proteins (UCH37 or UCHL3) that have dual activity. Whereas, the corresponding residue of the proteins that only have mono activity was an aspartic acid (D). For the second residue, an aspartic acid (D) was found in all the proteins (UCH37 or UCHL3) that have dual activity. Whereas, the corresponding residue of the proteins that only have mono activity was a glutamic acid (E) (Figure 3.10).
### Table 3.1 Ub and Nedd8 activity profiles of UCH37 and UCHL3 orthologues

<table>
<thead>
<tr>
<th>Accession # (GI)</th>
<th>Hydrolysis activity</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PfUCH54</td>
<td>23496100</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>PkUCH37</td>
<td>221056068</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>TsUCH37</td>
<td>*</td>
<td>X ✓</td>
</tr>
<tr>
<td>MmUCH37</td>
<td>4878011</td>
<td>X ✓</td>
</tr>
<tr>
<td>HsUCH37</td>
<td>7706753</td>
<td>X ✓</td>
</tr>
<tr>
<td>PfUCHL3</td>
<td>282403542</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>ScYUH1</td>
<td>1015802</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>DmUCHL3</td>
<td>17136836</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>HsUCHL3</td>
<td>5174741</td>
<td>✓ ✓</td>
</tr>
</tbody>
</table>

**Table 3.1 Ub and Nedd8 activity profiles of UCH37 and UCHL3 orthologues**

List of UCH37 (UCH-L5) and UCHL3 enzyme orthologues that have been tested for ubiquitin hydrolysis activity and/or Nedd8 hydrolysis activity. Accession numbers are displayed as NCBI gene ID codes (GI). The RACE-confirmed TsUCH37 sequence as yet does not have a GI ID (asterisk). Where applicable, the publication in which the activity was demonstrated is listed. Where no publication is listed, protein activity was tested either by myself or other members of Katerina Artavanis-Tsakonas’ lab (unpublished data).
Table 3.2

<table>
<thead>
<tr>
<th>Accession # (GI)</th>
<th>Nedd8</th>
<th>Ub</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>PfUCH37</td>
<td>23496100</td>
<td>✓</td>
<td>✓</td>
<td>N</td>
<td>D</td>
<td>Y</td>
<td>K</td>
</tr>
<tr>
<td>PkUCH37</td>
<td>221056068</td>
<td>✓</td>
<td>✓</td>
<td>N</td>
<td>D</td>
<td>Y</td>
<td>K</td>
</tr>
<tr>
<td>TsUCH37</td>
<td>*</td>
<td>✓</td>
<td>D(12)</td>
<td>E(32)</td>
<td>H</td>
<td>I(148)</td>
<td>P</td>
</tr>
<tr>
<td>MmUCH37</td>
<td>4878011</td>
<td>X</td>
<td>✓</td>
<td>D</td>
<td>E</td>
<td>H</td>
<td>F</td>
</tr>
<tr>
<td>HsUCH37</td>
<td>7706753</td>
<td>X</td>
<td>✓</td>
<td>D</td>
<td>E</td>
<td>H</td>
<td>F</td>
</tr>
<tr>
<td>PfUCHL3</td>
<td>282403542</td>
<td>✓</td>
<td>✓</td>
<td>N</td>
<td>D</td>
<td>K</td>
<td>Q</td>
</tr>
<tr>
<td>ScYUH1</td>
<td>1015802</td>
<td>✓</td>
<td>✓</td>
<td>N</td>
<td>D</td>
<td>K</td>
<td>Q</td>
</tr>
<tr>
<td>DmUCHL3</td>
<td>17136836</td>
<td>✓</td>
<td>✓</td>
<td>N</td>
<td>D</td>
<td>K</td>
<td>Q</td>
</tr>
<tr>
<td>MmUCHL3</td>
<td>7578956</td>
<td>✓</td>
<td>✓</td>
<td>N</td>
<td>D</td>
<td>C</td>
<td>Q</td>
</tr>
<tr>
<td>HsUCHL3</td>
<td>5174741</td>
<td>✓</td>
<td>✓</td>
<td>N</td>
<td>D</td>
<td>C</td>
<td>Q</td>
</tr>
</tbody>
</table>

Table 3.2 UCH37 and UCHL3 residues that cluster with Ub/Nedd8 activity profile

List of UCH37 (UCH-L5) and UCHL3 enzyme orthologues that have been tested for ubiquitin hydrolysis activity and Nedd8 hydrolysis activity. Protein sequences were aligned and their residues that cluster based on dual deubiquitinating and deNeddylating activity were identified (numbered 1-5). Residues 1 and 2 are conserved in identity. Residues 3-5 are conserved in hydrophobicity/hydrophilicity. Residues 2 and 4, highlighted in red, are also important for ScYUH1 interaction with Ub (Table 3.3). Residue 1, highlighted in yellow, is also important for PfUCHL3 interaction with Nedd8 (Table 1.D). This residue was mutated in 6His-TsUCH37 (D>N).
Figure 3.10

Residues that cluster based on activity (listed in Table 3.2)
Residues that are important for YUH1 interaction with Ub (listed in Table 3.3)

Figure 3.10 Alignment of UCH37 and UCHL3 orthologues

Orthologues of UCH37 and UCHL3 DUBs were aligned using MUSCLE parameters [209]. The UCH37 orthologous proteins were aligned: P. falciparum GI: 23496100, P. knowlesi GI: 221056068, T. spiralis RACE-confirmed (no accession number), M. musculus GI: 4878011, H. sapiens GI: 7706753. The UCHL3 orthologues were aligned: P. falciparum GI: 282403542, S. cerevisiae GI: 282403542, D. melanogaster GI: 17136836, M. musculus GI: 7578956, H. sapiens GI: 5174741. Residues that cluster in alignment based on dual deubiquitinating and deNeddylating activity are numbered 1-5 and boxed in red. Residues predicted by Artavanis-Tsakonas et al. [104] to be involved in the direct binding of PfUCHL3 to Nedd8 (Table 3.4) are boxed in blue. The bar graph above the sequence indicates the level of identity within the alignment.

1. PIUCH54
2. PIUCH56
3. PIUCH57
4. MmUCH37
5. HsUCH37
6. PIUCHL3
7. ScYUH1
8. MmUCH37
9. MmUCH37
10. HsUCH37

MUSCLE alignment, Genieus
Further analyses were carried out to try and understand the structural basis of the direct interaction between 1. a DUB and Ub and 2. a DUB and Nedd8.

1. The crystal structure of human UCH37 in complex with Ub has not been solved. Instead, the crystal structure of the *S. cerevisiae* YUH1 in complex with Ub (reported by Johnston et al.) was used to identify residues important for direct DUB to Ub contact [243]. These residues were then identified in each orthologue of UCH37 and UCHL3 listed in Table 3.1, using the multi-orthologue alignment in Figure 3.10. Table 3.3 lists the residue number of YUH1, and each corresponding residue of each orthologue predicted to be involved in the binding of Ub. 11 residues were identified, 2 of which were also identified to cluster based on dual deubiquitinating and deNeddylating activity (Table 3.2). These were residues D35 and Q153 of YUH1.

2. Artavanis-Tsakonas et al. used the crystal structure of PfUCHL3 in complex with Ub-VME to model Nedd8 in the position of the Ub ([104], Figure 3.11). Residues that appeared important for direct contact between PfUCHL3 and Nedd8 were identified. 4 residues of PFUCHL3 were predicted to be involved in direct DUB to Nedd8 contact: glutamic acid (E) 11, asparagine (N) 13, glutamic acid (E) 153 and aspartic acid (D) 157 (Table 3.3). These residues were then aligned with the other orthologues listed in Table 3.1 for which deubiquitinating and deNeddylating activity has been tested.
Chapter 3: Results

Table 3.3

<table>
<thead>
<tr>
<th>Accession # (GI)</th>
<th>Nedd8</th>
<th>Ub</th>
<th>E12</th>
<th>D35</th>
<th>L58</th>
<th>Q84</th>
<th>C90</th>
<th>F149</th>
<th>T151</th>
<th>Q153</th>
<th>T183</th>
<th>L185</th>
<th>N225</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIUCH54</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIUCH37</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TsUCH37</td>
<td></td>
<td>X</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MnuUCH37</td>
<td>4876011</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HsUCH37</td>
<td>7706753</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIUCHL3</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ScYUH1</td>
<td>1015802</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DmUCHL3</td>
<td>17136836</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MnuUCHL3</td>
<td>7578056</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HsUCHL3</td>
<td>5174741</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

List of UCH37 (UCH-L5) and UCHL3 enzyme orthologues that have been tested for ubiquitin hydrolysis activity and Nedd8 hydrolysis activity. Protein sequences were aligned and the residues that are important for the interaction between ScYUH1 and Ub (as identified by Johnston et al. EMBO 1999 [243]) were listed. Residue numbers represent the position in the ScYUH1 protein. The corresponding residue in the orthologous sequences (when aligned) was listed. Residues highlighted in red are also listed in Table 3.2 because they cluster based on dual deNeddylating and deubiquitinating activity.

Structural data from Johnston et al. EMBO 1999

Table 3.3 UCH37 and UCHL3 residues predicted to be important for Ub contact
Table 3.4 UCH37 and UCHL3 residues predicted to be important for Nedd8 contact

List of UCH37 (UCH-L5) and UCHL3 enzyme orthologues that have been tested for ubiquitin hydrolysis activity and Nedd8 hydrolysis activity. The residues predicted to be important for the interaction between PfUCHL3 and Nedd8 are numbered. Protein sequences were aligned and residues of orthologous sequences were listed as they correspond to the PfUCHL3 residue. The residue highlighted in yellow was also identified in Table 3.2 as clustering based on dual deNeddylating and deubiquitinating activity.

<table>
<thead>
<tr>
<th>Accession # (GI)</th>
<th>Nedd8</th>
<th>Ub</th>
<th>E11</th>
<th>N13</th>
<th>E153</th>
<th>D157</th>
</tr>
</thead>
<tbody>
<tr>
<td>PfUCHL3</td>
<td>23496100</td>
<td>✔</td>
<td>✔</td>
<td>E</td>
<td>N</td>
<td>D</td>
</tr>
<tr>
<td>PkUCH37</td>
<td>221056068</td>
<td>✔</td>
<td>✔</td>
<td>E</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>TsUCH37</td>
<td>*</td>
<td>X</td>
<td>✔</td>
<td>E</td>
<td>D(12)</td>
<td>D</td>
</tr>
<tr>
<td>MmUCH37</td>
<td>4878011</td>
<td>X</td>
<td>✔</td>
<td>E</td>
<td>D</td>
<td>T</td>
</tr>
<tr>
<td>HsUCH37</td>
<td>7706753</td>
<td>X</td>
<td>✔</td>
<td>E</td>
<td>D</td>
<td>T</td>
</tr>
<tr>
<td>PfUCHL3</td>
<td>282403542</td>
<td>✔</td>
<td>✔</td>
<td>E</td>
<td>N</td>
<td>E</td>
</tr>
<tr>
<td>ScYUH1</td>
<td>1015802</td>
<td>✔</td>
<td>✔</td>
<td>E</td>
<td>N</td>
<td>E</td>
</tr>
<tr>
<td>DmUCHL3</td>
<td>17136836</td>
<td>✔</td>
<td>✔</td>
<td>E</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>MmUCHL3</td>
<td>7578958</td>
<td>✔</td>
<td>✔</td>
<td>E</td>
<td>N</td>
<td>E</td>
</tr>
<tr>
<td>HsUCHL3</td>
<td>5174741</td>
<td>✔</td>
<td>✔</td>
<td>E</td>
<td>N</td>
<td>E</td>
</tr>
</tbody>
</table>

Artavanis-Tsakonas et al. JBC 2010
Figure 3.11
Model of Nedd8 in the position of Ub in complex with PfUCHL3

![Model of Nedd8 interaction with PfUCHL3](image)

Image modified from Artavanis-Tsakonas et al. JBC 2010

**Figure 3.11 Model of Nedd8 interaction with PfUCHL3**

A structural model of PfUCHL3 bound to PfNedd8 was built using the Ub-VME-bound crystal structure (Artavanis-Tsakonas et al. JBC 2010 [104]). The structure of PfNedd8-VME (magenta) was superimposed onto the Ub-VME (green)-bound PfUCHL3 active site (grey). Four residues of PfUCHL3 were predicted to be involved in direct DUB to Nedd8 contact (highlighted in yellow and boxed in red): glutamic acid (E) 11, asparagine (N) 13, glutamic acid (E) 153 and aspartic acid (D) 157.
It was observed that one of these residues was also shown to be important for the binding of YUH1 to Ub as shown in Table 3.3 (PfUCHL3 E11 and YUH1 E12). PfUCHL3 residue N13, was also identified as being conserved only in enzymes with dual deubiquitinating and deNeddylating activity (listed as residue number 2 Table 3.1, Figure 3.10). Here, the conserved residue at the corresponding position in the proteins that only have deubiquitinating activity is aspartic acid (D). Not only does this residue cluster based on dual versus mono activity, it is also predicted to be involved in the direct contact between a PfUCHL3 and Nedd8 but not in the direct contact between YUH1 and Ub. The question was therefore raised: would the mutation of this residue in TsUCH37 from D to N, cause a gain in deNeddylating activity?

Site-directed mutagenesis (substitution) of 6His-TsUCH37 D12>N was carried out to produce a mutant coding sequence (Figure 3.12 A). The mutant protein was expressed and purified as described earlier for wild-type 6His-TsUCH37 (section 3.2). Ub-AMC and Nedd8-AMC assays were then used to assess whether or not the TsUCH37 D12N mutant had gained dual activity for Ub and Nedd8 (AMC assays were described earlier, section 3.2). TsUCH37 D12N showed Ub-AMC hydrolysis comparable to TsUCH37 wild-type, but was not able to hydrolyse Nedd8-AMC (Figure 1.12 B and C). It is therefore clear that the residue TsUCH37 D12 is not determinative of dual Ub and Nedd8 activity.

Other members of our group have continued this study. The other residues of the UCH37 orthologues that appear to cluster based on activity (listed in Table 3.1) are undergoing mutagenesis. These mutants will be screened to identify the structural and evolutionary basis of dual Ub and Nedd8 activity in UCH37 and UCHL3 DUBs.
Chapter 3: Results

Figure 3.12

A 6His-TsUCH37 mutagenesis of D12 to N

B Ub-AMC hydrolysis by TsUCH37 WT and D12N

C Nedd8-AMC hydrolysis by TsUCH37 WT and D12N

Figure 3.12 Mutagenesis and Ub/Nedd8 activity profiling of TsUCH37

A. Recombinant 6His-TsUCH37 mutant was cloned, where residue D12 was mutated to N. Recombinant mutant protein was expressed and purified. B. The Ub-AMC and C. Nedd8-AMC hydrolysis activity of the mutant 6His-TsUCH37 D12N was tested. Reactions were monitored for 20 minutes. Relative fluorescence unit (RFU) measurements at 467 nm were taken at minute intervals. The hydrolysis of Ub-AMC and Nedd8-AMC by each sample was measured in triplicate. Data are presented as the mean of triplicate values from a single experiment and the error bars represent SEM. Consistent data were obtained from similar experiments performed using different concentrations of protein and different protein batches. PfUCHL3 and the wild-type 6His-TsUCH37 were assayed as positive controls for activity. Each protein was pre-incubated with the cysteine protease inhibitor NEM as a negative control.
Although a lower eukaryote and a parasite, based on these analyses the _T. spiralis_ TsUCH37 has more in common with the mammalian orthologues than the _Plasmodium_ orthologues. The human and mouse UCH37 are proteasome associated DUBs that play an important role in proteasomal degradation of substrate proteins [33,233]. It was therefore hypothesised that the association of _T. spiralis_ UCH37 with the proteasome may have been conserved throughout evolution to higher eukaryotes. TsUCH37-associated proteins were therefore analysed, with a focus on _T. spiralis_ proteasome subunits.

### 3.5. Analysis of proteins that associate with TsUCH37

The structure of the 26S proteasome is displayed in Figure 3.14 C. The 26S proteasome is made up of the 19S cap structure that binds the substrate protein, and the 20S catalytic core ATP-dependent protease that degrades the substrate protein as it passes through the hollow core [21]. The 19S regulatory particle is made up of a lid and a base. The lid structure contains the regulatory subunits Rpn3, 5-9, 11,12 and 15. In-between this and the catalytic core is the 19S base, which is composed of the regulatory subunits Rpn10, 1, 2 and Rpn13, otherwise known as ADMR1 (usually called Rpn13 in yeast). The 20S core is made up of proteolytic alpha and beta subunits that mediate substrate degradation.

In human cells, HsUCH37 is often found in direct association with ADRM1, a protein that is referred to as the ubiquitin receptor. ADRM1 associates with the proteasome 19S base subunit Rpn2 [32]. During the MRes project, HA-Ub-VME was reacted with _T. spiralis_ muscle larvae lysate, and probe-protein complexes were immuno-precipitated using anti-HA antibodies (MRes report 2010, [1]). When the products of immuno-precipitation were analysed by LC/MS/MS, a number of putative _T. spiralis_ DUBs were identified. These included TsUCH37. A number of other putative _T. spiralis_ proteins were also identified by LC/MS/MS. Although these proteins had co-precipitated with HA-Ub-VME and anti-HA resin, they were not themselves DUBs. These proteins may either have been experimental contaminants, other proteins able to specifically associate with the ubiquitin component of HA-Ub-VME, or proteins able to specifically associate with DUBs that were in complex with HA-Ub-VME. One of these proteins was a putative _T. spiralis_ ADRM1 orthologue.

It was hypothesised that if TsUCH37 were found also to associate with the _T. spiralis_ ADRM1 it would also be able to associate with the _T. spiralis_ proteasome. The aim was to investigate this by co-precipitating the _T. spiralis_ ADRM1 from lysate of _T. spiralis_
muscle larvae, using commercially available anti-human ADRM1 antibodies. The products of co-precipitation would then be analysed for the presence of TsUCH37 using anti-human UCH37 antibodies in an immuno-blot. The *T. spiralis* and human ADRM1 share 36.6% of the same residues, and the *T. spiralis* and human UCH37 share 45.9% of the same residues making it plausible that the antibodies would cross-react with the *T. spiralis* orthologues. The experiment could also be carried out in the reverse order. In other words, immuno-precipitation could be carried out using anti-UCH37 antibodies followed by immuno-blot with anti-ADRM1 antibodies (Figure 3.13 A). The antibodies had to therefore be tested first for cross-reactivity with the *T. spiralis* orthologues (Figure 3.13 B and C).

Commercial anti-human UCH37 and ADRM1 antibodies were tested for reactivity with lysate of *T. spiralis* muscle larvae. As a control for reactivity, human embryonic kidney HEK 293T cell lysate was also tested. HEK 293T cells are a cell line often used in the laboratory because they are robust and easy to culture. They were used throughout the project to represent human cell protein expression.

Although a protein at an expected size (human ADRM1 weighs 42.2 kDa, and human UCH37 weighs 35.9 kDa) was observed in the HEK lysate, the antibodies did not cross react with recombinant purified TsUCH37 or with *T. spiralis* lysate (Figure 3.13 B and C). These antibodies could therefore not be used for further analyses of *T. spiralis* ADRM1. Instead, recombinant 6His-TsUCH37 was used to analyse the potential association of the protein with ADRM1. Because *T. spiralis* proteins did not cross-react with the anti-human ADRM1 antibodies used, lysate from human HEK 293T cells was used to investigate whether or not recombinant *T. spiralis* TsUCH37 could associate with the human ADRM1. This interaction could then be monitored using the anti-human ADRM1 antibodies.
Figure 3.13 Analysis of TsUCH37 co-precipitation with ADMR1

A. Schematic of the experimental methods. B. Immuno-blots (IB) analysis of *T. spiralis* muscle larvae (ML) lysate, 6His-TsUCH37 and HEK 293T cell lysate using anti-human UCH37 antibodies. (a) Approximate size of human and *T. spiralis* UCH37. C. Immuno-blot analysis of *T. spiralis* muscle larvae lysate and HEK 293T cell lysate using anti-human ADRM1 antibodies. (b) Approximate size of human and *T. spiralis* ADRM1. D. DB71 stain of Ni-NTA 6His-TsUCH37 co-precipitation with HEK 293T cell lysate. E. Ni-NTA 6His-TsUCH37 co-precipitation with HEK cell lysate followed by immunoblot using anti-human ADRM1 antibodies. (a) Approximate size of human ADRM1. (b) Approximate size of human 6His-TsUCH37. All protein marker sizes are indicated in molecular weight (kDa).
Recombinant purified 6His-TsUCH37 was bound to Ni-NTA resin. This was then incubated with HEK lysate under conditions that would preserve protein-protein interactions. As a control, native Ni-NTA resin, with no 6-His-TsUCH37, was incubated with HEK lysate. The resin was then washed, and associated proteins were eluted. To assess whether or not any E. coli contaminants remained after the original purification process of 6His-TsUCH37 and to observe any non-specific interactions between these contaminants and the anti-ADRM1 another control was included. This was purified recombinant 6His-TsUCH37 alone, bound to Ni-NTA (with no cell lysate added). All the eluates were then separated by SDS-PAGE, transferred to PVDF membrane and visualised using a dye called direct blue 71 (DB71) that is similar to coomassie but binds to proteins on PVDF membranes (Figure 3.1 D). Proteins were then analysed by immuno-blots using anti-(Hs)-ADRM1 antibodies (Figure 3.1 E). No reactivity was observed with 6His-TsUCH37 alone. After co-precipitation of 6His-TsUCH37 (bound to Ni-NTA resin) with HEK lysate, a band was observed at the expected size for human ADRM1 (42.2 kDa). However, the same band was observed after the co-precipitation of HEK lysate with naked Ni-NTA resin. It appeared possible that this band was enriched (but not significantly) in the 6His-TsUCH37 co-precipitation sample. The binding of human ADRM1 to Ni-NTA, both in the absence and the presence of 6His-TsUCH37, suggests that human ADRM1 can non-specifically interact with Ni-NTA under conditions designed to preserve protein-protein interactions. I propose that human ADRM1 is able to non-specifically bind to either the nickel or the agarose of the Ni-NTA resin. It was therefore not possible to use these reagents to verify direct interactions between ADRM1 and either endogenous TsUCH37, or recombinant 6His-TsUCH37.

A possible follow-up approach would have been to produce T. spiralis-specific antibodies to either the TsUCH37, TsADRM1 or indeed to other T. spiralis proteasome components. However, a mass spectrometry-based approach was adopted for 3 reasons:

1. The peptide coverage of TsADRM1 identified during LC/MS/MS analysis of HA-Ub-VME bound T. spiralis proteins was 11%. This did not provide much sequence upon which potential immunogen sequences could be mapped.

2. Since the annotation was only predicted, the TsADRM1 coding sequence would first have to be confirmed using RACE-PCR.

3. A collaboration with Eric Spooner at the mass spectrometry facility at The Whitehead Institute was already underway.
Mass spectrometry was therefore used to analyse the *T. spiralis* proteins that co-precipitate with recombinant 6His-TsUCH37, and search for possible *T. spiralis* ADRM1 or proteasome components within the results.

A large quantity of recombinant purified 6His-TsUCH37 was bound to Ni-NTA resin, which was then incubated with lysate of *T. spiralis* muscle larvae under conditions that would preserve protein-protein interactions. Three controls were included: 1. Naked Ni-NTA resin with no recombinant protein bound was incubated with *T. spiralis* lysate to control for non-specific interactions. 2. purified recombinant 6His-TsUCH37 alone was bound to Ni-NTA to control for contaminants of the purification process. 3. Crude *T. spiralis* lysate with no Ni-NTA or 6His-TsUCH37 to observe the SDS-PAGE profile of all *T. spiralis* proteins. Ni-NTA resin was washed, co-precipitated proteins were separated by SDS-PAGE and visualised by staining with colloidal coomassie (Figure 3.14 A). Proteins were manually excised from the gel that appeared only in the Ni-NTA-6His-TsUCH37 plus *T. spiralis* lysate sample, and not in the control samples. Excised proteins were analysed by LC/MS/MS. Using SEQUEST, the identified peptides were then matched against the annotated draft genome of *T. spiralis* (Genome Institute at Washington University). All mass spectrometry and SEQUEST screening was carried out by Eric Spooner at the Whitehead Institute in Boston.

Matches were made to 12 different putative *T. spiralis* proteasome components (Figure 3.14 A and B and appendix 1, and all co-precipitated proteins are listed in the supplementary appendix 1) of the 26S proteasome. These 12 proteasome proteins were amongst the highest scoring matches (in terms of unique peptide count) of all proteins that co-precipitated, suggesting an increased likelihood that they were specific to TsUCH37 association. For example, the 4th highest scoring match of all co-precipitated proteins was the putative *T. spiralis* 19S base subunit protein, Rpn1. Rpn1 matched 28 exclusive unique peptides found in band 2 of the gel (only), giving a coverage of 39% of a 102 kDa (911 amino acids) protein (data not shown). Including this result, the most abundant *T. spiralis* putative proteasome proteins were those that are found in the 19S base of the proteasome, such as Rpn1, Rpn2 and Rpt1. This is the part of the mammalian proteasome that interacts with the mammalian UCH37. Other matches were made to proteins found in the 19S lid: Rpn3 and Rpn6, and proteins of the 20S catalytic core such as the alpha and beta ATPase subunits (Figure 3.14 B). Results also included a PSME3 orthologue otherwise known as the PA28 subunit of the immuno-proteasome [244]. Three unique peptides matching the *T. spiralis* putative ADRM1 orthologue were also identified (GI:316977948). The annotated TsADRM1 sequence that matched these peptides was the same as that previously
identified by HA-Ub-VME immuno-precipitation of *T. spiralis* muscle larvae lysate during the MRes project [1]. In this analysis, the peptides identified by co-precipitation covered 10% of the putative TsADRM1 sequence of a protein predicted to weigh 46 kDa (data not shown). When compiled with peptide data from the MRes project HA-Ub-VME immuno-precipitation experiment, peptides identified by mass spectrometry covered 14.4% of the putative *T. spiralis* ADRM1 sequence. Proteins were often found to migrate during SDS-PAGE in a similar manner to human proteasome proteins as previously reported by Lee et al. [245] (Figure 3.14 D). For example, 2 of the largest subunit proteins, Rpn1 and Rpn2 were found to have the least electrophoretic mobility as they are amongst the largest of the proteasome subunits.

Data suggests that *T. spiralis* UCH37 can associate with the putative *T. spiralis* ADRM1 and with multiple proteasome subunits, but mostly those from the 19S base. This association with the proteasome may be via the interaction with ADRM1, as is observed in mammalian systems. It is therefore likely that the *T. spiralis* UCH37 also functions as a proteasome interaction partner. If so this function has been conserved throughout evolution, from nematodes to mammals. In mammals, proteasome associated UCH37 is essential for survival. The knockout of mouse UCH37 is lethal at the embryonic stage [36]. It was therefore hypothesised that TsUCH37 may also be essential for the survival of the parasite, and if so, may be a potential drug target to be considered in the development of novel anthelmintics.
Figure 3.14 Analysis of TsUCH37 co-precipitation with the T. spiralis proteasome

A. Co-precipitation of 6His-TsUCH37 with lysate of T. spiralis muscle larvae. Bands that appeared only in the test sample (Ni-NTA, 6His-TsUCH37 and T. spiralis muscle larvae lysate) and not in the control lanes, were manually excised (numbered 1-19). These proteins were analysed by LC/MS/MS by Eric Spooner (Whitehead Institute). Band 9 is marked in red and was the location of peptides that matched the putative T. spiralis ADRM1. Protein marker sizes are indicated in molecular weight (kDa).

B. Putative T. spiralis proteasome components identified by LC/MS/MS

<table>
<thead>
<tr>
<th>Band</th>
<th>Putative T. sp proteasome component</th>
<th>Number of unique peptides</th>
<th>Predicted molecular weight (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rpn2</td>
<td>18</td>
<td>156</td>
</tr>
<tr>
<td>2</td>
<td>Rpn1</td>
<td>28</td>
<td>102</td>
</tr>
<tr>
<td>4</td>
<td>Rpn6, Rpn10, PSME3</td>
<td>4, 4, 13</td>
<td>50, 45, 29</td>
</tr>
<tr>
<td>5</td>
<td>PSME3</td>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>PSME3</td>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td>7</td>
<td>Rpt1, Rpn3</td>
<td>7, 6</td>
<td>18, 59</td>
</tr>
<tr>
<td>8</td>
<td>Rpt1, alpha1, alpha2, beta1, PSME3</td>
<td>18, 4, 2, 3</td>
<td>18, 218, 13, 34, 28</td>
</tr>
<tr>
<td>9</td>
<td>ADRM1, Rpn10</td>
<td>3, 6</td>
<td>46, 6</td>
</tr>
<tr>
<td>10</td>
<td>alpha2, PSME</td>
<td>4, 13</td>
<td>26, 28</td>
</tr>
<tr>
<td>11</td>
<td>Rpt1</td>
<td>4</td>
<td>39</td>
</tr>
<tr>
<td>12</td>
<td>Rpt1, alpha6, PSME3</td>
<td>4, 3, 13</td>
<td>36, 13, 28</td>
</tr>
<tr>
<td>13</td>
<td>alpha6</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>PSME3</td>
<td>13</td>
<td>28</td>
</tr>
</tbody>
</table>

C. The structure of the proteasome. The ADRM1 subunit (Rpn13) is red.

D. SDS-PAGE of the human 26S proteasome

Image modified from Weissman et al, Nat Rev Mol Cell Biol 2011
Image modified from Lee et al, J BC 2012

Chapter 3: Results
3.6. Investigation of the potential of TsUCH37 as a drug target

*T. spiralis* cannot be genetically modified. Therefore genetic analyses, such as specific gene knockout or gene or transcript silencing could not be carried out. Instead, the potential of TsUCH37 as a drug target was investigated using available inhibitor compounds that were previously shown to be specific to UCH DUBs. Two small compounds were found to be available. The small molecule inhibitor compound WP1130 has been shown to inhibit UCH37 in human cells [246]. LDN-57444 has been shown to inhibit UCH-L1 and to a lesser extent, UCH-L3 in human cells [247,248].

Considering the level of functional conservation of UCH37, and the level of structural conservation between UCH domain DUBs, it was hypothesised that WP1130 and LDN-57444 may also inhibit TsUCH37. This was first tested in vitro using Ub-AMC hydrolysis assay.

WP1130 was kindly supplied by William Bornmann of the University of Texas. WP1130 (degrasyn) is a small tyrphostin compound that has been shown to inhibit the activity of USP9x, USP5, USP14, UCHL1 and UCH37, but not UCHL3 [246], Figure 3.15 A). The effect of WP1130 on the in vitro activity of recombinant human TsUCH37 was tested using a Ub-AMC assay. The effect on human UCH37 and UCHL1 were also tested to confirm in vitro inhibition of the mammalian DUBs using this method. Purified recombinant proteins were incubated with WP1130 at the indicated concentrations, before adding Ub-AMC (Figure 3.15 B, C and D). As a positive control for inhibition, proteins were pre-incubated with NEM only and, as a negative control for inhibition, with the drug solvent (DMSO) only. Fluorescence measurements were taken every minute for 20 minutes. In contrast to previously reported data, no inhibition by WP1130 of Ub hydrolysis activity of either protein was observed by Ub-AMC assay. Since no inhibition of TsUCH37 was observed, the second compound LDN-57444 was used for further experiments.
Figure 3.15  

A. The structure of the small molecule UCH enzyme inhibitor WP1130. The Ub-AMC hydrolysis activities of B. human UCH37, C. human UCHL1 and D. 6His-TsUCH37 were tested in the presence of WP1130 at the indicated concentrations. Relative fluorescence unit (RFU) measurements at 467 nm were taken at minute intervals for 20 minutes. Each protein was pre-incubated with the cysteine protease inhibitor NEM as a positive control for inhibition. Proteins were assayed with DMSO only as a negative control for inhibition. The hydrolysis of Ub-AMC by each sample was measured in triplicate. Data are presented as the mean of triplicate values from a single experiment and the error bars represent SEM. Consistent data were obtained from similar experiments performed using different concentrations of protein and different protein batches.
Chapter 3: Results

LDN-57444 (compound 30) is a cell permeable isatin o-acyl oxime (Figure 3.16 A). LDN-57444 exhibits active site-directed inhibition of HsUCHL3 and, with greater potency, HsUCHL1 [247,248] Treatment of mammalian neuronal cells with LDN-57444 causes a build up of highly ubiquitinated proteins [249].

Recombinant TsUCH37, and recominant HsUCH37 as a control, were incubated with either 1 mM LDN-57444 (solubilised in DMSO) or with DMSO alone. Proteins were then mixed with Ub-AMC, before a fluorescence measurement was taken every minute for 20 minutes. Pre-incubation of recombinant TsUCH37 with NEM was assayed as a positive control for inhibition and to confirm cysteine-specific activity. Samples incubated with DMSO alone, showed Ub hydrolysis activity from TsUCH37 and UCHL1 (Figure 3.16 B). Almost complete inhibition by 1 mM of LDN-57444 of UCHL1 was observed, whereas partial inhibition of TsUCH37 was observed. A titration of LDN-57444 (50, 100 and 500 µM) showed specific, concentration-dependent inhibition of recombinant TsUCH37 activity compared to the DMSO control (Figure 3.16 C).

Although LDN-5744 inhibits TsUCH37 in vitro, it is also able to inhibit mammalian UCH-L1 and UCH-L3 [247]. It is therefore not a specific inhibitor of UCH37, rather a specific inhibitor of UCH domain DUB enzymes. Putative orthologues of the T. spiralis L3 enzyme (but not UCHL1) can be identified by BLAST analysis of the T. spiralis annotated protein database. In addition, a putative T. spiralis UCH-L3 was identified by mass spectrometry analysis of proteins immuno-precipitated by the HA-Ub-VME probe. T. spiralis muscle larvae therefore express at least 2 UCH domain DUBs. For this reason, LDN-57444 could not be used to specifically target TsUCH37 in vivo. It was therefore used to investigate whether or not the UCH DUB enzyme family of T. spiralis might be essential for the survival of the muscle stage parasite.
Figure 3.16  

A. The effect of LDN-57444 on TsUCH37 Ub hydrolysis activity

B. Ub-AMC hydrolysis by TsUCH37 and HsUCHL1

C. Ub-AMC hydrolysis by TsUCH37


Figure 3.16 The effect of LDN-57444 on TsUCH37 Ub hydrolysis activity

A. The structure of the small molecule UCH enzyme inhibitor LDN-57444. B. The Ub-AMC hydrolysis activities of human UCHL1 and TsUCH37 were tested in the presence of 1 mM of LDN-57444. As a negative control for inhibition, proteins were incubated with DMSO only. As a positive control for inhibition, 6His-TsUCH37 was incubated with the cysteine protease inhibitor NEM. C. 6His-TsUCH37 was tested in the presence of a titration of LDN-57444 (concentrations are indicated). Relative fluorescence (RFU) measurements at 467 nm were taken at minute intervals for 20 minutes. 6His-TsUCH37 was pre-incubated with the cysteine protease inhibitor NEM as a positive control for inhibition. 6His-TsUCH37 was assayed with DMSO only as a negative control for inhibition. The hydrolysis of Ub-AMC by each sample was measured in triplicate. Data are presented as the mean of triplicate values from a single experiment and the error bars represent SEM. Consistent data were obtained from similar experiments performed using different concentrations of protein and different protein batches.
Chapter 3: Results

_T. spiralis_ muscle larvae were isolated, cultured and immediately incubated with LDN-57444 over a total period of 96 hours. Parasite viability was then measured using the MTT assay. The MTT viability test is a quantitative colorimetric assay based on the tetrazolium salt, 3-[4,5-diethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide [250,251]. MTT is yellow in colour. It is reduced by dehydrogenase enzymes of metabolically active cells to insoluble purple crystals of formazan. Formation of formazan does not occur in dead cells and is directly proportional to the number of live cells. Formazan crystals can be solubilised by adding organic solvents to the parasite culture, and the absorbance of the purple supernatant represents cell viability. Over a period of 96 hours, muscle larvae were incubated with either LDN-57444 solubilised in DMSO or DMSO alone, providing a negative control for inhibition (100% viability). An MTT assay was then carried out for each sample after 24 hours, 48 hours and 96 hours of incubation. An MTT assay of heat-killed larvae that cannot produce formazan (larvae heated to 70°C for 10 minutes) was taken as a positive control for inhibition for each time-point (0% viability). After 24 hours of incubation with LDN-57444, a decrease in motility and a change in morphology were observed, with more larvae becoming uncoiled and stationary (an early sign of _T. spiralis_ muscle larvae death) than the DMSO control sample. When quantified by the MTT assay, a significant reduction in the viability of the parasites incubated with 100 µM LDN-57444 was observed, compared to the DMSO control sample (Figure 3.17). After 48 and 96 hours of incubation, the viability of larvae was significantly reduced by 50 µM and 100 µM of LDN-57444. By 96 hours, the viability of muscle larvae incubated with 100 µM LDN-57444 was reduced by 75%.
Figure 3.17

The viability of *T. spiralis* muscle larvae was measured using the MTT assay. Larvae were incubated with DMSO alone as a negative control for inhibition of viability (100% viability) then with 50 or 100 µM of LDN-57444. MTT assays were taken at 24 hours (A), 48 hours (B) and 96 hours (C) incubation. As a positive control for inhibition of viability, larvae were killed by heat treatment (HT) at each time-point. All MTT assays were carried out in triplicate. The mean percentage viability of the DMSO control is plotted, with standard error of the mean as error bars. Data was analysed using a two-way ANOVA test. Statistically significant data (P<0.005) is indicated with an asterisk.

Results show that TsUCH37 activity can be targeted with small molecule inhibitors and that UCH DUBs, including TsUCH37, may be essential for the survival of *T. spiralis* muscle larvae. This set precedent for the further investigation of TsUCH37 as a specific drug target. A useful drug for therapeutic use during parasitic infection must target the parasite enzyme but not target the mammalian orthologue. In order to separately target the parasite orthologue of a conserved protein, it is useful to have the crystal structures of both proteins to identify structural differences that could be exploited in drug identification or design.

In order to find, design or develop effective enzyme inhibitors, crystal structures must first be resolved. The crystal structures of the full length and the catalytic domain human UCH37 have been previously reported [25] [252], [253]. In order to compare the structure of the human UCH37 with TsUCH37, the clone that I had made and used for protein synthesis was given to Chittaranjan Das, the supervisor of a crystallography laboratory at Purdue University (US). Using protocols developed during my project, Marie Morrow, a member of the crystallography group, was able to express and purify the full length and the catalytic domain residues (1-226) of TsUCH37. This led to the successful determination of the crystal structure of both the full length TsUCH37, and the catalytic domain of TsUCH37, each in complex with Ub-VME (appendix 7). This angle of the project has continued in the Das laboratory, as the group endeavors to investigate the structural basis of the specific functions of deubiquitinating enzymes.
Summary

3.1. The correct open reading frame of the *T. spiralis* DUB TsUCH37 was hypothesised using bioinformatics analyses and confirmed experimentally using RACE techniques.

3.2. The coding sequence of TsUCH37 was cloned and recombinant His-tagged TsUCH37 was expressed and purified.

3.3. The *in vitro* Ub hydrolysis activity of recombinant TsUCH37 was validated.

3.4. TsUCH37 was not able to hydrolyse Nedd8-AMC. The residue D12 was predicted to be important for the specific interaction with Nedd8. The substitution of this residue to N, the residue found in Nedd8 hydrolases, did not enable Nedd8 hydrolysis by TsUCH37.

3.5. TsUCH37 was co-precipitated with numerous *T. spiralis* proteasome components, including the putative *T. spiralis* ADRM1.

3.6. The compound LDN-57444 inhibited TsUCH37 Ub-AMC hydrolysis activity, and significantly reduced the viability of *T. spiralis* muscle larvae in culture.
Discussion

The initial aim of the project was to characterise TsUCH37, a DUB expressed by *T. spiralis* muscle larvae, by cloning the coding sequence and expressing recombinant protein. In order to do so, the true coding sequence had to first be confirmed. This was achieved using bioinformatics analysis, including gene predictions and alignments, followed by RACE-PCR. Once sequenced, RACE-PCR products provided experimental evidence that the gene prediction using AUGUSTUS software represented the true transcript of the TsUCH37 gene. This sequence did not include the extra 2 protein domains annotated in the database as being upstream of the peptidase_C12 DUB domain. The RACE-confirmed start and stop codons aligned well with multiple UCH37 orthologues. These results demonstrate that *T. spiralis* has a unique genome arrangement pattern. During the draft annotation 1 prediction programme, FGENESH *de novo*, was used to identify *T. spiralis* genes [232]. Because the FGENESH algorithm has been developed using other well-studied organisms, when used alone to scrutinise the *T. spiralis* genome, it may not necessarily be accurate for predicting *T. spiralis* genes.

During expression in *E. coli*, 6His-TsUCH37 was observed in a greater proportion in inclusion bodies of the bacteria. Methods of purification from inclusion bodies under denaturing conditions were therefore adopted, allowing an increased level of purity to be achieved and ensuring that the denatured enzyme could not be degraded by other proteases. On the down side, experimentally refolding proteins from a denatured state does not guarantee that they will reach a native and active confirmation. For this reason, the activity of TsUCH37 had to be confirmed before any further experiments investigating the function of the protein could be carried out.

6His-TsUCH37 was able to hydrolyse Ub-AMC at a physiologically relevant rate, confirming that the protein was refolded into the correct confirmation and that the activity was dependent on an active site cysteine. Recombinant 6His-TsUCH37 was therefore deemed a reliable reagent for the characterisation of the *T. spiralis* DUB.

The yeast, malaria parasite, mouse and human orthologues of the highly conserved UCH domain DUB, UCHL3, and the highly conserved human UCH domain DUB UCHL1, can process both Ub and Nedd8 [102-104,241,254-256]. Although most UCH37 orthologues, are specific to Ub, 2 parasite UCH37 orthologues, the *Plasmodium falciparum* UCH54 and the *Plasmodium knowlesi* UCH37, are able to hydrolyse both Ub and Nedd8. The function of the dual activity of the parasite UCH37 orthologues is
unknown. *P. knowlesi* has a UCHL3 gene and *P. falciparum* has been reported to express PfUCHL3 [102]. Furthermore, when analysed by BLAST both organisms have orthologous genes for the mammalian deNeddylation enzymes DEN1 (deNeddylation enzyme 1) and NEDP1 but not for SENP8 (SUMO/sentrin specific peptidase family member 8). It was hypothesised that the dual activities of the malaria parasite UCH37 proteins may serve a parasite-specific advantage. The *T. spiralis* parasite orthologue, TsUCH37 was therefore tested for its ability to hydrolyse Nedd8. Like the human orthologue HsUCH37, TsUCH37 was not able to hydrolyse Nedd8-AMC.

To try and explain the structural basis of dual versus mono activity, patterns of activity-based clustering of UCH37 and UCHL3 residues in multiple orthologues were highlighted. Five residues were identified that are conserved in only the enzymes that possess dual activity. These analyses were then combined with structural modelling analyses that predicted 4 residues to be important for the direct contact between PfUCHL3 and Nedd8 [104]. One residue was highlighted by both analyses that was not shown to be important for the interaction between YUH1 and Ub, and therefore possibly Nedd8-specific [243]. This residue was therefore mutated in the *T. spiralis* TsUCH37 coding sequence to match the corresponding residue found in dual deubiquitinating and deNeddylation enzymes. The TsUCH37 D12N mutant was unable to hydrolyse Nedd8 confirming that this residue is not a determinant of Nedd8 hydrolysis. This is not to say, however, that this residue is not predictive of dual activity.

These studies are on-going and the remaining 4 residues of the UCH37 orthologues listed in Table 1.A are currently undergoing mutagenesis. Residues will first be mutated to an alanine residue (because alanine is chemically inert) to confer the importance of the residue in substrate recognition or binding by assessing the effect of the mutation on Ub and Nedd8 hydrolysis activity. If found to be essential for activity, the residue will then be mutated to match the corresponding residue of the enzymes that have dual activity. These mutants will continue to be screened for Ub-AMC and Nedd8-AMC hydrolysis activity, with the aim of identifying the structural and evolutionary basis of dual Ub and Nedd8 activity in UCH37 and UCHL3 DUBs. *S. cerevisiae* (yeast) only express 1 UCH domain DUB, YUH1 [257]. YUH1 is most closely related to the mammalian UCHL3, but also shows homology with mammalian UCHL1 and UCH37 (in order of identity). YUH1 also has dual Ub and Nedd8 activity [242]. This may suggest that as lower eukaryotes evolved to express a wider range of UCH enzymes, the dual activity of the UCH37 orthologue was lost. The functional significance of the dual activity of PfUCH54 and PkUCH37 is unknown, and whether or not these DUBs are in
involved in the regulation of Nedd8 conjugation in the parasite remains to be determined.

Results suggested that the function of the *T. spiralis* parasite UCH37 may be more related to the function of the mammalian orthologues than the malaria parasite orthologues. The interaction of mammalian UCH37 with the proteasome enhances its Ub hydrolysis activity [32,258]. This interaction occurs via the direct binding of UCH37 to the proteasome subunit ADRM1. Alone human UCH37 is able to hydrolyse monoubiquitin, with ADRM1 this activity is 2-3-fold increased. Evidence suggests that UCH37 contains a C-terminal tail that acts in an auto-inhibitory manner [32,233]. This tail contains the binding site for ADRM1, and it is hypothesised that the binding of ADRM1 relieves the auto-inhibitory effect of the tail. UCHL1, UCHL3 and YUH1 do not associate with the proteasome and can only hydrolyse monoubiquitin [24,36,233,243,259,260]. Only in the presence of both ADRM1 and the 19S proteasome complex, is UCH37 able to hydrolyse di-ubiquitin [233]. The interaction of UCH37 with the proteasome therefore determines the specific function of this DUB as a proteasome-associated polyubiquitin hydrolase.

The co-precipitation of TsUCH37 with proteins of *T. spiralis* muscle larvae revealed that 6His-TsUCH37 associates with a large number of proteasome components, including the putative *T. spiralis* ADRM1. For an undetermined reason, human ADRM1 was found to non-specifically bind to Ni-NTA resin. It is therefore possible that TsADRM1 can also non-specifically bind to Ni-NTA resin. Considering that co-precipitation analyses were carried out using Ni-NTA suspended 6His-TsUCH37, it cannot be concluded that the presence of ADRM1 is due to a specific interaction with TsUCH37. However ADRM1 was also identified alongside TsUCH37 after immuno-precipitation of HA-Ub-VME with *T. spiralis* lysate during the MRes project (MRes report 2010, [1]). Since the HA-Ub-VME immuno-precipitation method did not involve Ni-NTA, TsADRM1 may have been immuno-precipitated in association with TsUCH37, or in association with the ubiquitin of the HA-Ub-VME probe itself via the pleckstrin like receptor for ubiquitin (Pru) domain that is present in ADRM1 orthologues [261]. The direct interaction between TsUCH37 and TsADRM1 would therefore require validation.

Numerous other putative *T. spiralis* proteasome subunits were co-precipitated with 6His-TsUCH37, the most abundant of which are found in the 19S base of the proteasome. In the mammalian system, ADRM1 associates with Rpn2 (subunit of the 19S base) anchoring UCH37 to the proteasome. Data therefore strongly suggests that TsUCH37 can be proteasome associated. It is therefore likely that the function of TsUCH37, as a proteasome associated DUB has been conserved throughout evolution.
Controlled proteasomal degradation of polyubiquitinated proteins is essential for cell homeostasis. Silencing UCH37 in human adenocarcinoma cells leads to apoptosis and the knockout of mouse UCH37 in vivo is lethal [36,262]. Knockdown of the drosophila UCH37 orthologue, which is also proteasome associated, causes death at an early stage of development [240] [263,264]. Deletion of YUH1 however, which is not proteasome associated, has no phenotype in yeast [265,266]. Since TsUCH37 is also able to associate with the T. spiralis proteasome, it was hypothesised that TsUCH37 might be essential for the survival of the T. spiralis parasite. TsUCH37 was therefore investigated for its potential as a drug target in vitro and then in parasites in culture.

I was unable to show that WP1130, a small molecule inhibitor previously shown to inhibit a number of UCH enzymes, can inhibit UCHL1, UCH37 or TsUCH37 activity by Ub-AMC assay. It is possible that this discrepancy in Ub-AMC data to that already published is due to differences in concentrations of recombinant protein used for Ub-AMC assay. A nM range for protein concentration was used by Kapuria et al. whereas a µM range was used in my experiments [246]. This was because below the µM range, no Ub-AMC hydrolysis activity was observed with DMSO only (negative control for inhibition). My protein preparations may have been less active than those reported by Kapuria et al. possibly containing a lower proportion of correctly folded protein. I therefore decided that LDN-57444 would be tested for the effect on TsUCH37 instead, and indeed, a potency of inhibition was observed similar to that for human UCHL1 [249].

Although the activity of recombinant 6His-TsUCH37 was specifically inhibited by LDN-57444, the compound is also known to inhibit the human UCH37, UCHL1 and UCHL3. If drugs were to target TsUCH37, or other parasitic DUBs for therapeutic effect during infection, they would have to be developed to select only the parasite protein and not the human orthologue. Considering that UCHL3 orthologue is annotated in the T. spiralis genome, and a putative T. spiralis UCHL3 was identified in muscle larvae lysate [1], it cannot be concluded that the inhibitory effect of LDN-57444 on T. spiralis muscle larvae viability was solely due to TsUCH37 inhibition. LDN-57444 however, is a specific UCH inhibitor, and does significantly reduce the viability of T. spiralis muscle larvae. Some UCH inhibitor compounds are able to specifically discriminate between the different UCH DUBs, such as the UCHL1 inhibitor 1 which does not inhibit UCHL3 despite 21% identity between the 2 proteins [267]. This shows that it is possible to specifically target different UCH DUBs, despite their level of conservation. Indeed, shortly after the experiments presented in this chapter were carried out, a specific inhibitor of UCH37, and USP14 (another mammalian proteasome-associated DUB) was discovered [82]. This drug has been recently shown to induce apoptosis in bone...
marrow cancer cells. This drug would make an interesting candidate for following up the *T. spiralis* viability studies, to determine if proteasome-associated DUBs, rather than all UCH domain DUBs, are essential for *T. spiralis* survival.

To identify structural specificities that can be exploited in drug identification or design, it is useful to have the crystal structures of proteins. Both the full length and the catalytic domain of TsUCH37 in complex with Ub-VME were solved (Morrow *et al.*) as a result of a collaboration with Chittaranjan Das [3]. This data allowed for the identification of differences between the human UCH37 and the *T. spiralis* UCH37 when the 2 structures were overlaid (appendix 7). The crossover loop of TsUCH37 remained disordered even in a packed crystal, where in the human UCH37 structure the loop was ordered. This disorder is not observed in the previously solved substrate-bound structures of human UCH-L1, UCH-L3, or in the unbound structure of human UCH37 [25,268,269]. Evidence suggests that the crossover loop of UCH DUBs, may be involved in substrate filtering, where a longer crossover loop allows the entry of larger substrates into the active site of the enzyme [270]. Differences in the confirmation of crossover loops between UCH37 orthologues, may therefore also confer specificity of inhibitors, especially if these inhibitors were active-site targeted.

Subtle changes in the electrostatic charges of the catalytic cleft were also observed, where the active-site cysteine of TsUCH37 flipped in orientation compared to its position in the human UCH37 structure. These changes may be TsUCH37-specific, or they may occur upon binding to Ub-VME. In which case, the equivalent human structure of HsUCH37\textsuperscript{cat}-Ub-VME would have to be solved for direct comparison. So far, data is only available for the unbound form of human UCH37 [25,252,253]. It is possible that the conformational changes observed and presented here for TsUCH37 would also occur in the substrate-bound form of human UCH37. Greater scrutiny of the direct comparison would be necessary before any conclusions on selective drug targeting could be made. The crystallography group at Purdue University continue to study the structure function relationship of deubiquitinating enzymes, using proteins such as TsUCH37 as informative models.

Results presented in chapter 3, demonstrate the initial stages of the project that aimed to investigate the role of the ubiquitin pathway during infection by *T. spiralis* from the point of view of the parasite. I then wanted to use the remainder of the PhD project to use the tools and skills presented during this chapter, to shift focus and investigate the role of the ubiquitin pathway in the direct host-parasite interactions during infection by *T. spiralis*. 

Chapter 3: Results
Chapter 4:
Looking for DUBs and ubiquitin conjugation enzymes in the secreted proteins of *T. spiralis* muscle larvae
Chapter 4: Looking for DUBs and ubiquitin conjugation enzymes in the secreted proteins of *T. spiralis* muscle larvae

Chapter 3 describes experiments that were designed to study the endogenous ubiquitin proteasome system of the parasitic nematode, *Trichinella spiralis*. A successful system for characterising functional *T. spiralis* DUB enzymes was established. This system was used to begin an investigation into the role of the ubiquitin proteasome system in the direct host-parasite interactions that occur during infection by *T. spiralis*. To do so, *T. spiralis* secreted proteins, the main mediators of host-parasite interaction, were studied. Following this, a multitude of experimental approaches were used to analyse the secreted proteins of *T. spiralis* muscle larvae for deubiquitinating, deNeddylating and ubiquitin conjugation activity.

**Background**

Pathogens use many strategies to directly manipulate host biochemistry to their own advantage during infection. One of these strategies involves pathogen-derived proteins that target and alter the activity of the host ubiquitin-proteasome pathway. This can lead to the modification of host biological processes such as immunity to infection, thus promoting the survival of the pathogen [271], [84]. Pathogen proteins can target the host proteasome, host deubiquitinating enzymes (DUBs) or host ubiquitin conjugation machinery, namely E1, E2 and E3 enzymes [85,89]. Alternatively, pathogens can express their own ubiquitin proteasome components that are able to mimic, interact with and modify the host ubiquitin proteasome system.

Prokaryotic organisms do not have their own ubiquitin-proteasome system. It was therefore surprising to discover that some prokaryotic pathogens, viruses and bacteria, express proteins with ubiquitin-proteasome system activity (Table 4.1, [93,96,97,99,100,272-275]). Because there is no ubiquitin pathway endogenous to the pathogen with which these proteins can interact, they are thought to specifically target the host ubiquitin-proteasome system. It is thought that these collaborations can lead to the ubiquitination and subsequent degradation of either host-derived or bacteria-derived proteins involved in the infection process.
Eukaryotic pathogens such as parasites have their own ubiquitin pathway, as was demonstrated for the mammalian parasitic nematode, *T. spiralis* during chapter 3. Human apicomplexan parasites also express a functional ubiquitin proteasome pathway and parasite-derived DUBs have been characterised for *Plasmodium falciparum* and *Toxoplasma gondii* [102,103]. It is not clear however whether or not these DUBs are specifically involved in the infection process because separating proteins that are specifically host-targeted from those that are parasite's ‘own’ is a particular challenge. This is especially true when studying the intracellular stages of parasites such as *Plasmodium falciparum*. For parasitic worms, along with the parasite surface proteins, the parasite secreted proteins (SP) are thought to be the chief mediators of host-parasite interaction during infection. Although apicomplexan parasites secrete proteins into the host cell, it is extremely difficult to isolate these proteins from host cell proteins, and not possible to culture the parasite without the host cell.
Table 4.1 List of pathogen-derived ubiquitin-proteasome enzymes

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Protein</th>
<th>Activity</th>
<th>Host-targeted</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Virus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adenovirus</td>
<td>adenain</td>
<td>DUB</td>
<td>✓</td>
<td>Balakirev 2002 J Virol</td>
</tr>
<tr>
<td>adenovirus</td>
<td>E4orf6</td>
<td>E3</td>
<td>✓</td>
<td>Querido 2001 Genes Dev</td>
</tr>
<tr>
<td>herpesvirus</td>
<td>UL36</td>
<td>DUB</td>
<td>✓</td>
<td>Kattenhorn 2005 J Virol</td>
</tr>
<tr>
<td>herpesvirus</td>
<td>ICP0</td>
<td>E3</td>
<td>✓</td>
<td>Everett 2006 J Virol</td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chlamydia</td>
<td>ChlaDub1</td>
<td>DUB</td>
<td>✓</td>
<td>Misaghi 2006 Mol Microbiol</td>
</tr>
<tr>
<td>salmonella</td>
<td>SopA</td>
<td>E3</td>
<td>✓</td>
<td>Zhang 2006 Mol Microbiol</td>
</tr>
<tr>
<td>salmonella</td>
<td>AvrA</td>
<td>DUB</td>
<td>✓</td>
<td>Ye 2007 Am J Pathol</td>
</tr>
<tr>
<td>yersinia</td>
<td>YopJ</td>
<td>DUB</td>
<td>✓</td>
<td>Orth 2000 Science</td>
</tr>
<tr>
<td><strong>Parasite</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. falciparum</td>
<td>UCH54</td>
<td>DUB</td>
<td>?</td>
<td>Artavanis-Tsakonas 2006 Mol Microbiol</td>
</tr>
<tr>
<td>P. falciparum</td>
<td>UCHL3</td>
<td>DUB</td>
<td>?</td>
<td>Frickel 2007 Cell Microbiol</td>
</tr>
<tr>
<td>T. gondii</td>
<td>UCHL3</td>
<td>DUB</td>
<td>?</td>
<td>Frickel 2007 Cell Microbiol</td>
</tr>
<tr>
<td>T. spiralis</td>
<td>UCH37</td>
<td>DUB</td>
<td>?</td>
<td>White 2011 PLoS NTD</td>
</tr>
</tbody>
</table>

**Table 4.1 Pathogen-derived ubiquitin pathway enzymes**

List of pathogen-derived ubiquitin pathway enzymes and the publication in which their identification and host-targeting strategy during infection is reported.
The parasitic nematode *Trichinella spiralis* secretes proteins into the host during both the extracellular stages of infection and the intracellular stage of chronic infection, where larvae invade and reside inside mature skeletal muscle cells [169,171]. *T. spiralis* muscle larvae can be isolated from infected muscle tissue and cultured without host tissue. The secreted proteins can then be collected from culture supernatants and studied separately to host proteins [116,276]. Since these proteins are actively secreted by the parasite it is hypothesised that they are specifically host targeted. Furthermore, during the muscle stage of infection, *T. spiralis* induces changes in host cell transcription that lead to changes in the host cell cycle and differentiation state [159,162,163,165]. The ubiquitin-proteasome system heavily regulates the proliferation and differentiation of eukaryotic cells [5,52]. It was therefore hypothesised that *T. spiralis* muscle larvae SP contains proteins that are able to interact with and modulate the host ubiquitin-proteasome system.

The experiments carried out during my MRes project showed that the inhibitor probe HA-Ub-VME can be successfully used for identifying functional *T. spiralis* DUBs [1]. The inhibitor probe HA-Ub-VME is composed of the human Ub protein with an N-terminal haemagglutinin (HA) epitope tag and a C-terminal electrophilic group, VME [223,226]. Due to the high level of conservation of ubiquitin amongst eukaryotes, these probes have been used to successfully identify DUBs from a number of pathogenic organisms of medical importance, including the human apicomplexan parasites *Plasmodium falciparum* and *Toxoplasma gondii* [102,103]. They have also been used to identify Ub-proteasome components expressed by bacteria and viruses [93,96]. These probes were therefore used to look for DUB activity in the secreted proteins of the muscle larvae of *T. spiralis*.

Most DUB enzymes are cysteine (thiol) based in activity. When a protein sample containing a cysteine (thiol)-based DUB(s) is reacted with HA-Ub-VME, the ubiquitin is recognised by the DUB, and the electrophilic group (or ‘warhead’), VME reacts with the thiol of the DUB active-site cysteine, forming a covalent thioether bond and ‘trapping’ the DUB(s) ([223,226] Figure 2.1). These electrophilic warheads may vary, and for each warhead a specific affinity for different DUB enzymes is observed. Some even show affinity for E2 and E3 enzymes [277]. To control for cysteine-specific activity, the protein sample of interest can be pre-incubated with N-ethylmaleimide (NEM). NEM also reacts with the thiol group of the active site cysteine, rendering the cysteine protease inactive. Once a sample has been reacted with the probe, in the presence and absence of NEM, anti-HA epitope tag antibodies can then be used to immuno-precipitate or immuno-blot probe-enzyme complexes for further analyses.
In this chapter *T. spiralis* secreted proteins were analysed for DUB and deNeddylation activity using HA-Ub-VME, and using an additional probe based on the ubiquitin-like protein Nedd8. Following this, a number of techniques including modified inhibitor probe assays, *in vitro* ubiquitin conjugation assays, bioinformatics and tandem mass spectrometry were then used to investigate the presence of E1, E2 or E3 activity in the *T. spiralis* secreted proteins.
Chapter 4: Results

Figure 4.1

A. HA-Ub-VME

B. HA-Ub-DUB

* active site cysteine

** irreversible thioether bond

Figure 4.1 The mechanism of the HA-Ub-VME probe

A. HA epitope tag and an N-terminal electrophilic group (VME). A DUB recognises the Ub of HA-Ub-VME and the warhead, VME, reacts with the active site cysteine, forming an irreversible covalent thioether bond and creating a HA-Ub-DUB complex. Curly arrows indicate the movement of an electron pair. To control for cysteine-specific activity, the protein sample of interest can be pre-incubated with N-ethylmaleimide (NEM). NEM reacts with the active site cysteine rendering the DUB inactive. B. Once a sample has been reacted with the probe, in the presence or absence of NEM, anti-HA tag antibodies can then be used to immuno-precipitate HA-Ub-DUB complexes.
Chapter 4 Results:

Looking for DUB and deNeddylating activity in *T. spiralis* secreted proteins

4.1. Production of the ubiquitin-based inhibitor probe HA-Ub-VME

The vector for expression of the inhibitor probe was prepared previously as described by Hemelaar and Borodovsky *et al.* and kindly donated by Hidde Ploegh (Whitehead Institute for Biomedical Research) [223,256]. Briefly, the coding sequence of human ubiquitin, lacking both the C-terminal glycine (Ub$_{75}$) and a stop codon, was cloned with an N-terminal haemagglutinin (HA) tag, into the multiple cloning site of pTYB1 vector. This allowed for transcription of the coding sequence in-frame with a C-terminal intein gene from *Saccharomyces cerevisiae* and a chitin-binding domain (CBD) from *Bacillus circulans* (Figure 4.2 A). The chitin-binding domain allows affinity purification of the fusion protein on chitin resin (Figure 4.2 B). The intein gene encodes an intein domain that allows protein cleavage after addition of a thiol reagent like 2-mercaptoethane sulfonate Na (MESNa). HA-Ub-intein/CBD was expressed in *E. coli* as described (materials and methods section 2.23). A fusion protein of approximately 60 kDa was observed after SDS-PAGE analysis of the bacterial lysate (Figure 4.2 D (a)). The fusion protein was then purified on chitin resin before cleavage of the intein domain was induced by adding the thiol reagent MESNa (Figure 4.2 C). This eluted HA-Ub-MESNa (approximately 11 kDa) from the column, leaving the chitin binding domain on the resin (Figure 4.2 D (b)). HA-Ub-MESNa was concentrated and reacted with the electrophilic ‘warhead’ glycine-VME in the presence of the activating reagent N-Hydroxysuccinimide (NHS) to form HA-Ub-VME (Figure 4.1 E).

In order to remove inactive probe, un-reacted warhead, un-reacted HA-Ub-MESNa and impurities, the probe was then further purified by ion exchange chromatography (FPLC) using a strong cation exchange (MonoS) column (Figure 4.3 A). HA-Ub-MESNa was bound to the FPLC MonoS column in buffer A. Buffer A is a sodium acetate-based buffer at pH 4.5, optimised for the binding of the HA-Ub-VME to the MonoS column based on its isoelectric point of 5.85. Once bound and washed on the column in buffer A, buffer B (buffer A supplemented with 1M NaCl), was slowly introduced to gradually increase the concentration of Na$^+$ ions. This changes the net charge of the bound protein resulting in its elution from the column. Protein was then collected fraction by fraction. As FPLC fractions were collected, the UV absorbance of each fraction at 280 nm (measured in milli-absorbance units, mAU) was taken to
determine the protein content. Two small peaks in mAU were observed (Figure 4.3 A). FPLC fractions were analysed by SDS-PAGE and coomassie staining to confirm the presence of protein at the expected size of 11 kDa (data not shown).

Fractions containing probe were concentrated and tested for activity. The reactivity of *T. spiralis* TsUCH37 with HA-Ub-VME was previously confirmed during my MRes project [1]. Therefore the recombinant purified 6His-TsUCH37 was used as a test for probe activity. Complex formation of the probe with 6His-TsUCH37 was analysed by SDS-PAGE and coomassie staining (Figure 4.3 B). 6His-TsUCH37-probe complexes were observed as an electrophoretic mobility 'up-shift' of approximately 11 kDa. This up-shift was not present when 6His-TsUCH37 alone was analysed, however a smaller band was consistently observed that is likely to be a breakdown product or contaminant in the 6His-TsUCH37 sample. Fraction A15 did not contain any probe. Fractions B5 to B1 contained active probe, for which a band of that corresponded to the 6His-TsUCH37-probe complex was observed (Figure 4.3 B (a)). An additional larger band was observed for fractions B4 to B2. Fraction B11 contained some active probe, although some un-reacted probe was still observed at approximately 11 kDa, indicating that this peak may represent mainly un-reacted HA-Ub-MESNa, or inactive HA-Ub-VME. Fraction C5 did not contain any probe.
Figure 4.2 Production of the HA-Ub-VME probe

A. Human (Hs) ubiquitin with an N-terminal HA tag was cloned into pTYB1 creating a C-terminal intein/chitin binding domain fusion protein (CBD). The vector was kindly donated by Hidde Ploegh (Whitehead Institute). B. The HA-Ub-intein/CBD fusion protein was purified using chitin resin that binds to the CBD. C. The addition of the MESNa mediates the cleavage of ubiquitin from the intein/CBD, releasing HA-Ub-MESNa. D. HA-Ub-intein/CBD was expressed by E. coli (a) and purified using the chitin column and on-column cleavage technique, forming HA-Ub-MESNa (b). E. HA-Ub-MESNa was ligated to glycine-VME forming HA-Ub-VME.
Figure 4.3

**A.** HA-Ub-VME was purified by ion exchange chromatography. The absorbance (mAU) of fractions eluted from the column by buffer B was measured as they were collected. The shaded fractions were tested for activity: A15 (blue), B11 (green), B5-B1 (yellow) and C5 (lilac).

**B.** Fractions were tested for reactivity with 6His-TsUCH37. Reactions were analysed by SDS-PAGE and coomassie staining. As a control TsUCH37 was analysed alone (b). Probe-TsUCH37 complexes (a) demonstrate a shift in electrophoretic mobility (a). An unidentified signal was observed (c) and probe alone (d).

**C.** The effect of NEM and isopropanol on TsUCH37 and HA-Ub-VME reactions was analysed using probe fraction B3. Reactions were analysed by coomassie staining and His blot. Probe-TsUCH37 complexes (a) demonstrate a shift in electrophoretic mobility (a) above the unreacted TsUCH37 signal (c). NEM also caused a shift in electrophoretic mobility of TsUCH37 (b).
There is some evidence that Ub is able to dimerise in a non-covalent manner [278]. The aggregation of a protein under reducing conditions during SDS-PAGE, would imply SDS resistant aggregation. The reducing agent dithiothreitol (DTT) is included in both the probe reaction buffer and the protein (SDS) loading buffer and theoretically, disulphide linkages are prevented in this buffer. However DTT is not included in SDS-PAGE running buffer, and there is evidence that some proteins are able to aggregate during the stacking stage of SDS-PAGE [279]. If I propose that the probe is able to non-covalently aggregate, the extra unexpected bands that were observed may represent a probe dimer weighing approximately 22 kDa, and a probe dimer in complex with 6His-TsUCH37, weighing approximately 60 kDa (38 kDa + 22 kDa).

To test for cysteine-specific reactivity, probe fraction B3 was then further tested for reactivity with and without the cysteine protease inhibitor NEM. Preliminary experiments indicated that either the NEM, or the NEM vehicle isopropanol, was having an effect on the electrophoretic mobility of the proteins themselves (data not shown). An experiment was therefore set up to verify the activity of the probe fraction B3 with 6His-TsUCH37, and to demonstrate the effect of NEM on protein electrophoretic mobility (Figure 4.3 C). This time, coomassie stain and anti-His-tag antibodies were used to analyse the reactions by immuno-blot. Figure 4.3 C shows that isopropanol itself has no effect on the TsUCH37 mobility, and that probe fraction B3 indeed reacts with TsUCH37, forming a complex of an expected size. This reactivity is silenced by NEM and NEM also causes the un-reacted TsUCH37 to have a reduced mobility in SDS-PAGE, even in the absence of the probe. This observation had to be taken into account in subsequent experiments so as not to overlook subtle changes in the mobility of proteins when testing Ub-based inhibitor probes.

4.2. Reaction of ubiquitin-based inhibitor probes with T. spiralis secreted proteins

In addition to the HA-Ub-VME probe, another probe that was previously produced by Hemelaar et al. was kindly provided Hidde Ploegh [226]. This probe is based on another human Ub-like modifier protein, Nedd8. DeNeddylation enzymes remove Nedd8 from substrate proteins using the same mechanism as deubiquitinating enzymes. The probe contained a FLAG epitope tag for immuno-blol/precipitation and the electrophilic warhead, vinyl sulphone (VS), FLAG-Nedd8-VS. The mechanism of reactivity of this probe with deNeddylation enzymes is as described earlier for HA-Ub-
VME in Figure 4.1. The putative *T. spiralis* Nedd8 orthologue (GI: 339239065) shows 65.6% identity to human Nedd8 (Figure 4.4 E).

The secreted proteins from *T. spiralis* muscle larvae were collected and 300 ug were reacted with HA-Ub-VME and FLAG-Nedd8-VS, in the presence or absence of the cysteine protease inhibitor NEM as a control. Reactions were then separated by SDS-PAGE and analysed by immuno-blot using anti-HA-tag antibodies and anti-FLAG-tag antibodies (Figure 4.4 B and C). Most bands observed after the reaction of HA-Ub-VME with *T. spiralis* secreted proteins were present in both the samples (with and without NEM), indicating either probe aggregation, or that reactivity with the probe was not cysteine dependent and therefore unlikely to be DUB specific (Figure 4.5 B). There were however a few bands that appeared to have been present only in the NEM-free sample. No cysteine-dependent reactivity was observed after the reaction of *T. spiralis* secreted proteins with FLAG-Nedd8-VS (Figure 4.4 C).

To confirm whether or not the signal observed after the reaction of *T. spiralis* secreted proteins with HA-Ub-VME (in the absence of NEM) might be DUB-specific, the experiment was scaled-up. A large amount (3 mg) of *T. spiralis* secreted protein was reacted with HA-Ub-VME. Protein-probe complexes were then immuno-precipitated and purified using anti-HA affinity resin. These were eluted from the resin and analysed by SDS-PAGE. This experiment was repeated a 2nd time. Proteins from the 1st experiment were visualised by coomassie staining and proteins from the 2nd experiment were visualised by silver staining (Figure 4.5 A and B respectively). All bands present in the coomassie-stained gel were manually excised for analysis. From the silver-stained gel, bands that appeared present in only the NEM-free sample, and corresponding areas of the NEM-present sample, were manually excised for analysis. This constituted 19 bands from the coomassie-stained gel labelled 1-19 and 13 bands from the silver-stained gel labelled A-M (Figure 4.5 A and B respectively). These proteins were analysed by tandem mass spectrometry (LC/MS/MS). Peptides were then matched against the *T. spiralis* draft genome annotation database using SEQUEST software, [280]. In bands 1-19, a total of 100 putative *T. spiralis* proteins were identified and in bands A-M, a total of 57 putative *T. spiralis* proteins were identified. A list of all the putative *T. spiralis* protein matches can be found in the supplementary appendices 2 and 3. Eric Spooner of the Whitehead Institute, Boston carried out the mass spectrometry analysis and SEQUEST searches.
Figure 4.4 Reaction of T. spiralis secreted proteins with the HA-Ub-VME probe

A. Schematic of the reaction of a DUB with the HA-Ub-VME probe. The DUB recognises Ub and the active-site cysteine attacks the modified (VME) C-terminal glycine. An irreversible thioether bond is formed between the DUB and the glycine-VME. Curly arrows indicate the movement of an electron pair. Anti-HA antibodies with react with the HA epitope tag of the HA-Ub-DUB complex, allowing immuno-blot (IB) or immuno-precipitation (IP) analyses. B. Reaction of the secreted proteins of T. spiralis muscle larvae (T.sp SP) with HA-Ub-VME. C. Reaction of T. spiralis secreted proteins (SP) with FLAG-Nedd8-VS. Reactions were carried out in the presence and absence of the cysteine protease inhibitor NEM and analysed by immuno-blot using anti-HA or anti-FLAG antibodies. A possible enrichment of signal was observed after the reaction with HA-Ub-VME in the absence of NEM (red arrows). All molecular weights are indicated in kDa. D. Alignment of human (Hs) ubiquitin with T. spiralis (Ts) ubiquitin shows 98% identity. E. Alignment of human (Hs) Nedd8 with T. spiralis Nedd8 shows 59.2% identity.
Figure 4.5  Proteomic analysis of the HA-Ub-VME reaction with *T. spiralis* secreted proteins

The secreted proteins (SP) of *T. spiralis* muscle larvae were reacted with HA-Ub-VME, in the presence and absence of the cysteine protease inhibitor NEM, followed by immunoprecipitation (IP) using anti-HA antibodies. Two separate reactions were carried out and separated by SDS-PAGE. The first was analysed by coomassie staining (A) and the second by silver staining (B). Bands that appeared in the NEM negative sample only were identified and manually excised (numbers or letters) along with the corresponding areas of the NEM positive sample. Proteins in these bands were analysed by tandem mass spectrometry (LC/MS/MS). All molecular weights are indicated in kDa. C. The putative *T. spiralis* ubiquitin conjugating enzyme TsUBE2N (GI: 316965577) was identified in band 5 of the coomassie-stained HA-Ub-VME/*T. spiralis* SP experiment. Two peptides covering 21% of the putative protein sequence were identified. The UBCc (ubiquitin conjugation) domain is boxed in red.
Although many putative *T. spiralis* proteins were identified (see supplementary appendix 2 and 3 for all LC/MS/MS data), none were annotated as DUB or deNeddyllating enzymes. All protein matches annotated as ‘uncharacterised’ were further analysed for conserved domain architecture by BLAST and SMART (Simple Modular Architecture Research Tool, [210,211]). No DUB or deNeddyllase domains were identified. However, 3 ubiquitin-related protein sequences were identified (Table 4.2). The first match, the ‘Ubiquitin family member protein (fragment, (GI: 339233028), is the putative *T. spiralis* ubiquitin protein itself, which is labelled as a fragment because the annotated sequence contains a single (mono) ubiquitin followed immediately by a fragment of another ubiquitin whose sequence is incomplete. In order to determine the source of the *T. spiralis* ubiquitin protein, *T. spiralis* secreted proteins and lysate were separated by SDS-PAGE and reacted with antibodies raised against human ubiquitin by immuno-blot. For positive controls, HEK 293T lysate and lysate of *T. spiralis* muscle larvae were analysed in parallel. *T. spiralis* secreted protein contained many proteins that were able to react with the anti-Ub antibodies at various sizes, and a protein at approximately 10 kDa that would correspond to the *T. spiralis* Ub itself (Figure 4.6). Peptides matching the *T. spiralis* ubiquitin family protein may therefore have derived from the human Ub-based probe itself, since the level of identity between the human Ub and the *T. spiralis* Ub is 97.4%, or from *T. spiralis* secreted proteins that are themselves modified by ubiquitin.
Table 4.2 Table of ubiquitin-related *T. spiralis* protein matches

<table>
<thead>
<tr>
<th>Putative <em>T. spiralis</em> protein</th>
<th>Accession Number</th>
<th>Predicted molecular Weight</th>
<th>No. of unique peptides</th>
<th>Coomassie</th>
<th>Silver-stain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(- NEM)</td>
<td>(+NEM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bands 1-11</td>
<td>Bands 12-19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-NEM)</td>
<td>(+NEM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bands A-F</td>
<td>Bands G-M</td>
</tr>
<tr>
<td>RWD domain protein</td>
<td>gj</td>
<td>316972581 (+3)</td>
<td>77 kDa</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>ubiquitin</td>
<td>gj</td>
<td>339233028 (+2)</td>
<td>23 kDa</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>UBE2N</td>
<td>gj</td>
<td>316965577 (+1)</td>
<td>20 kDa</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>

Table 4.2 Ubiquitin-related proteins identified by proteomic analysis of the HA-Ub-VME reaction with *T. spiralis* proteins

Table showing the 3 ubiquitin-related proteins that were identified by LC/MS/MS analysis of the proteins that were immuno-precipitated from 2 separate reactions (coomassie-stained or silver-stained) of the secreted proteins of *T. spiralis* muscle larvae with the HA-Ub-VME probe.
Human cell (HEK 293T) lysate and *T. spiralis* lysate were included as controls to analyse the secreted proteins (SP) of *T. spiralis* muscle by SDS-PAGE and immuno-blot (IB) using anti-ubiquitin antibodies. All molecular weights are indicated in kDa.
The 2nd ubiquitin-related protein identified by LC/MS/MS was annotated as a RING finger and WD-repeat (RWD) domain protein, GI: 316972581 with a predicted size of 77 kDa. RWD domains are RING finger and WD-repeat domains, and are often found in E3 ligases [281,282]. Although RWD proteins have been shown to interact with ubiquitin, their function is not well understood. When further analysed by BLAST alignment, the N-terminal portion of the putative T. spiralis RWD protein sequence matched another T. spiralis protein, annotated as the 53 kDa excretory-secretory antigen (GI: 805126). This protein was previously characterised as a dominant T. spiralis antigen whose sequence and size has been confirmed by proteomic analysis [196,283,284]. In these studies the authors do not observe an RWD domain as part of the sequence. When multi-orthologous alignments were carried out the sequence aligned with those from various species as 2 separate proteins. The N-terminal region aligned with a T. spiralis sequence that does not contain the RWD domain, the 53 kDa excretory-secretory antigen that is conserved amongst other species of Trichinella only. The C-terminal region aligned with RWD domain proteins from various other species. Peptides identified by LC/MS/MS only aligned with the N-terminal region.

Reactivity with the probe was found in both the NEM-free sample and the NEM-present sample, and thus was not cysteine-dependent (Figure 4.5 A and B). Furthermore, peptides that matched the annotated RWD protein were found most predominantly in bands C and J of the silver-stained gel and bands 2 and 13 of the coomassie-stained gel that represents a molecular weight of approximately 55 kDa (indicating a protein of 44 kDa in complex with an 11 kDa probe). These peptides only matched the N-terminal half of the sequence, and not any of the RWD domain that comprises the C-terminal portion of the annotated sequence. It therefore could be that this sequence has been mis-annotated.

The 3rd Ub-related protein was annotated as a putative ubiquitin-protein ligase that, when further analysed by BLAST, actually contained a UBCc domain. UBCc domains are characteristic of ubiquitin conjugation (rather than ligase) enzymes. The putative T. spiralis E2 GI: 316965577 is predicted to weigh 20.5 kDa, but matched 2 unique peptides found in band 5 (corresponding to approximately 40 kDa) of the coomassie-stained gel (Figure 4.5 A). The 2 unique peptides covered 21% of the putative T. spiralis E2 sequence (Figure 4.5 C). When analysed by BLAST alignment the sequence was found to be orthologous to the human E2 enzyme UBE2N (S. cerevisiae UBC13 and C. elegans ubc-1). Mammalian UBE2N is involved in lysine-63-linked polyubiquitin chain assembly [285,286]. This kind of ubiquitin chain is not predominantly involved in protein degradation, rather other proteasome-independent
processes such as the transcriptional activation of target genes [287,288]. When the experiment was repeated, peptides from the silver-stained gel did not match this protein, despite bands D, E, K and L being excised from a region of the gel that could correspond to band 5 of the coomassie-stained gel (Figure 4.5 B).

The sequence was then further analysed by bioinformatics programmes to ascertain the likelihood of true secretion of the protein by *T. spiralis*. The putative *T. spiralis* UBE2N sequence predicts a 179 amino acid protein. When analysed for protein architecture and conserved domains by SMART, residues 32 to 176 span the conserved UBCc domain (Figure 4.5 C). SignalP is an algorithm that identifies signal peptide cleavage sites and therefore signal peptides for membrane localisation, translocation or secretion [219]. SignalP analysis of the putative TsUBE2N did not identify a signal peptide or cleavage site. WolfPSORT is a programme that predicts sub-cellular localisation motifs based on sequence orthology. This includes searching for orthology with sequences that are often involved in secretory pathways, and proteins that are often located extracellularly [217]. WolfPSORT gives a list of potential locations of the protein, in order of a score of likelihood. For example, the wolfPSORT score for the human secreted protein insulin is as follows:

extr: 32.0 (only one possible location for this protein - extracellular)

WolfPSORT predicted the putative TsUBE2N to be predominantly located in the cytoplasm or targeted to the mitochondria or nucleus:

wolfPSORT score: mito: 22.0, cyto: 7.0, cyto_nucl: 5.0

The human UBE2N and the putative *T. spiralis* UBE2N are 87.2% identical. An antibody was purchased that was raised against a portion of the human UBE2N that shows 75% identity (Figure 4.7 A). The secreted proteins of *T. spiralis* muscle larvae, lysate of muscle larvae and lysate of HEK 293T cells were reacted with the antibody (Figure 4.7 B). Proteins were also reacted with antibodies (kindly provided by Kleoniki Gounaris) that react with an abundant *T. spiralis* secreted protein, the 5’nucleotidase (5’NT) enzyme, GI: 22656349, previously characterised by Gounaris et al. [197]. Both *T. spiralis* secreted proteins and *T. spiralis* lysate reacted with the known secreted 5’NT at approximately 60 kDa. The human UBE2N is 17 kDa. Proteins from HEK lysate, of approximately 17 kDa, reacted with the anti-HsUBE2N, confirming specificity of the antibody. Although proteins in the *T. spiralis* lysate reacted with anti-UBE2N antibodies, showing a doublet band at approximately 15 kDa, there was no reactivity in the *T. spiralis* SP. Results indicated that although *T. spiralis* muscle larvae express
a UBE2N homologue, it is unlikely that it is secreted. *T. spiralis* muscle larvae were
cultured for 4 days and the secreted proteins were harvested at the end of each day.
By the 4th day, a significant proportion of parasites may have begun to die. Therefore
peptides matching TsUBE2N may have been released from dead or dying larvae, and
not present as a result of the active secretion of the protein.
Figure 4.7 Analysis of TsUBE2N

A. Alignment of human (Hs) UBE2N and T. spiralis (Ts) putative UBE2N showing 87.2% identity. To analyse secretion of TsUBE2N by T. spiralis, an antibody raised against a portion of the human protein (immunogen sequence shaded in yellow) was reacted with T. spiralis secreted proteins (SP), T. spiralis lysate and human cell (HEK) lysate (B). As a control, the same samples were reacted with an antibody raised against an abundant T. spiralis secreted protein, 5’NT. Anti-5’NT was kindly donated by Kleoniki Gounaris (Gounaris et al. 2004 [197]).
4.3. Verification of the absence of DUB and deNeddylating activity in the *T. spiralis* secreted proteins

No DUB or deNeddylating activity could be detected in the secreted proteins of *T. spiralis* muscle larvae using HA-Ub-VME and FLAG-Nedd8-VS. To verify the absence of DUB and deNeddylating enzymes in *T. spiralis* secreted proteins, AMC hydrolysis assays were carried out. A description of the mechanism of an AMC assay can be found in section 3.3. AMC conjugates of the human orthologue of ubiquitin and Nedd8 were used to test for ubiquitin and Nedd8-specific hydrolysis activity in secreted proteins of *T. spiralis* muscle larvae. Lysate of *T. spiralis* muscle larvae was tested as a positive control, and purified recombinant PfUCHL3, the *Plasmodium falciparum* UCH-L3 orthologue, (at a saturating concentration) was tested as an additional positive control because it was previously shown to have both Ub and Nedd8 hydrolysis activity [103]. HEK lysate, *T. spiralis* lysate and *T. spiralis* SP were all used at 2 mg/ml. Protein samples were mixed with each AMC substrate and fluorescence, corresponding to the release of AMC from the conjugate, was measured every minute for 15 minutes (Figure 4.8 A and B). Lysate of *T. spiralis* muscle larvae gave rise to Ub-AMC and Nedd8-AMC hydrolysis activity that was comparable to the activity of PfUCH-L3 or HEK lysate (Figure 4.8 A and B). No Ub-AMC or Nedd8-AMC hydrolysis activity was observed in *T. spiralis* secreted proteins, consistent with the lack of activity measured by inhibitor probe assays. All activities were silenced by pre-incubation of the proteins with NEM, denoting cysteine dependent hydrolysis.

Using the Ub- and Nedd8-based inhibitor probes and Ub and Nedd8 AMC conjugates, no DUB or deNeddylating activity was detected in *T. spiralis* secreted proteins. Immunoprecipitation of HA-Ub-VME-SP complexes did however suggest the potential presence of a ubiquitin conjugating (E2) enzyme TsUBE2N, the secretion of which by *T. spiralis* could not be confirmed by immuno-blot. This prompted further investigation into the possibility of the secretion of Ub-conjugation machinery by *T. spiralis* muscle larvae.
Figure 4.8

A  Ub-AMC hydrolysis by T. spiralis SP

B  Nedd8-AMC hydrolysis by T. spiralis SP

Figure 4.8 Testing for Ub and Nedd8 hydrolysis activity of T. spiralis secreted proteins

The secreted proteins (SP) of T. spiralis muscle larvae were tested for reactivity with (A) Ub-AMC and (B) Nedd8-AMC. T. spiralis lysate and the P. falciparum dual deNeddylating and deubiquitinating enzyme PfUCHL3 were also tested as controls. All samples were pre-incubated with the cysteine protease inhibitor NEM as a negative control for cysteine dependent activity. Fluorescence (representing the release of free AMC) was measured in relative fluorescence units (RFU) over time (minutes). The hydrolysis of Ub-AMC and Nedd8-AMC by each sample was analysed in triplicate. Data are presented as the mean of triplicate values from a single experiment and the error bars represent SEM. Consistent data were obtained from similar experiments performed using different concentrations of protein and different protein batches.
Looking for ubiquitin activating (E1), conjugating (E2) and ligating (E3) activity in *T. spiralis* secreted proteins

4.4. Production of modified ubiquitin-based inhibitor probes

Ubiquitin-based inhibitor probes, including Ub-VME, can interact with E1, E2 and E3 enzymes as well as DUBs, albeit with a lower affinity, the level of which depends on the electrophilicity of the warhead [277]. Ubiquitin activating (E1), conjugating (E2) and ligase (E3) enzymes are also reliant on a catalytic cysteine for transient interaction with Ub during ubiquitination. The greater the electrophilicity of the warhead of a ubiquitin-based inhibitor probe, the greater the affinity for ubiquitin conjugation enzymes; E1, E2 and E3s. Therefore, the purpose of these experiments was to produce HA-tagged-Ub-based probes with more electrophilic warheads than VME. Love *et al.* demonstrated that 2 warheads, glycine 2,6-trifluoromethylbenzyloxymethyl-ketone (TF$_3$BOK) and alpha-amino-beta-lactone (Lac) have a higher affinity for Ub conjugation enzymes than vinylmethylester (VME, [277]).

These warheads were also shown to react with a largely different panel of proteins to the VME warhead. These warheads were therefore chosen based on their divergent affinities to the VME, which had already been tested with the *T. spiralis* secreted proteins.

The TF$_3$BOK and Lac warheads were synthesised in collaboration with Jennifer Ward under the supervision of Ed Tate at the Institute of Chemical Biology, Imperial College London (Figure 4.9 and 4.10 A). HA-Ub-MESNa was produced as described earlier (section 4.1) and reacted with each warhead at pH 8, under the direction of the methods previously reported by Love *et al.* A significant level of insolubility of each warhead was observed, thus reducing the ratio of warhead to HA-Ub-MESNa during the ligation step. Ligation samples were then purified by FPLC as described earlier (Figure 4.9 and 4.10 B). Fractions with peak absorbance were analysed by SDS-PAGE and coomassie staining to confirm presence of the probe at the expected size (Figure 4.9 and 4.10 C). An additional larger band was also observed, possibly representing aggregates/ doublets. Love *et al.* reported the identification of the human UCH37 (UCH-L5) using both TF$_3$BOK and Lac warheads and TsUCH37, the *T. spiralis* UCH37 orthologue, was originally identified using the human Ub-based probe HA-Ub-VME. It was hypothesised that the lac and TF$_3$BOK probes would also be able to react with TsUCH37 and the His-tagged recombinant protein was used to test for probe activity. Probe containing fractions were reacted with 6His-TsUCH37 and
analysed by SDS-PAGE and immuno-blot using anti-HA antibodies. The previously produced HA-Ub-VME probe (section 4.1) was also reacted with 6His-TsUCH37 as a positive control and as a negative control, 6His-TsUCH37 was pre-incubated with NEM before addition of the fraction (Figure 4.9 and 4.10 D).

No probe-6His-TsUCH37 complexes were detected in any fraction, and therefore no active TF$_3$BOK or Lac probes were present. In order to try and explain the lack of active probe, the warheads were then analysed by mass spectrometry (LC/MS) by Jennifer Ward. When the probes are produced they are in a slightly acidic state, they are then buffered to pH 8 for the ligation reaction. The stability of the warheads at pH 8 was therefore analysed (Jennifer Ward MRes research manuscript, Institute of Chemical Biology 2012). It was found that significant degradation of the TF$_3$BOK warhead occurred within 1 hour of incubation at pH 8. Methods of ligation and FPLC were adapted for optimisation, but multiple attempts to produce active HA-Ub-TF$_3$BOK and HA-Ub-Lac failed. In order to troubleshoot and optimise the conditions required for stabilisation of the warheads, in depth analysis of the composition of the products at each production and ligation step would be necessary. This would require extensive LC/MS and NMR analysis. Due to time constraints the production of the TF$_3$BOK and Lac probes was therefore not continued and other methods were sought to try and confirm whether or not Ub conjugation enzymes are secreted by *T. spiralis*. 
Chapter 4: Results

Figure 4.9

A. HA-Ub-MESNa was ligated to the TF₃BOK (glycine 2,6-trifluoromethylbenzyloxy-methyl-ketone) warhead. B. HA-Ub-TF₃BOK probe was purified by ion exchange chromatography. The absorbance (mAU) of fractions eluted from the column by buffer B was measured as they were collected. C. The following fractions were analysed by SDS-PAGE for the presence of probe: the flow through X1, the wash A3, B10-B5 (green), B2 (blue), C3-C6 (yellow), C8-C10 (lilac), D9 (not shown on trace) and the original ligation reaction that was loaded onto the column (load). Fractions contained a protein at the expected size of the probe (b) and a possible aggregate protein of double the size (a). The marker (M) shows the molecular weights in kDa. D. Fraction B8 was tested for reactivity with 6His-TsUCH37. Reactions were analysed by SDS-PAGE and immuno-blot (IB) using anti-HA antibodies. As a control 6His-TsUCH37 was reacted with HA-Ub-VME probe. (a) A complex larger than the expected probe-TsUCH37 complex (possible aggregation complex or contaminant), (b) probe-TsUCH37 complex (c), possible probe aggregates, (d) expected size of probe. For each probe, 6His-TsUCH37 was also pre-incubated with the cysteine protease inhibitor NEM as a control.
Figure 4.10

A. HA-Ub-MESNa was ligated to the Lac (alpha-amino-beta-lactone) warhead. B. HA-Ub-TF$_3$BOK probe was purified by ion exchange chromatography. The absorbance (mA) of fractions eluted from the column by buffer B was measured as they were collected. C. The following fractions were analysed by SDS-PAGE for the presence of probe: the flow through X1 and X2, the wash A4, B10-B8 (green), B4 (pink), C1 (yellow), C6-C12 (lila), D9 (blue) and the original ligation reaction that was loaded into the column (load). Fractions contained a protein at the expected size of the probe (b) and a possible aggregate protein of double the size (a). The marker (M) shows the molecular weights in kDa. D. Fraction B8 was tested for reactivity with 6His-TsUCH37. Reactions were analysed by SDS-PAGE and immuno-blot (IB) using anti-HA antibodies. As a control 6His-TsUCH37 was reacted with HA-Ub-VME probe. (a) A complex larger than the expected probe-TsUCH37 complex (possible aggregation complex or contaminant), (b) probe-TsUCH37 complex (c), possible probe aggregates, (d) expected size of probe. For each probe, 6His-TsUCH37 was also pre-incubated with the cysteine protease inhibitor NEM as a control.
4.5. An *in vitro* ubiquitin conjugation assay of *T. spiralis* secreted proteins

An *in vitro* ubiquitin conjugation assay involves a mixture of all the components required for the enzymatic cascade that leads to the ubiquitination of a protein substrate. Either the substrate or the ubiquitin can be tagged (Figure 4.11 A). This allows detection of the tag by immuno-blot and visualisation of an upward shift in size of the substrate protein if ubiquitination has taken place. This may be observed as a smear if the protein has multiple ubiquitination sites and/or can be modified with polyubiquitin chains of various lengths. The E1 enzyme requires ATP to activate free ubiquitin, before passing the activated ubiquitin onto the E2 enzyme. Therefore the addition of ATP to the reaction buffer initiates the ubiquitin conjugation cascade.

Streptavidin is a protein from the bacterium *Streptomyces avidinii* that binds with a high affinity to biotin. Streptavidin-HRP conjugates can therefore be used in a similar manner to antibodies to blot for biotin-tagged proteins.

An *in vitro* ubiquitin conjugation reaction was carried out to determine if *T. spiralis* secreted proteins contain proteins with E1, E2 or E3 activity, using biotin-tagged human ubiquitin. All reactions were then separated by SDS-PAGE and analysed by binding to streptavidin-HRP and visualisation of the HRP using chemiluminescence. The positive control ubiquitin conjugation mixture contained the human recombinant E1 enzyme, the human recombinant E2 UBE2L3 (UbcH7), and the human recombinant E3 parkin. In the presence of the UbcH7, parkin is able to ubiquitinate itself (auto-ubiquitination). Unmodified parkin is 51 kDa. When all 3 enzymes were mixed together, with biotin-Ub and ATP, a smear was observed corresponding to the ubiquitinated forms of parkin (Figure 4.11 B, panel a). For a negative control, the E2 was removed from the reaction mixture and no smear was observed (Figure 4.11 B, panel b).

*T. spiralis* SP were then substituted into the reaction mixture for either the E1, the E2 or the E3 (Figure 4.11 B, panels c, d and e). A mixture containing *T. spiralis* SP and the E1 but no E2 or E3 was also set up, and a mixture containing only *T. spiralis* SP and biotin-Ub (Figure 4.11 B, panels f and g). Biotin-Ub conjugation was observed when *T. spiralis* SP was used instead of the E2 and the E3 (Figure 4.11 B, panels c and d). Conjugation was not observed when *T. spiralis* SP was used instead of the E1. This suggests that proteins in *T. spiralis* SP have Ub conjugation and ligation activities but not E1 activity. There was also no signal observed when *T. spiralis* SP were mixed with the E1 only, indicating that the human E2 is required for the *T. spiralis* secreted proteins to carry out E3 activity and the human E3 is required...
for *T. spiralis* secreted proteins to carry out E2 activity (Figure 4.11 B, panel f). In other words, *T. spiralis* SP have E2 and E3 activity, but it appears unlikely that these activities can work together without coming into contact with ubiquitin conjugation enzymes from an external source. It was therefore hypothesised that ubiquitination of substrates by these proteins requires mammalian host E2 or E3 partners. Efforts were then made to try and identify the proteins responsible for the E2 and E3 activity.
Figure 4.11

A. Schematic of the ubiquitin conjugation assay. Using ATP, the E1 enzyme activates biotin-tagged Ub and passes it on to the E2. The E2 enzyme passes Ub-biotin on to the E3, which catalyses the ligation of Ub-biotin to a substrate protein. For this assay the human E1, human E2 UBE2L3 and the human E3 parkin were used. HRP conjugated streptavidin can bind to the biotin-tag, thus allowing detection of ubiquitinated proteins by blot. B. An in vitro ubiquitination assay was carried out using the E1, E2, E3, and Ub-biotin (a) as a positive control. (b) The assay was carried out in the absence of the E2. (c) The secreted proteins of T. spiralis muscle larvae were assayed in the place of the E2 and the E3 (c). The assay was carried out with T. spiralis SP in the absence of the E1 (d), in the absence of the E2 and E3 (e) and in the absence of the E1, E2 and E3 (g). All reactions were separated by SDS-PAGE and analysed by streptavidin-HRP blot. Proteins sizes are indicated in kDa.
Identification of the proteins responsible for the ubiquitin conjugating (E2) and ligating (E3) activity in *T. spiralis* secreted proteins

4.6. Bioinformatics-based prediction of the *T. spiralis* secretome

Firstly, a bioinformatics approach was adopted to look at *T. spiralis* proteins and build a list of putative *T. spiralis* secreted proteins. Within this, searches for candidate *T. spiralis* ubiquitin-related enzymes were carried out. There are a multitude of bioinformatics programmes that have been developed to analyse eukaryotic proteins and predict secretion signals (signal peptides) or localisation motifs. Analysis of proteins by these programmes gives an overall score of the likelihood of secretion of the protein. All non-redundant (RefSeq) *T. spiralis* annotated protein sequences (a total of 16404 proteins) were analysed by 3 different sequence-based programmes: iPSORT, wolfPSORT and signalP. WolfPSORT was described earlier. SignalP is an algorithm that identifies signal peptide cleavage sites and therefore signal peptides for membrane localisation, translocation or secretion \[219\]. IPSORT predicts subcellular localisation sites by predicting N-terminal sorting signals. This includes identifying signals required for sorting proteins to the mitochondria or externally, i.e. for extracellular secretion \[289\].

All 16404 putative *T. spiralis* proteins were systematically searched against each of the 3 programmes. Derek Huntley at the Imperial College Bioinformatics Support Service carried out initial mass bioinformatics searches. Proteins that gave a positive score for one programme, were then searched against the other 2 programmes, until all possible combinations of positive score was obtained for each protein. Table 4.3 lists all combinations of a positive score for all 3 programmes out of all 16404 proteins. For example, 321 *T. spiralis* proteins were identified and listed as positive for iPSORT only, 776 proteins were positive for both iPSORT and wolfPSORT, and 750 proteins were positive for all 3 programmes: iPSORT, wolfPSORT and signalP. In total, 1948 out of a total 16404 proteins had a positive score (for an N-terminal sorting signal) when searched against iPSORT, 8824 had a positive ‘extracellular’ score according to wolfPSORT, and signalP predicted that 1060 proteins have a signal peptide.
Table 4.3 Bioinformatics-based prediction of *T. spiralis* secreted proteins

<table>
<thead>
<tr>
<th></th>
<th>iPSORT</th>
<th>wolfPSORT</th>
<th>signalP</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPSORT</td>
<td>321</td>
<td>776</td>
<td>101</td>
</tr>
<tr>
<td>wolfPSORT</td>
<td>7134</td>
<td>164</td>
<td></td>
</tr>
<tr>
<td>signalP</td>
<td></td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>total</td>
<td>1948</td>
<td>8824</td>
<td>1060</td>
</tr>
</tbody>
</table>

* number of sequences with a positive score for all 3 programmes

Table 4.3 Table of *T. spiralis* proteins predicted to be secreted

All non-redundant *T. spiralis* annotated protein sequences were analysed by 3 different prediction programmes that predict signal peptides (iPSORT and signalP) or sub-cellular localisation motifs (wolfPSORT). The number of sequences that were predicted as positive for each programme alone, and in combination with each of the other programmes, is listed. For example, 7134 sequences were predicted to be secreted by wolfPSORT only (but not by signalP and iPSORT, and 776 different sequences were predicted to be secreted by both wolfPSORT and iPSORT (but not signalP) etc.
Each list of proteins with a positive score (for all combinations of searches) was then further analysed by keyword searches, BLAST for orthology and by SMART for conserved domain architecture [210,211]. This identified 130 proteins that are related to the ubiquitin-proteasome system or contain domains that can associate with Ub and that iPSORT wolfPSORT and/or signalP predict to be secreted (all results are found in appendix 2). To refine the list, all proteins whose wolfPSORT score of ‘extracellular’ was listed as the most likely location (highest score) were identified (Table 4.4). This was because a vast number of proteins were predicted to have signal peptides by iPSORT and signalP that may not have been signal peptides for secretion, rather signal peptides for membrane localisation. Using WolfPSORT in this way (taking the highest scoring location only) was considered the most stringent prediction algorithm for *T. spiralis* proteins.

4.7. Further analysis of 2 putative *T. spiralis* E3 proteins predicted to be secreted by bioinformatics

Searches were carried out to find commercially available antibodies to the mammalian orthologues of the Ub-related proteins whose wolfPSORT score of ‘extracellular’ was the highest (listed in Table 4.4). Alignments of the mammalian orthologues with the putative *T. spiralis* sequences identified 2 antibodies that were predicted to cross react (Table 4.5). Two antibodies were purchased that were predicted to cross react with:

1. A putative *T. spiralis* HECT-domain protein, GI: 339253068, orthologous to the human Ub E3-ligase, NEDD4-L (neural precursor cell expressed, developmentally down-regulated 4-like) [290].

2. A putative WD domain, G-beta repeat-containing domain protein orthologous to human WDR48 [291].

HECT and domains both have Ub-ligase activity [292]. WD-repeat domains consist of short amino acid (approximately 40) repeats and are often found in E3 ligase enzymes [293].

Anti-NEDD4-L and anti-WDR48 were reacted by immuno-blot with 10 ug of *T. spiralis* SP, *T. spiralis* muscle larvae lysate and with human cell lysate as a positive control (HEK 293T cell lysate). The predicted sizes of each orthologue compared to the human protein can be seen in Table 4.5. No reactivity with the anti-WD48 antibody was observed in the *T. spiralis* SP at the predicted size of 115.3 (Figure 4.12 B). Both
antibodies showed reactivity with proteins of the predicted size in HEK lysate, although additional reactivity was observed at various sizes with the anti-WDR48 antibodies. Significant reactivity with the *T. spiralis* SP was observed for the anti-NEDD4-L antibodies (Figure 4.12 A). The predicted size of the putative *T. spiralis* NEDD4-L is 96 kDa. In *T. spiralis* SP, a strong band was observed between 55 and 70 kDa. Reactivity at the same size was observed in the *T. spiralis* lysate, although an additional larger band (just below 100 kDa) was also observed. In HEK lysate, a band was observed at an expected size of human NEDD4-L of approximately 100 kDa, however 2 additional bands were also observed, one corresponding to the reactive band in *T. spiralis* secreted proteins. Considering that the reactivity in *T. spiralis* secreted proteins was so significant, and that thus far many potential mis-annotations in the *T. spiralis* draft genome had been experienced, the source of the reactivity with the anti-NEDD4-L antibody was further investigated. This was done using a method of protein fractionation called ‘salting out’.

The solubility of a protein is dependent on the salt concentration of its solvent. Once the salt concentration has reached the critical concentration for a particular protein with a particular hydrophobicity profile, the protein will precipitate by 'salting out' of the solution. Proteins in a mixture can therefore be separated (or fractionated) by incrementally increasing the salt concentration of the solution step by step. The precipitate is then recovered at each step. This method was used to fractionate *T. spiralis* secreted proteins using a saturated ammonium sulphate solution. Precipitated secreted protein fractions were re-solubilised in aqueous buffer, and separated by SDS-PAGE. Proteins were analysed by immuno-blot using the anti-NEDD4-L antibody described above. This method was repeated a number of times to narrow down the range of ammonium sulphate salt concentration required to precipitate the anti-NEDD4-L-reactive proteins (data not shown). Once optimised, 500 µg of *T. spiralis* secreted proteins was fractionated at the appropriate concentrations of ammonium sulphate. A small portion of each of these fractions was analysed by immuno-blot using anti-NEDD4-L (Figure 4.12 C). The secreted proteins alone (20 µg) were analysed in parallel as a positive control, showing now 3 reactive bands between 50 and 70 kDa. Of the three reactive bands, the largest and smallest were found to precipitate out of a solution of 34% ammonium sulphate, and the smallest again at 44%. The middle band was found to precipitate in every fraction between 64% and 80% ammonium sulphate (Figure 4.12 C).

*T. spiralis* secreted proteins were then precipitated using 34%, 44% and 69% ammonium sulphate. Precipitated fractions were separated by SDS-PAGE and 5%
was analysed by immuno-blot using anti-NEDD4-L, and 95% was visualised by silver staining (Figure 4.12 D). The immuno-blot was used as a guide to manually excise 6 protein bands from the silver-stain gel that contained the anti-NEDD4-L-reactive proteins. These proteins were then analysed by LC/MS/MS and peptide data was searched against all annotated *T. spiralis* proteins using SEQUEST. Eric Spooner at the Whitehead Institute, Boston, carried out LC/MS/MS and SEQUEST analyses. No peptides from any of the 6 bands, matched the putative *T. spiralis* HECT domain protein GI: 339253068, the NEDD4-L orthologous sequence (appendix 3). Although putative *T. spiralis* protein matches were made, none of these protein sequences contained HECT domains when further analysed by BLAST and SMART.

WolfPSORT predicted many *T. spiralis* ubiquitin-proteasome components to be secreted, including the putative ubiquitin ligase TsNEDD4L. Because it was not possible to experimentally verify the secretion of TsNEDD4L using the fractionation and immuno-blot technique, this method was not continued.
Table 4.4 Table of ubiquitin-related *T. spiralis* proteins predicted to be secreted with wolfPSORT ‘extracellular’ score

<table>
<thead>
<tr>
<th>Programme</th>
<th>GI accession #</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>iPSORT, wolfPSORT and signalP</td>
<td>339238271</td>
<td>E3 ubiquitin-protein ligase RNF34</td>
</tr>
<tr>
<td></td>
<td>339245151</td>
<td>RWD domain protein</td>
</tr>
<tr>
<td>iPSORT and wolfPSORT</td>
<td>339235861</td>
<td>E3 ubiquitin-protein ligase synoviolin-A</td>
</tr>
<tr>
<td></td>
<td>339262132</td>
<td>E3 ubiquitin-protein ligase sia-1</td>
</tr>
<tr>
<td></td>
<td>339233898</td>
<td>ubiquitin-conjugating enzyme E2</td>
</tr>
<tr>
<td></td>
<td>339264723</td>
<td>RING-H2 finger protein ATL1D</td>
</tr>
<tr>
<td></td>
<td>339236113</td>
<td>polycomb group RING finger protein 1</td>
</tr>
<tr>
<td></td>
<td>339264545</td>
<td>bromodomain and WD repeat-containing protein 2</td>
</tr>
<tr>
<td></td>
<td>339238813</td>
<td><strong>WD domain, G-beta repeat-containing domain protein (WD48)</strong></td>
</tr>
<tr>
<td></td>
<td>339242135</td>
<td>bromodomain and WD repeat-containing protein 2</td>
</tr>
<tr>
<td></td>
<td>339253068</td>
<td><strong>putative HECT-domain protein (NEDD4-L)</strong></td>
</tr>
<tr>
<td>wolfPSORT and signalP</td>
<td>339235867</td>
<td>E3 ubiquitin-protein ligase synoviolin-A</td>
</tr>
<tr>
<td></td>
<td>339263624</td>
<td><strong>putative ubiquitin--protein ligase (UBE2N)</strong></td>
</tr>
</tbody>
</table>

Table 4.4 Ubiquitin-related *T. spiralis* proteins predicted to be secreted by wolfPSORT

Table shows the ubiquitin-related proteins (containing conserved ubiquitin interaction domains) that were predicted by wolfPSORT to be most likely located extracellularly. Commercially available antibodies were found for testing 2 of these *T. spiralis* proteins for secretion (highlighted in yellow).
Table 4.5 Table of antibodies for testing *T. spiralis* SP predicted to be secreted

<table>
<thead>
<tr>
<th>Protein</th>
<th>Antibody</th>
<th>MW of human (kDa)</th>
<th>Predicted MW of <em>T. spiralis</em> (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEDD4-L</td>
<td>anti-HsNEDD4L Protein Tech group: 13690-1-AP</td>
<td>100.3</td>
<td>96</td>
</tr>
<tr>
<td>WDR48</td>
<td>anti-HsWDR48 Sigma: HPA038421</td>
<td>76.2</td>
<td>115.3</td>
</tr>
</tbody>
</table>

Table 4.5 Antibodies to ubiquitin-related proteins predicted to be secreted by *T. spiralis*

Commercially available antibodies that are likely to cross-react with the putative *T. spiralis* E3, NEDD4-L and the putative *T. spiralis* WDR48 protein.
Figure 4.2 Analysis of the secretion of NEDD4-L and WDR48 by *T. spiralis*

The secreted proteins (SP) of *T. spiralis* muscle larvae, *T. spiralis* lysate, and human (HEK) cell lysate were analysed by immuno-blots (IB) for reactivity with anti-NEDD4-L antibodies (A) and anti-WDR48 antibodies (B). (A) The predicted size of the *T. spiralis* protein, (b) the expected size of the human protein and (c) the observed size of the signal in *T. spiralis* SP. C. Gradient ammonium sulphate (NH)$\textsubscript{4}$SO$_{4}$ precipitation (fractionation) of the secreted proteins of *T. spiralis* muscle larvae. Fractions were separated by SDS-PAGE and analysed by immuno-blot using anti-NEDD4-L antibodies. Three reactive bands were observed (1, 2 and 3). The fractionation by precipitation using 34%, 44% and 69% (NH)$_{4}$SO$_{4}$ (circled in red) was scaled up. D. Of each fraction, 5% was analysed by immuno-blot using anti-NEDD4-L antibodies, and 95% was analysed by silver staining. The anti-NEDD4-L signal was lined up with proteins of the silver-stained gel. These 6 areas were manually excised (red boxes) for analysis by tandem mass spectrometry (LC/MS/MS) by Eric Spooner (Whitehead Institute).
4.8. Proteomic analyses of all *T. spiralis* secreted proteins by tandem mass spectrometry

Thus far, a number of different methods were used to try and identify the proteins responsible for the E2 and E3 activity detected in *T. spiralis* secreted proteins by a ubiquitin conjugation assay. None of these methods however proved successful. It was therefore decided that the total secreted proteins would be analysed by tandem mass spectrometry (LC/MS/MS), and the data would be searched for conserved E2 and E3 domains.

*T. spiralis* muscle larvae were isolated from infected skeletal muscle tissues. Larvae were washed multiple times to remove host muscle tissue. When *T. spiralis* muscle larvae die, they become uncoiled and float in culture media. Larvae that are alive readily sediment. It is therefore possible to remove dead parasites from a mixture. Dead larvae were removed from live larvae to prevent non-secreted proteins being released from dead or dying parasites. To ensure a low level of parasite death, larvae were cultured for 24 hours only, before the culture supernatant was collected. Day 1 secreted proteins were purified from the culture supernatant. A large proportion of *T. spiralis* secreted proteins are modified with sugar moieties (glycosylations). The dominant moiety on *T. spiralis* secreted proteins is called tyvelose [294-296]. Muscle larvae SP (50 ug) were therefore subjected to treatment with the deglycosylating enzyme PNGaseF to remove all N-linked glycosylations from glycoproteins to obtain more clarity of the actual size of a protein by SDS-PAGE. Deglycosylated SP were separated by SDS-PAGE, alongside 10 ug of untreated, glycosylated proteins, under conditions compatible for LC/MS/MS (Figure 4.13 A). The entire lane of proteins was manually excised from the gel and dissected as annotated in Figure 4.13 A. Each mid-section (numbered) portion was analysed by highly sensitive LC/MS/MS. Mass spectrometry was carried out by Steve Gygi and Michael Weekes at the Harvard Medical School, Boston. Data was then searched against the *T. spiralis* UniProt proteome, its reverse complement and common contaminants of SDS-PAGE and mass spectrometry. Matches were made to all of the *T. spiralis* putative secreted proteins identified during a previous study of the *T. spiralis* secretome by Robinson et al. however many additional new matches were also made [196]. Because the larvae were isolated from rat muscle tissue, peptides were also searched against the rat proteome, and only 3 matches to rat proteins were made. Data was consistent with previous *T. spiralis* secreted protein data found in the literature. For example, the most abundantly secreted protein according to this study was analysed by BLAST. It
was found to match putative *T. spiralis* nudix hydrolase enzymes, and so may be related to the 5' nucleotidase (5'NT) enzyme previously characterised by Gounaris *et al.* to be an abundant secreted protein [197].

Matches were made to 3 ubiquitin-related proteins, shown in Table 4.6. The first match, the ‘Ubiquitin family member protein (fragment)’ is the putative *T. spiralis* Ub protein orthologue, which is labelled as a fragment because the annotated sequence contains a single (mono) Ub followed immediately by a fragment of another Ub whose sequence is incomplete (Uniprot: E5RYP1, GI: 339233028). In many organisms, tandem Ub genes exist, where Ub is transcribed in tandem and later hydrolysed by DUBs to form mono-ubiquitin [297-299]. This fragment sequence may therefore represent an incomplete version of the *T. spiralis* tandem Ub coding sequence, since the annotation was generated from contig scaffolds of the *T. spiralis* genome.

Peptides matching this sequence were found in all segments of the gel except number 3 (Figure 4.13 A). This reinforces the evidence presented in Figure 4.6 that shows how *T. spiralis* proteins are secreted with ubiquitin attached.

The second match was orthologous to the RING finger domain protein (E3) RNF13 (Uniprot: E5S4Y0, GI: 339237319) and has a predicted size of 58 kDa. The human orthologue of RNF13 is a protein that is found in the membranes of the ER, the golgi, the nucleus and lysosomes. Human RNF13 is a RING finger protein that has Ub ligase (E3) activity and has been linked to cancer development of pancreatic cells [300]. When analysed by BLAST, it was found that the domain architecture of the *Trichinella spiralis* putative open reading frame was slightly different to the consensus structure of the orthologues. The *T. spiralis* sequence contains both an N-terminal extension and a C-terminal extension that are not present in the orthologous sequences (Figure 4.13 C). The extended *Trichinella spiralis* N-terminal region contains an extra domain that is not present in the orthologues. This is an MIR domain, which stands for Mannosyltransferase Inositol 1,4,5-trisphosphate receptor (IP3R) and Ryanodine receptor (RyR). There are 3 types of protein in which the MIR domain has been found and there are therefore multiple different functions associated with these domains [301,302]. Peptides matching the putative TsRNF13 sequence were found in segment 2 of the gel, corresponding to a molecular weight of 20-30 kDa (Figure 4.13 A).
Chapter 4: Results

Figure 4.13

A. 50 ug of day-1 secreted proteins (SP) of *T. spiralis* muscle larvae were treated with the deglycosylating enzyme PNGaseF to remove glycosylations, 10 ug were not deglycosylated. These samples were separated by SDS-PAGE and the areas (boxed in blue and numbered) were manually excised and analysed by tandem mass spectrometry (LC/MS/MS) by Michael Weekes and Steve Gygi (Harvard Medical School).

B. A separate experiment was carried out where 60 ug of day-1 SP (not treated with PNGaseF) was separated by SDS-PAGE. The areas boxed in blue and numbered were analysed by LC/MS/MS. Two ubiquitin-related enzymes were identified in the bands indicated (red boxes) in both gels, the putative *T. spiralis* RNF13 (E3) and the putative *T. spiralis* UBE2L3 (E2).

C. Schematic of the structure of the putative TsRNF13 and (D) the putative TsUBE2L3 showing the conserved domains and their locations within each sequence.

Figure 4.13 Proteomic analysis of *T. spiralis* secreted proteins

A. 50 ug of day-1 secreted proteins (SP) of *T. spiralis* muscle larvae were treated with the deglycosylating enzyme PNGaseF to remove glycosylations, 10 ug were not deglycosylated. These samples were separated by SDS-PAGE and the areas (boxed in blue and numbered) were manually excised and analysed by tandem mass spectrometry (LC/MS/MS) by Michael Weekes and Steve Gygi (Harvard Medical School). B. A separate experiment was carried out where 60 ug of day-1 SP (not treated with PNGaseF) was separated by SDS-PAGE. The areas boxed in blue and numbered were analysed by LC/MS/MS. Two ubiquitin-related enzymes were identified in the bands indicated (red boxes) in both gels, the putative *T. spiralis* RNF13 (E3) and the putative *T. spiralis* UBE2L3 (E2). C. Schematic of the structure of the putative TsRNF13 and (D) the putative TsUBE2L3 showing the conserved domains and their locations within each sequence.
Table 4.6

<table>
<thead>
<tr>
<th>Uniprot</th>
<th>Protein name</th>
<th>Gene name</th>
<th>Unique peptides</th>
<th>Protein length (aa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E5RYP1</td>
<td>Ubiquitin family protein (Fragment)</td>
<td>Tsp_07290</td>
<td>2</td>
<td>210</td>
</tr>
<tr>
<td>E5S4Y0</td>
<td>RING finger protein 13</td>
<td>Tsp_05047</td>
<td>5</td>
<td>525</td>
</tr>
<tr>
<td>E5S8T6</td>
<td>Ubiquitin-conjugating enzyme E2 L3 (Fragment)</td>
<td>Tsp_00154</td>
<td>6</td>
<td>145</td>
</tr>
</tbody>
</table>

Table 4.6 Ubiquitin-related proteins identified by the proteomic analysis of *T. spiralis* secreted proteins

Table shows 3 ubiquitin-related proteins that matched peptides identified by LC/MS/MS in each separate analysis of the day-1 secreted proteins of *T. spiralis* muscle larvae. Unique peptides did not match any other *T. spiralis* protein.
When analysed by localisation prediction programmes, the following results for the putative TsRNF13 were:

- signalP: yes - signal peptide
- iPSORT: yes - signal peptide
- wolfPSORT: plas: 20.0, extr_plas: 15.5, extr: 9.0

WolfPSORT analysis indicates that the plasma membrane is the most likely location for the protein, but that it also contains motifs orthologous to proteins that are secreted. SignalP and iPSORT also predict a signal peptide. This may be a signal for translocation to a plasma membrane or for secretion. SMART also predicts that the protein contains a mid-sequence single transmembrane domain (Figure 4.13 C).

Another protein domain prediction programme, InterPro, predicts that the N-terminal half (containing the MIR domain) may be non-cytoplasmic, whereas the C-terminal half (containing the RING domain) may be cytoplasmic.

The final ubiquitin-related match was a putative protein annotated as a 'Ubiquitin-conjugating enzyme E2 L3 (Fragment)' that has a predicted size of 17 kDa. This sequence is an incomplete open reading frame for a UBE2L3 orthologue (Uniprot E5S8T6, GI:339240047). UBE2L3, also known as UbcH7 or E2-F1, is a UBCc domain ubiquitin conjugating enzyme (E2, [65,303]). The annotated fragment open reading frame for TsUBE2L3 is 145 amino acids long and lacks a start codon. The UBCc domain of the putative TsUBE2L3 spans the fragment from the first amino acid to the 140th amino acid (Figure 4.13 D). Peptides matching the putative TsUBE2L3 were also found in segment 2 of the gel (Figure 4.13 A).

When analysed by localisation prediction programmes, the following results for the putative TsUBE2L3 were:

- signalP: no signal peptide
- iPSORT: no signal peptide
- wolfPSORT: cyto: 23.0, pero: 3.0, extr: 2.0, nucl: 2.0

To validate these findings, the experiment was repeated using an even more sensitive method of LC/MS/MS (again, carried out by Steve Gygi and Michael Weekes). Day 1 SP was collected and concentrated as described for the first experiment. This time however, proteins were not treated with PNGaseF. This was to minimise protein
degradation that can occur during the long incubation at 37°C with PNGaseF. The secreted proteins were separated by SDS-PAGE, but in this experiment the running time was reduced, to ensure that even the smallest proteins were retained in the gel. Efforts to prevent contamination of the gel were also increased. Figure 4.13 B shows how the gel was divided up for LC/MS/MS. Again, peptides were identified and searched against the *T. spiralis* protein database using SEQUEST. In total, a smaller number of peptides were identified than in the first experiment. Peptides from both the first and second experiment were then combined and assembled into proteins. A good level of agreement in protein matches was observed (all data can be found in the supplementary appendix 4). A total of 333 *T. spiralis* proteins were identified, of which 5 are likely to be false positive results or proteins released from dead and dying parasites. It is likely that the false positive results are proteins from which only 1 peptide was identified. Less than 1% of the protein matches were rat proteins, demonstrating a clean isolation of the muscle larvae from host muscle tissue.

Both the putative TsUBE2L3 and the TsRNF13 were identified again in the second experiment. More peptides matching the TsUBE2L3 were identified in the second experiment than the first. Less peptides matching the TsRNF13 were identified in the second experiment than the first. Each protein matched unique peptides, i.e. peptides that did not match any other protein. Peptides matching the TsUBE2L3 were all found in segment 2 of the second gel, and peptides matching the TsRNF13 were found in segment 3 of the second gel. Peptides that matched other proteins as well (to a lesser extent) are referred to as razor peptides. In total from both sets of data, 6 unique peptides and no razor peptides matched the TsUBE2L3, and 5 unique peptides and no razor peptides matched TsRNF13 (Table 4.6). The peptide coverage of the 145 amino acid sequence of TsUBE2L3 was 22.8%. The coverage of the 525 amino acid sequence of TsRNF13 was only 2.5% (Figure 4.13 C and D). When cross-referenced with the bioinformatics analyses, both proteins were predicted by wolfPSORT to be located extracellularly.

All *T. spiralis* sequence matches labelled as 'uncharacterised' were further analysed using BLAST and SMART, and no other ubiquitin-proteasome-related proteins were identified. At this point, the putative TsUBE2L3 and the putative TsRNF13 were therefore the only proteins identified by LC/MS/MS that could be attributed to the E2 and E3 activity observed in *T. spiralis* secreted proteins. The purpose of experiments presented in chapter 5 was to then validate the secretion of these proteins and characterise their function.
Summary

4.1. The Ub-based inhibitor probe, HA-Ub-VME was produced and purified and its reactivity with DUB enzymes was verified using 6His-TsUCH37.

4.2. Although no deNeddylating activity was detected in *T. spiralis* secreted proteins using FLAG-Nedd8-VS, potential cysteine-dependent activity could be detected using HA-Ub-VME. When further analysed using HA-Ub-VME followed by LC/MS/MS, no DUB activity could be detected.

4.3. The lack of deubiquitinating and deNeddylating activity was verified using AMC conjugates of Ub and Nedd8.

4.4. The production of Ub-based inhibitor probes designed to have a higher affinity for ubiquitin conjugation enzymes (E1, E2 and E3), was unsuccessful.

4.5. *T. spiralis* secreted proteins demonstrated E2 and E3 activity as measured by an *in vitro* ubiquitin conjugation assay.

4.6. Bioinformatics analysis of *T. spiralis* annotated proteins identified 130 putative ubiquitin-proteasome proteins that were predicted to have an N-terminal signal peptide or be located extracellularly.

4.7. *T. spiralis* secreted proteins reacted with an antibody raised against a human orthologue of a *T. spiralis* E3 protein predicted to be located extracellularly, NEDD4-L. LC/MS/MS however could not confirm the presence of this protein in the secreted proteins.

4.8. Candidates for E2 and E3 activity in the secreted proteins were identified by LC/MS/MS, a putative *T. spiralis* UBE2L3 and a putative *T. spiralis* RNF13.
Discussion

Experiments in chapter 4 set out to determine whether or not the secreted proteins of *T. spiralis* muscle larvae possess activities of ubiquitin-proteasome system enzymes. The investigation began by looking for DUB and deNeddyrating activity in *T. spiralis* secreted proteins and this could not be identified. Experiments were then designed to look for ubiquitin conjugation activity.

Initially, production of the Ub-based inhibitor probe, HA-Ub-VME was carried out. Expression of the precursor of the inhibitor probe HA-Ub-intein/CBD, and purification of HA-Ub-MESNa was successful, however the final yield of purified active HA-Ub-VME was low. It is possible that some probe may have been lost during ion exchange chromatography due to inefficient binding to the column. Although purified active probe was detected, some of the fractions collected after ion exchange chromatography contained probe of an expected size that was not fully active. The VME warhead that eventually forms the covalent bond with the ubiquitin enzyme (DUB) contains an ester bond; a chemical group that is sensitive to both acid and base catalysed hydrolysis (Figure 4.14 A). Hydrolysis of the VME ester bond could occur either before or after ligation with the protein, and in both cases would render the probe less active and less able to react with DUBs. In addition, the HA-Ub-MESNa could degrade thus preventing ligation, and also reducing the final yield of active probe. For future optimisation of yield and activity, buffer conditions during ion exchange chromatography and ligation conditions would have to be further investigated. Because active inhibitor probes have an extremely high affinity for DUB enzymes, a sufficient volume of active HA-Ub-VME was collected for my experiments and these issues were not investigated here.
Figure 4.14 Inhibitor probe warhead degradation

A  VME

\[
\begin{array}{c}
\text{VME} \\
\text{Cysteine} \\
\text{Hydrolysis leads to reduced function}
\end{array}
\]

B  TF₃BOK

\[
\begin{array}{c}
\text{TF₃BOK} \\
\text{Cysteine} \\
\text{Hydrolysis leads to loss of function}
\end{array}
\]

B  Lac

\[
\begin{array}{c}
\text{Lac} \\
\text{Hydrolysis leads to loss of function}
\end{array}
\]

Figure 4.14 The degradation of warheads used to make Ub-based inhibitor probes

All probe warheads have ester bonds that can hydrolyse under acidic or basic conditions, either before or after ligation with the HA-Ub-MESNa. Each warhead reacts differently with the sulfhydryl group (SH) of the cysteine of the ubiquitin-specific enzyme. These reactions are not necessarily through reaction with the ester bond of the warhead. The diagram shows the mechanism of reactivity with a cysteine and the mechanism of the possible hydrolysis, which for VME (A) renders the warhead less reactive and for (B) TF₃BOK and (C) Lac, renders the warheads completely inactive. Curly arrows indicate the movement of an electron pair. Image modified from drawings supplied by Jennifer Ward.
Although no deNeddylating activity was observed using FLAG-Nedd8-VS, Using HA-Ub-VME it was not clear if DUB activity was present. By immuno-blot there were some suggestions of probe-protein complexes forming. For this reason, the HA-Ub-VME reaction with the secreted proteins was scaled-up for a more in-depth analysis. Despite 2 biological repeats, LC/MS/MS could not detect any putative *T. spiralis* DUBs in the proteins excised from the gel. Areas of the gel to be excised were selected for analysis based on the visual appearance of a band in the NEM-negative sample only. Being subjective, this process may not have been exhaustive, and DUBs may have been missed. Furthermore, a large number of *T. spiralis* secreted proteins were immuno-precipitated from both the NEM-negative and -positive samples. These proteins may have been non-specific contaminants, proteins able to specifically react with the probe in a cysteine-independent manner, or probe aggregates. It was possible that their LC/MS/MS signal masked signals from any DUB-probe complexes present.

Interestingly, LC/MS/MS did identify 3 ubiquitin-related proteins. One of these was ubiquitin itself, the source of which could have been from the probe or from ubiquitinated *T. spiralis* secreted proteins. The second was the putative *T. spiralis* E2 UBE2N, however peptides matching this protein were only identified after one of 2 experiments. They were found in a region of the coomassie-stained gel that did not correspond with the size of TsUBE2N in complex with the probe (which would be approximately 20 + the 11 probe (31) kDa if the annotation is correct). TsUBE2N was not predicted to be secreted by bioinformatics analysis, and although the anti-human UBE2N that was tested reacted with *T. spiralis* proteins in lysate of muscle larvae, it did not react with *T. spiralis* secreted proteins. Taken together, results suggest that it is unlikely that TsUBE2N is truly secreted by *T. spiralis* muscle larvae. It is possible that the protein was present as a contaminant released from dead or dying parasites, as the secreted proteins were collected after *T. spiralis* larvae had been cultured for 4 days by which time a significant reduction in viability is observed.

The final ubiquitin-related match, an RWD domain protein characterised throughout the literature as the 53 kDa excretory-secretory *T. spiralis* antigen (GI: 805126) was also dubious. The 53 kDa excretory-secretory *T. spiralis* antigen does not contain an RWD domain, which is the only ubiquitin-related domain in the sequence. Without the RWD domain, it is not clear how it would interact with the Ub of the HA-Ub-VME probe. It is therefore unlikely that this protein is transcribed and secreted as annotated, or that the annotation is incorrect. To confirm this it would be necessary to carry out RACE-PCR and find the true open reading frame of this gene.
Thus far, no conclusion could be reached on whether or not *T. spiralis* muscle larvae secrete deubiquitinating or deNeddylating enzymes. In the case of Nedd8 the lack of activity with the probe may have been due to differences between the human and *T. spiralis* Nedd8 sequences, and the failure of *T. spiralis* deNeddylating enzymes to recognise the human Nedd8-based probe. To clarify these points, AMC assays were carried out. Although lysate of *T. spiralis* muscle larvae was able to hydrolyse Ub-AMC and Nedd8-AMC, confirming an ability to interact with the human orthologues, *T. spiralis* secreted proteins were not able to hydrolyse either the Ub or the Nedd8. It was therefore concluded that, using the available reagents, no DUB or deNeddylating activity could be detected in the *T. spiralis* secreted proteins. It is therefore unlikely that these enzymes are secreted by the parasite.

Because pathogenic viruses and bacteria express ubiquitin conjugation enzymes (as well as DUBs), I next wanted to look for E1, E2 and E3 enzymes in *T. spiralis* secreted proteins. These enzymes ligate more readily with Ub-based inhibitor probes that contain more electrophilic warheads [277]. This is possibly due to the increased nucleophilicity of the catalytic cysteine of the enzymes, compared to DUB and deNeddylating enzymes. Of the warheads previously reported, TF$_3$BOK and Lac exhibited reactivity with a diverse list of E1, E2 and E3 enzymes from lysates of mouse lymphoma (EL4) and human mammary epithelial cell (HMLE) cells [277]. These were therefore chosen for testing the *T. spiralis* secreted proteins for E1, E2 and E3 activity.

Production of the HA-Ub probe, with both the TF$_3$BOK or Lac warheads was problematic. When synthesised, the warheads are in a slightly acidic state, in which they are more stable. When transferred to the conditions required for ligation (pH 8), warhead degradation was significant. This may have been due to hydrolysis of the ester bond of the warhead that can undergo acid or base catalysed hydrolysis as described earlier for HA-Ub-VME. In these cases, hydrolysis of the warhead at pH 8 may be happening more rapidly than the ligation reaction with HA-Ub-MESNa (Figure 4.14). It is also possible that the ligation event occurred, and the warhead underwent hydrolysis later, leaving degraded HA-Ub-TF$_3$BOK, or HA-Ub-Lac. When VME is degraded its activity is reduced but when TF$_3$BOK and Lac are degraded the warheads are rendered completely inactive and can no longer react with target proteins. This made their production more problematic than the HA-Ub-VME probe.

Although Love *et al.* detected E1, E2 and E3 enzymes using these probes, they were not originally designed to specifically capture E1, E2 and E3 enzymes. Ub-based
inhibitor probes were designed so that the active-site cysteine of a DUB enzyme reacts with the C-terminal end of the probe, thus reacting with the warhead and becoming trapped. Conversely, there is some evidence that during ubiquitin conjugation, the active-site cysteine of E1, E2 and E3 enzymes doesn’t react with the C-terminal residues of Ub, rather with mid-sequence residues of Ub [304,305]. This would mean that although ubiquitin conjugation enzymes are able to recognise the Ub component of the probe, the cysteine residue of the enzyme that then becomes covalently bound to the probe is not the active-site cysteine. This may explain the reduced affinity of Ub-based probes with ubiquitin conjugation machinery. Considering this, and considering the problems experience with the production of these probes, this method was not continued.

The next experiment, an in vitro ubiquitin conjugation assay, confirmed that *T. spiralis* secreted proteins were not able to ubiquitinate substrate proteins when supplemented with the human E1 alone. Interestingly however, ubiquitin conjugation activity was observed when *T. spiralis* secreted proteins were supplemented with either the human E2, UBE2L3 or the human E3, parkin. This confirmed that *T. spiralis* proteins can interact with the human E1 which shares 45.7% of residues with the *T. spiralis* putative E1 (GI: 339234521). Although in the presence of UBE2L3, parkin can undergo auto-ubiquitination, the *T. spiralis* secreted proteins also contain hundreds of theoretical ubiquitination substrates [306,307]. And since the signal being measured was the biotin labelled ubiquitin, the source of the observed signal may not necessarily be from auto-ubiquitination of the parkin. Where the secreted proteins were supplemented with the human E2 (UBE2L3) and the E1 only (no parkin E3), substrates will always be the *T. spiralis* secreted proteins themselves, unless the E2 can also auto-ubiquitinate in the presence of *T. spiralis* secreted proteins. Where the secreted proteins were supplemented with the human parkin and the E1 only (no UBE2L3), the signal may represent the auto-ubiquitination of the parkin or the ubiquitination of the *T. spiralis* secreted proteins themselves. In conclusion, *T. spiralis* proteins have E2 activity and E3 activity, and are not able to ubiquitinate amongst themselves in vitro, but are able to collaborate with the human proteins UBE2L3 and parkin. This has very interesting implications, considering that the secreted proteins come into contact with host proteins during infection. I hypothesise that these enzymes are specifically secreted to only ubiquitinate when in contact with the host partner enzyme, and to therefore directly interact with a host ubiquitin-proteasome system.
The purpose of the follow-up investigation was to confirm the identity of the secreted proteins that possess E2 and E3 activities. To try and predict this, the bioinformatics programmes iPSORT, wolfPSORT and signalP were used to analyse all annotated *T. spiralis* proteins. In total, 9291 putative *T. spiralis* proteins were predicted, by one or multiple programmes, to have an N-terminal signal peptide or be located extracellularly (the sum of all prediction combinations listed in Table 4.3). From previous reports and experimental experience, there are possibly a few hundred proteins secreted by *T. spiralis* muscle larvae [196,308-311]. I propose 2 main reasons for this vast discrepancy:

1. *T. spiralis* secrete proteins at all stages of their life cycle, and these proteins will be differentially expressed [115]. Undoubtedly the secreted proteins required for each stage will vary, and so for each few hundred proteins secreted at each stage, combined, a few thousand may be secreted throughout the entire life cycle.

2. N-terminal signal peptides do not only target proteins for secretion, but also target transmembrane proteins to their membrane destination. They may also serve as signals that target proteins to intracellular organelles, such as mitochondria, peroxisomes, lysosomes and the nucleus. Many *T. spiralis* proteins will fall into these categories, giving a false positive iPSORT and/or signalP score. Conversely, some *T. spiralis* proteins are secreted without a signal peptide, via an unconventional system of secretion that is not well-understood [312]. These proteins would therefore give a false negative iPSORT and/or signalP score.

Proteins predicted to be secreted were therefore analysed based on their wolfPSORT score only, and not on their signalP and iPSORT score. These proteins were then analysed to isolate only those whose function was annotated as being involved in the ubiquitin pathway. 130 sequences were identified. These were further analysed to isolate 13 ubiquitin-related sequences whose wolfPSORT score of ‘extracellular’ was the highest. Commercial antibodies to 2 of these 13 proteins looked likely to cross-react with the *T. spiralis* sequence and so *T. spiralis* secreted proteins was tested for reactivity. Only the anti-NEDD4 antibody reacted with *T. spiralis* secreted proteins. However, LC/MS/MS analysis of the reactive bands from *T. spiralis* secreted proteins, did not return any NEDD4-L matches. It is possible that the signal observed represented proteins that are able to non-specifically bind to the antibody, and this would explain the size discrepancy. Alternatively, it is possible that the NEDD4-L protein was present, but masked by more abundant secreted proteins of a similar size, thus failing to be detected by LC/MS/MS. At large, these techniques were not
considered fruitful for verifying the identity of the *T. spiralis* secreted proteins with E2 and E3 activity.

In the final attempt to find the enzymes with E2 and E3 activity, LC/MS/MS analysis of total *T. spiralis* secreted proteins was carried out. In 2005 Robinson *et al.* reported 2D electrophoresis and proteomic analysis by MALDI-TOF-MS and LC/MS/MC of selected abundant *T. spiralis* secreted proteins [196,309]. Since then, techniques in tandem mass spectrometry have developed and are much more sensitive. Robinson’s studies were not exhaustive, only selecting the most abundant proteins from a 2D SDS-PAGE gel. Furthermore, in 2005 the available database for matching *T. spiralis* peptides was composed of expressed sequence tag sequences (NCBI dbEST), whereas the annotated draft genome (containing structural and functional annotation where possible) is now available [232]. Earlier experiments suggested that after 4 days of culture, proteins from dead or dying larvae might be released into the culture supernatant, thus contaminating the secreted proteins. For this reason, methods were optimised to ensure that the vast majority of the secreted proteins collected for LC/MS/MS analyses were truly secreted, and that any protein released from dead or dying parasites was negligible. The experiment was also repeated to enhance the credibility of the data. In both experiments peptides matched the same 3 ubiquitin-related proteins. Interestingly, both the TsUBE2L3 fragment open reading frame, and the putative TsRNF13 were also predicted by wolfPSORT to be located extracellularly.

The mammalian orthologue of the putative E3 ligase RNF13 has a mid-sequence transmembrane domain that predominantly localises to the ER and golgi membranes. Proteases can cleave the protein from either side of the membrane, releasing either the N-terminal portion, which contains a PA domain or the C-terminal RING finger portion of the protein [313,314]. It is possible that the *T. spiralis* orthologue also undergoes a form of processing at the transmembrane domain, and that the resulting product is released and secreted at a smaller size than the full annotated sequence. This could explain why it was consistently found in a portion of the gel that did not correspond to its predicted size, suggesting a much smaller protein. It is also worth noting, that the peptides identified by LC/MS/MS only covered a small sequence at the N-terminus where there is an MIR domain but no RING finger domain. The RING finger domain is situated after the transmembrane domain in the C-terminal half of the protein. If the protease-processing theory was the true reason for the size discrepancy, then the cleaved, secreted portion of the protein would not contain the RING finger domain and therefore not be the protein responsible for E3 activity.
Peptides matching the E2 TsUBE2L3 were found in both experiments, in a region of the gel that would correspond to the predicted size of the protein. The annotated TsUBE2L3 sequence is a fragment of a full protein sequence, which translates to a protein of 17 kDa. Most orthologous UBE2L3 sequences that align with this fragment are proteins of 18 kDa, so it is possible that the fragment is very close to the full sequence anyway [65]. Interestingly, the human UBE2L3 was the human E2 used for the *in vitro* ubiquitin conjugation assay. Human UBE2L3 is an E2 partner for the human E3 parkin both *in vivo* and *in vitro*, and in collaboration with E1 they ubiquitinate the Parkinson’s disease-associated proteins synphilin-1 and alpha-synuclein [306,315,316]. It is therefore plausible that the *T. spiralis* secreted protein-derived UBE2L3, can also directly interact with the parkin E3 and the human E1 during the *in vitro* ubiquitin conjugation assay, leading to the ubiquitination of substrate proteins as were observed by immuno-blots.

Having identified TsUBE2L3 and TsRNF13 as the only candidates for the E2 and E3 activity in the *T. spiralis* secreted proteins, I wanted to validate their secretion, and investigate their function. Chapter 5 describes the characterisation of these candidates and the functional analysis of the *T. spiralis* E2, TsUBE2L3.
Chapter 5: Results

Chapter 5: Characterisation of ubiquitin pathway enzymes secreted by *T. spiralis* muscle larvae
Chapter 5: Results

Chapter 5:

Characterisation of ubiquitin pathway enzymes secreted by *T. spiralis* muscle larvae

Chapter 4 describes how ubiquitin conjugation and ubiquitin ligation activity was discovered in the secreted proteins of *T. spiralis* muscle larvae. A putative *T. spiralis* E2 enzyme, UBE2L3 and a putative *T. spiralis* E3 enzyme, RNF13 were identified. The purpose of the experiments described in chapter 5 was to further characterise these secreted *T. spiralis* enzymes and understand their potential function in host-parasite interactions during the infection of muscle cells by *T. spiralis*.

Background

A comprehensive description of the *T. spiralis* life cycle can be found in chapter 1 (section 1.13). Briefly, after the ingestion of *T. spiralis* contaminated meat, larvae develop into adults in the small intestine of their new host. The adults reproduce, giving rise to newborn larvae that migrate through the intestinal wall and enter the circulation. Newborn larvae are transported in the circulation to skeletal muscle tissues, where they invade terminally differentiated skeletal myofibres [127]. The parasite does not kill the muscle cell, rather it forms a complex in which the parasite develops and the host cell is transformed. It is thought that the *T. spiralis* surface and secreted proteins communicate with the host cell to stimulate its re-entry into the cell cycle, thus reversing its differentiation state (dedifferentiation) [156,162,163,169,170].

A comprehensive description of myogenesis can be found in the chapter 1 (section 1.14), however it is important to emphasise that there are very few examples in nature of terminally differentiated skeletal muscle cells that can dedifferentiate, or re-enter the cell cycle. It is thought that the spontaneous regeneration of limbs by amphibians such as the newt, occurs via the dedifferentiation of terminally differentiated muscle cells [147,148]. However mammalian skeletal muscle dedifferentiation has only been observed in response to infection by SV40 and Polyoma viruses and during the *in vitro* manipulation of muscle cell cultures [149-152] [153]. Indeed when *T. spiralis* newborn larvae invade host myofibres, this is exactly what happens. The larvae invade the cell without killing it. Once inside they begin to grow and develop into muscle larvae (L1 stage). During this time, the host
muscle myofibre begins to lose markers of muscle-specific terminal differentiation [131,156]. It also begins to up-regulate cell cycle factors and replicate DNA, in effect re-entering the cell cycle [158,159,162,165]. Although the host cell re-enters the cell cycle, it does not divide and becomes arrested in the G2/M phase [163]. This multinucleated cell has a unique identity, observed only in *T. spiralis* infection, and because it safely harbours the parasite until a chance for transmission occurs, it is referred to as a nurse cell. It is thought that these dramatic biological changes are induced by the parasite surface and secreted proteins, some of which enter the nuclei of the nurse cell [175,201,317]. Very few of these *T. spiralis* secreted proteins have been identified, and even fewer have been functionally characterised [197,202,203]. None have been assigned a mechanistic role in nurse cell development.

The differential balance of specific proteins required during each phase of the cell cycle, relies heavily on the regulation of protein degradation, which in eukaryotes is predominantly controlled by the ubiquitin-proteasome system [52,318]. It is therefore conceivable that the ubiquitin-proteasome system plays a role in *T. spiralis*-induced cell cycle re-entry and dedifferentiation of host muscle cells. Chapter 4 presents evidence that the secreted proteins of *T. spiralis* muscle larvae contain ubiquitin conjugation and ligation activity. This activity was only observed when the secreted proteins came into contact with human ubiquitin conjugation and ligation enzymes, UBE2L3 and parkin. Candidates for this E2 and E3 activity were then identified in the secreted proteins of *T. spiralis* muscle larvae by tandem mass spectrometry. Orthologues of both ubiquitin pathway enzymes identified have been previously linked to the regulation of the cell cycle. The putative *T. spiralis* E2 is orthologous to UBE2L3 (also known as UbcH7, E2-F1, L-UBC, UbcM4), a nuclear and cytoplasmic ubiquitin conjugation enzyme that plays a role in polyubiquitination, cell cycle control and transcription [306,319-321]. Substrates of this E2 in mammalian systems include the tumour suppressor p53, the NFκB precursor p105, the inflammatory cytokine TNF (tumour necrosis factor) and the metabolic enzyme glyceraldehyde-3-phosphate dehydrogenase [66,67,322-324]. It was also shown to interact with a rabbit E3 ligase (E3L) leading to the ubiquitination of some muscle-specific proteins [325]. The putative *T. spiralis* E3 is orthologous to RNF13, a ring finger ubiquitin ligase found to be over-expressed in pancreatic cancer [326].

The final aim of the project was therefore to confirm the identity of the proteins responsible for the E2 and E3 activities, and to try and characterise their effect on
Chapter 5: Results

muscle cells. Chapter 5 describes how the coding sequences of the candidate E2 and E3 enzymes were analysed using RACE-PCR. The function of the candidate E2 enzyme was then investigated in a mammalian skeletal muscle cell line, C2C12 whereby the effect of TsUBE2L3 on muscle cell morphology, protein synthesis, the cell cycle and proliferation was investigated.

NB: Throughout the text, I use the term myofibre when referring to terminally differentiated muscle cells in vivo, and the term myotube when referring to terminally differentiated muscle cells developed in vitro. Thus, terminally differentiated C2C12 cells are always referred to as myotubes.
Chapter 5 Results:

5.1. Validation of the secretion of TsUBE2L3 (E2 enzyme) and TsRNF13 (E3 enzyme) by *T. spiralis* muscle larvae

Tandem mass spectrometry identified only 2 ubiquitin-related enzymes in the secreted proteins of *T. spiralis* muscle larvae, a putative *T. spiralis* E2 enzyme, UBE2L3 and a putative *T. spiralis* E3, RNF13. These proteins were therefore the only candidates for the *in vitro* ubiquitin conjugation and ubiquitin ligation activity observed in the *T. spiralis* SP.

The putative *T. spiralis* E3 (TsRNF13) sequence was first aligned with the human orthologue for closer analysis, showing 21.6% identity (Figure 5.1 A). The putative RNF13 protein was identified during mass spectrometry by peptides covering only 2.5% of the 525 amino acid sequence. These peptides did not match any of the conserved RING finger domain of the sequence as predicted by SMART [210,211]. To verify the secretion of TsRNF13 by *T. spiralis* muscle larvae, a search was carried out to find commercial antibodies that were raised against the mammalian orthologue that, based on identity between the immunogen sequence and the corresponding *T. spiralis* sequence, appeared likely to cross-react. Because of the low level of identity between the mammalian and *T. spiralis* orthologues, no commercial antibodies could be found to fit the criteria and so the secretion of the E3 enzyme by *T. spiralis* could not be verified.

The putative *T. spiralis* UBE2L3 sequence however, contains 62.3% of the same residues as the human UBE2L3 (Figure 5.1 B). A commercial antibody raised against a region of the human protein that has a high identity with the *T. spiralis* protein was purchased. *T. spiralis* muscle larvae were isolated from infected muscle tissues as described (section 2.2, [116]) and cultured for 24 hours before the secreted proteins were collected and concentrated. Day 1 *T. spiralis* SP were analysed by immuno-blot using anti-UBE2L3 antibodies. *T. spiralis* muscle larvae lysate and human cell lysate (HEK 293T cells) were also reacted with the anti-UBE2L3 antibodies as positive controls (Figure 5.1 C).
Figure 5.1  

A. Alignment of the putative T. spiralis (Ts) E3 (RNF13) protein sequence with the human (Hs) orthologue (homology and identity shaded in grey and dark grey respectively). The MIR domain is boxed in green, the transmembrane domain in blue and the RING finger (E3) domain in red. The peptide sequence identified by LC/MS/MS (highlighted in yellow) covered 2.5% of the 524 amino acid sequence.

B. Alignment of the putative T. spiralis (Ts) E2 (TsUBE2L3, GI:316975361) fragment protein sequence with the human isoform 1 (Hs) orthologue. The UBCc domain is boxed in red. Peptide sequences identified by LC/MS/MS are yellow. The catalytic cysteine residue of the UBCc domain is marked with a star symbol. The immunogen sequence of the HsUBE2L3 against which anti-UBE2L3 antibodies were raised, is boxed in black.

C. Anti-UBE2L3 antibodies were reacted with the secreted proteins (SP) of day-1 T. spiralis muscle larvae, human (HEK) cell lysate and T. spiralis lysate by immunoblot (IB). The same samples were reacted with anti-tubulin antibodies as a control for secretion.

Figure 5.1 Analysis of the ubiquitin-related enzymes identified by the proteomic analysis of T. spiralis secreted proteins.

A. Alignment of the putative T. spiralis (Ts) E3 (RNF13) protein sequence with the human (Hs) orthologue (homology and identity shaded in grey and dark grey respectively). The MIR domain is boxed in green, the transmembrane domain in blue and the RING finger (E3) domain in red. The peptide sequence identified by LC/MS/MS (highlighted in yellow) covered 2.5% of the 524 amino acid sequence. 

B. Alignment of the putative T. spiralis (Ts) E2 (TsUBE2L3, GI:316975361) fragment protein sequence with the human isoform 1 (Hs) orthologue. The UBCc domain is boxed in red. Peptide sequences identified by LC/MS/MS are yellow. The catalytic cysteine residue of the UBCc domain is marked with a star symbol. The immunogen sequence of the HsUBE2L3 against which anti-UBE2L3 antibodies were raised, is boxed in black.

C. Anti-UBE2L3 antibodies were reacted with the secreted proteins (SP) of day-1 T. spiralis muscle larvae, human (HEK) cell lysate and T. spiralis lysate by immuno-blot (IB). The same samples were reacted with anti-tubulin antibodies as a control for secretion.
In the human cell lysate, one prominent band was observed in the human cell lysate, corresponding to the predicted size of the human UBE2L3 isoform 1 that weighs 17.9 kDa [324]. In the lysate of *T. spiralis* muscle larvae, 2 prominent bands between 15 and 25 kDa were observed. Other less prominent bands were also observed. In the *T. spiralis* SP, the most prominent band was observed between 15 and 25 kDa, matching the smaller band in the *T. spiralis* lysate, and the band in HEK cell lysate. In the *T. spiralis* SP, a second less abundant protein was observed just below the most prominent band that may have represented a partially degraded from of the larger protein. The fragment coding sequence of the putative TsUBE2L3 translates into a protein with a predicted molecular weight of 16.9 kDa, which would approximately correspond to the most prominent band observed in the *T. spiralis* SP.

To test for proteins that might have been release from dead or dying larvae, i.e. not actively secreted, the samples were reacted with anti-tubulin antibodies that have a high affinity for both the human and the *T. spiralis* tubulin. Tubulin is a highly conserved structural protein that was not found during the proteomic analyses (by mass spectrometry) of *T. spiralis* SP presented in chapter 4. It would therefore only be present in the culture supernatant if released from dead or dying larvae.

Both the mass spectrometry data presented in chapter 4, and the immuno-blot data presented here provide evidence for the expression and secretion of the TsUBE2L3 by *T. spiralis* muscle larvae. This enzyme may be responsible for the E2 activity observed in *T. spiralis* SP. I therefore wanted to investigate the potential function of the E2 in *T. spiralis* SP, and the potential role for this function in the host-parasite interactions during the muscle stages of *T. spiralis* infection in mammals. To do so, I decided to clone the *T. spiralis* UBE2L3 gene for expression in mammalian muscle cells. It was therefore necessary to first confirm the coding sequence for this protein. This was done using the RACE-PCR method described in chapter 3 (section 3.1).

### 5.2. Confirmation of the full coding sequence of TsUBE2L3 by RACE-PCR

Rapid amplification of cDNA ends by PCR (RACE-PCR) is a technique used to confirm the full sequence of a gene transcript by amplifying and sequencing the 5’ end of the transcript (containing the start codon) and the 3’ end of the transcript (containing the stop codon). Multiple attempts to amplify the coding sequence of the putative *T. spiralis* E3, RNF13, failed. Attempts were made to optimise methods until
it was concluded that the annotation of the sequence might be incorrect. It was therefore not possible to confirm the coding sequence of the *T. spiralis* E3 candidate and, since its secretion could not be verified either, studies continued to focus on the *T. spiralis* E2, TsUBE2L3.

The annotated coding sequence for TsUBE2L3 exists as a fragment, with a missing portion at the 5’ end that does not contain a start codon. The human orthologue of UBE2L3 has 3 isoforms, isoform 1 (GI: 373432685, 465 bp, 154 aa, 17.9 kDa) isoform 3 (GI: 373432681, 369 bp, 122 aa, 14.1 kDa) and isoform 4 (GI: 373432683, 639 bp, 212 aa, 24 kDa), although in mice it is thought that isoform 3 and 4 function as pseudogenes [327] (Figure 5.2 A). The most abundant protein in HEK 293T cells corresponded to the size of isoform 1 at approximately 18 kDa (Figure 5.1 C). The human orthologue of UBE2L3 can also be modified with both poly and monoubiquitin and is itself degraded by the proteasome (appendix 4) [320,324]. In *T. spiralis* lysate, signals from proteins of 2 distinct sizes were observed during immuno-blots using anti-UBE2L3 antibodies (Figure 5.1 C). These bands may represent different isoforms of the *T. spiralis* protein or, if the regulation of UBE2L3 by ubiquitination is conserved in *T. spiralis*, may represent ubiquitinated forms of the protein. The full transcript sequence expressed by *T. spiralis* muscle larvae had to be experimentally confirmed by RACE-PCR before cloning could be carried out.

Total RNA was extracted and purified from *T. spiralis* muscle larvae using Trizol and general methods of phenol/chloroform precipitation. Primers for RACE-PCR were designed based on the region of the TsUBE2L3 sequence where peptides were identified during mass spectrometry of *T. spiralis* SP. A gene-specific reverse primer (GSRP) was designed to amplify the 5’ end, and a gene-specific forward primer (GSFP) was designed to amplify the 3’ end (Figure 5.2 and 5.3 respectively). RACE was then carried out using the FirstChoice RLM-RACE kit from Ambion and the GeneRacer kit from Life Technologies. *T. spiralis* RNA was processed as instructed by the manufacturer and reverse transcribed to cDNA before undergoing 2 rounds of nested PCR using the RACE 5’ primers that anneal to the 5’ RACE oligo, paired with the gene-specific reverse primers, and the 3’ RACE oligo paired with the gene-specific forward primers. PCR reactions were carried out using a Taq DNA polymerase enzyme that assembles sequences leaving an adenine DNA base overhang on each 3’ end of the complementary strands of the PCR product. PCR products were separated by agarose gel electrophoresis and visualised under UV light after incubation with ethidium bromide (data not shown). DNA bands that appeared to correspond with the potential size of the 5’ end fragment (based on
orthology with one of the 3 human HsUBE2L3 isoforms), and the potential size of the 3’ end (based on the annotated TsUBE2L3 sequence), were manually excised from the agarose gel. DNA was purified from these bands and ligated into the pGEMTeasy cloning vector using methods of TA cloning. The pGEMTeasy vector containing the 5’ or 3’ DNA fragment insert was then sequenced by Beckman Coulter Genomics.

The 5’ sequence matched the annotated fragment sequence, and continued upstream for 27 bases until an in-frame start codon (Figure 5.2 B). There was then another small sequence of 14 bases before the 5’ RACE oligo sequence began. This additional sequence aligned with a portion of the same T. spiralis contig sequence that the TsUBE2L3 fragment aligned with (contig scaffolds: ABIR00000000.2 [232]). This sequence may have represented a 5’ untranslated region (5’ UTR) of the transcript. The sequence that contained the start codon however, aligned much further upstream of the contig, suggesting a large splice site in between the annotated fragment sequence and the start site of this potential isoform of TsUBE2L3. This would explain the difficulty in predicting this gene during the draft annotation. The additional 14 bases of this sequence may represent a 5’ untranslated region of the transcript, or represent genomic DNA contamination (although no introns were present in the coding sequence). The RACE-PCR-generated 5’ end sequence was aligned with the human sequence for HsUBE2L3 isoform 1 (Figure 5.2 A), and many other orthologous sequences (data not shown). The RACE-PCR-confirmed start site aligned very closely with these orthologous sequences.
Figure 5.2  Life Technologies GeneRacer amplification of 5’ cDNA ends of TsUBE2L3

A. The RACE-PCR sequencing data (RACE-PCR sequence) of the 5’ end of TsUBE2L3 cDNA aligned with the existing fragment annotated coding sequence. Figure shows the position of the RACE oligo, the RACE 5’ forward primer-binding site, the start site of the full coding sequence of the gene, the custom gene-specific reverse primer (GSRP)-binding site and the cloning vector (pGEMTeasy) sequences.
Chapter 5: Results

The 3’ end sequence was 100% identical to the annotated fragment sequence, including the predicted stop codon. Beyond this was additional sequence that, when analysed by BLAST, matched genomic DNA of *T. spiralis* possibly representing the 3’ untranslated (UTR) region of the transcript. The full RACE-PCR-confirmed sequence, from start to stop, aligned very closely with these orthologous sequences (Figure 5.3 B).

The entire RACE-PCR-generated sequence was matched by BLAST against the *T. spiralis* genome. The sequence was found to match the reverse strand of contig 0.377. Contig 0.377 was then analysed by AUGUSTUS gene prediction software. AUGUSTUS can use informative constraints or ‘hints’ based on the gene arrangement in a given organism. When the contig is analysed using *T. spiralis* as a hint, a gene containing part of the RACE-PCR confirmed sequence is predicted. This sequence only partially aligns with UBE2L3 orthologues. *Brugia malayi* is a parasitic nematode responsible for lymphatic filariasis in humans. The predicted *B. malayi* UBE2L3 orthologous protein sequence shows 44.7% identity with the *T. spiralis* UBE2L3 (notably less than shared between the human and *T. spiralis* orthologues). When the contig is analysed by AUGUSTUS using *Brugia malayi* as a hint, 2 gene arrangements are predicted. The first and smallest aligns closely with the human UBE2L3 isoform 3, translating to a putative *T. spiralis* protein of 16.6 kDa. The second aligns closely with the human UBE2L3 isoform 1, translates to a protein of 18 kDa and is exactly the same sequence identified by RACE-PCR (data not shown). The 2 gene arrangements may represent different *T. spiralis* UBE2L3 isoforms. Alignments of the RACE-PCR confirmed translated sequence with multiple protein orthologues, as well as the human isoform 1 protein sequence, showed that the start site of the protein was well conserved (Figure 5.3 B). This sequence was therefore taken as the true open reading frame for TsUBE2L3 as expressed and secreted by *T. spiralis* muscle larvae.
Chapter 5: Results

Figure 5.3  Ambion RLM-RACE amplification of 3’ cDNA ends of TsUBE2L3

A. The RACE-PCR sequencing data (RACE-PCR sequence) of the 3’ end of TsUBE2L3 cDNA aligned with the existing fragment annotated coding sequence. Figure shows the position of the RACE oligo, the RACE 3’ forward primer-binding site, the stop site of the full coding sequence of the gene, the custom gene-specific forward primer (GSFP)-binding site and the cloning vector (pGEMTeasy) sequences. The sequence contained a 3’ continuation after the stop codon that was identified as T. spiralis genomic DNA (gDNA).

B. The annotated fragment (incomplete) coding sequence for TsUBE2L3 (GI:316975344) was aligned with the full RACE-PCR confirmed sequence from start to stop, compiled from both 5’ RACE-PCR and 3’ RACE-PCR data, and with the human isoform 1, 3 and 4 UBE2L3 coding sequences.

# = isoform number
5.3. Differentiation of the C2C12 mouse skeletal muscle cell line from myoblasts into myotubes

The next aim of the project was to investigate the potential role for the *T. spiralis* UBE2L3 protein in nurse cell formation during muscle stages of *T. spiralis* infection. To do so experiments were carried out to clone the confirmed open reading frame and express the protein in a mammalian skeletal muscle cell line.

C2C12 is a mouse myoblast cell line often used as a laboratory model for studies on skeletal muscle cell biology. C2C12 myoblasts were cultured to confluency and induced to differentiate using differentiation media (low-serum media lacking in growth factors), where they were observed to fuse, forming large elongated myotubes that aligned in orientation, representing their successful differentiation in culture (Figure 5.4 A). Terminal differentiation of myoblasts into skeletal myotubes is characterised by the expression of muscle specific proteins such as the early differentiation factor myogenin and the late differentiation factor myosin heavy chain (MHC) [328,329]. To verify the differentiation of wild-type C2C12 myoblasts into myotubes, anti-myogenin and anti-MHC antibodies were reacted with C2C12 lysates by immuno-blot (Figure 5.4 B). Since there are at least 4 different types of MHC expressed in mammalian skeletal muscle cells, an antibody was used that reacts specifically with MHC IIx/2x. This protein is encoded by the MYH1 gene. Antibodies raised against tubulin were also reacted with the lysates as a loading control. Despite the expression of tubulin as expected, no myogenin expression was observed in undifferentiated C2C12 myoblasts. Both myogenin and MHC expression was observed in differentiated C2C12 myotubes, confirming the successful differentiation of C2C12 cells in culture using this method.
Chapter 5: Results

Figure 5.4

A. C2C12 mouse myoblasts were successfully cultured and differentiated using differentiation media. After 4 days mature myotubes were observed. B. The immature state of the myoblasts was confirmed by immuno-blot (IB) using anti-myogenin (differentiation marker) antibodies. C. The differentiation of myotubes was confirmed by immuno-blot (IB) using anti-myogenin and anti-MHC (myosin heavy chain differentiation marker) antibodies. The same cell lysate samples were reacted with anti-tubulin antibodies as a protein loading control.
5.4. Cloning and expression of TsUBE2L3-HA in C2C12 skeletal muscle cells

During infection, *T. spiralis* newborn larvae invade already terminally differentiated myofibres [330]. The larvae then develop into muscle stage larvae (L1 stage) inside the myofibre [162]. To understand the effect of the *T. spiralis* secreted proteins (specifically the effect of TsUBE2L3) on the muscle cell, it was necessary to study differentiated myotubes rather than undifferentiated myoblasts. Differentiated myotubes are in a state of cell cycle arrest and do not divide. It is therefore not possible to transfect these cells using conventional methods of cell transfection, which require cell division events to take up the vector DNA. Lentiviruses can infect and mediate the integration of DNA in both dividing and non-dividing cells such as terminally differentiated skeletal muscle cells [331]. For this reason lentiviral vectors were chosen for the genetic manipulation of C2C12 myotubes.

Genetically modified lentiviruses can be produced to package genetic material of interest. The packaging process takes place in the packaging cell line HEK 293T (Figure 5.5 C). Inside the packaging cell, proteins required for the production of lentivirus particles are expressed, and the genetic material of interest is transcribed into mRNA. This mRNA is packaged inside the virus particle. Virus particles collected from the culture supernatant are then used to infect target cells and deliver the mRNA. This is reverse transcribed to DNA before being integrated into the genome of the cell for the stable expression of the gene of interest. This process of infection is called transduction (rather than transfection).

The first lentiviral vector to be tested was the HIV-derived pCSGW tdTomato, kindly provided by Ilaria Nisoli and Hugh Brady (Imperial College). This (largely ‘home-made’) vector was chosen because it contains the gene for the fluorescent reporter, tdTomato under the translational control of an internal ribosome binding site (IRES) sequence. An IRES sequence allows the initiation of transcription of the gene to be active at all times. The expression of the bright red fluorescent protein tdTomato, allows for visual confirmation that the genetic material packaged into the lentivirus has been successfully transduced. The RACE-PCR-confirmed coding sequence of TsUBE2L3 with a C-terminal haemagglutinin (HA) tag (18 kDa) was cloned into pCSGW (using NotI and BamHI restriction sites) downstream of the HIV LTR 5’ promoter sequence and upstream of the IRES tdTomato sequence (Figure 5.5 A) [332]. Lentivirus containing pCSGW tdTomato TsUBE2L3-HA RNA and lentivirus containing pCSGW tdTomato only RNA (empty vector control, Figure 5.5 B) was
produced in HEK 293T cells (Figure 5.5 C). C2C12 myoblasts were grown to confluency and induced to differentiate for 3 days using differentiation media. Half of the cells were transduced with the TsUBE2L3-HA virus and half of the cells were transduced with the empty vector control virus using polybrene and centrifugation to enhance transduction efficiency. After 72 hours (allowing time for infection, integration and expression) cells were observed using fluorescence microscopy. At both the packaging stage in HEK 293T cells (Figure 5.5 D) and the transduction stage in C2C12 myotubes (Figure 5.5 E), red fluorescence representing tdTomato expression was observed. Using pCSGW tdTomato transduction efficiencies varied between experiments, and did not exceed 50% of the total cells (based on the proportion of red fluorescent cells). This may have been due to the disruptive process of lentivirus transduction. Firstly, polybrene is toxic and although the transduction media was removed as soon as possible, C2C12 myotubes appeared less healthy after exposure to polybrene. Secondly, cells were centrifuged with the lentivirus, this is known as a ‘spinfection’. Although this also greatly improves transduction efficiency, C2C12 myotubes were structurally damaged in the process. Furthermore, after 3 days of differentiation followed by 72 hours of transduction, myotubes began to die. For these reasons this system of transduction was discontinued and another lentiviral vector system was employed.

It was decided that transduction would be greatly enhanced in dividing myoblasts because cells could be expanded after transduction. However, the aim of the study was to investigate the effect of the *T. spiralis* protein on myotubes (rather than myoblasts) and so I wanted to prevent expression of TsUBE2L3-HA until cells were already differentiated. The Lenti-X Tet On Advanced vector system from Clontech was chosen for two reasons: 1. because the vectors contain drug selection markers which would allow for the development of a stable myoblast cell line and 2. because the vectors mediate a doxycycline-controlled inducible system of expression, which could be switched on only after differentiation of the cell line.
Figure 5.5 Cloning TsUBE2L3-HA into the pCSGW lentiviral vector

A. The RACE-confirmed TsUBE2L3 coding sequence with a C-terminal HA tag was cloned into pCSGW downstream of the HIV-1 LTR promoter sequence and upstream of an IRES (internal ribosome entry site) tdTomato (red fluorescent reporter) sequence. B. The empty pCSGW tdTomato vector was used as a control. C. Schematic showing how lentivirus particles are produced in HEK 293T cells to package genetic material of interest. These lentivirus particles are then used to infect (or transduce) target cells, into which the genetic material is injected, becoming reverse transcribed to cDNA that is able to integrate into the genome of the target cell for expression of the protein of interest. D. Lentivirus particles carrying the TsUBE2L3-HA tdTomato genetic material and the empty vector tdTomato genetic material were made in HEK 293T packaging cells that were observed by light and fluorescence microscopy. E. Lentivirus particles were used to transduce mature C2C12 myotubes that were observed by light and fluorescence microscopy.
The system involves 2 vectors, a control vector that expresses a doxycycline-sensitive protein, and an expression vector into which the gene of interest is cloned (Figure 5.6 A, B and C). The pLVX Tet On control vector contains a neomycin resistance cassette and the pLVX Tight Puro expression vector with a puromycin resistance cassette. The coding sequence of TsUBE2L3 with a C-terminal haemagglutinin (HA) tag was cloned into the multiple cloning site of pLVX Tight Puro expression vector using BamHI and NtI restriction sites. This positions the gene of interest downstream of the pTight modified Tet-responsive promoter. The pLVX Tet On control vector contains a gene for the transactivator protein rtTA-Advanced. RtTA-Advanced contains a DNA-binding TetR domain. In the presence of the drug doxycycline (DOX), the rtTA-Advanced protein is activated and binds to the modified Tet-responsive promoter of the pLVX Tight Puro expression vector. This drives the expression of the gene of interest (either TsUBE2L3-HA or nothing from the empty vector Figure 5.6 A, B and C).

Three lentivirus particle lines were produced in HEK 293T cells: 1. the pLVX Tet On control vector, 2. the pLVX Tight TsUBE2L3-HA and 3. the pLVX Tight empty vector control (Figure 5.6 D). C2C12 myoblasts were transduced with a mixture of the control vector lentivirus paired with each expression vector lentivirus (virus 1+2, TsUBE2L3-HA or virus 1+3, empty vector). Cells were then selected for dual neomycin and puromycin resistance. Two cell lines were produced, one carrying both pLVX Tet On and pLVX Tight empty vector DNA, and another carrying pLVX Tet On and pLVX Tight TsUBE2L3-HA DNA. Once selected, myoblast cell lines were maintained under a low level of neomycin and puromycin selection. Cells were grown to confluency and then induced to differentiate for 3 days using differentiation media. Doxycycline was then added to the media to induce the expression of either no protein (empty vector) or the recombinant T. spiralis E2 enzyme, TsUBE2L3-HA. Expression of the HA-tagged E2 in C2C12 myotubes was confirmed by immuno-blots of the HA epitope tag at 4 time-points over a 48-hour period after the addition of doxycycline to the culture media (Figure 5.6 D). Reactivity to anti-HA antibodies, representing the expression of TsUBE2L3-HA, was observed at the expected size of approximately 18 kDa at 18, 22 and 48 hours post-doxycycline induction, but not at T0 before doxycycline was added. No anti-HA reactivity was observed in the empty vector control cells confirming the specificity of the signal. Anti-tubulin antibodies were also reacted as a control for equal protein loading.
Figure 5.6 Cloning TsUBE2L3-HA into the inducible pLVX lentiviral vector

A. The Lenti-X Tet-On Advanced Expression System control vector, pLVX-Tet-On containing the coding sequence for the rtTA protein under the constitutive control of a CMV promoter. B. The RACE-confirmed TsUBE2L3 coding sequence with a C-terminal HA tag was cloned into the Lenti-X Tet-On Advanced Expression System expression vector pLVX-Tight-Puro, downstream of a doxycycline (DOX)/rtTA responsive pTight promoter and upstream of a neomycin resistance cassette. When DOX is added, the DOX/rtTA protein complex binds to the pTight promoter inducing transcription of TsUBE2L3-HA C. The empty pLVX Tight-Puro vector was used as a control. D. Schematic of the production of lentivirus in HEK 293T cells followed by the transduction of C2C12 myotubes, and the selection of neomycin and puromycin resistant myoblasts for the generation of a stable cell line. E. Immuno-blot (IB) analyses of TsUBE2L3-HA expression in myoblast cell lines, 0, 18, 22 and 48 hours post DOX induction using anti-HA antibodies and anti-tubulin antibodies as a loading control.
Figure 5.7  pLVX inducible system expression of TsUBE2L3-HA in C2C12 myoblasts

A. Experiment set-up: neomycin and puromycin-selected empty vector and TsUBE2L3-HA C2C12 myoblast cell lines. Half the cells were induced using doxycycline for 24 hours (+DOX) and half were not induced. Cells were then analysed by immuno-fluorescence (IFA).

B. Myoblasts were processed for IFA and incubated with DAPI (to stain nuclei) and anti-HA antibodies (with Alexa-488 secondary antibodies). Samples were visualised at 20x magnification using confocal microscopy. Figure shows the brightfield image (BF), the DAPI (blue) the Alexa-488 (green), the overlay and the overlay plus the BF image.
Doxycycline-induced expression of TsUBE2L3-HA, in both undifferentiated myoblasts (Figure 5.7) and differentiated myotubes (Figure 5.8), was also confirmed by immuno-fluorescence using anti-HA antibodies (raised in rat) and Alexa-488 (green) conjugated anti-rat antibodies. All immuno-fluorescence experiments were visualised by confocal microscopy. Minimal background staining was observed in control samples prepared with the anti-rat Alexa-488 secondary antibodies only (no anti-HA antibodies, data not shown). No significant fluorescence was observed in the empty vector myoblasts or myotubes, before or after induction. Although no fluorescence was observed in pLVX Tight TsUBE2L3-HA myoblasts pre-induction, some fluorescent signals were observed by immuno-fluorescence in the pLVX Tight TsUBE2L3-HA myotubes pre-induction (data not shown). This ‘leaky’ expression was not observed at T0 post-induction by immuno-blot using anti-HA antibodies (Figure 5.6 E). Expression of TsUBE2L3-HA in undifferentiated myoblasts was observed as cytoplasmic and in many cells also nuclear (Figure 5.8, 5.9 and 5.10). When cytoplasmic, signals were observed throughout the cells, with occasional pockets of diminished signal. It was noted that the anti-HA signal was more often nuclear in cells that were in the process of dividing. Occasionally, the signal inside the dividing nuclei was observed to localise in patches (Figure 5.10). In myotubes, the signal appeared to be predominantly cytoplasmic, and rarely nuclear. Finally, not all cells reacted with the anti-HA antibodies. This was surprising, considering that the myoblasts cell line was selected and maintained using neomycin and puromycin. In theory, all cells of the TsUBE2L3-HA cell line should contain the gene for the T. spiralis protein. The variation in the intensity of the signal from cell to cell was most likely to be an artefact of the IFA process, however some cells contained no signal at all suggesting irregular lentivirus transduction or promoter-driven expression.

Initial observations also suggested that there were subtle differences in the morphology of the C2C12 cell line induced to express TsUBE2L3-HA compared with the empty vector cell line. This was therefore further investigated.
Figure 5.8 pLVX inducible expression of TsUBE2L3-HA in C2C12 myotubes

A. Experiment set-up: neomycin and puromycin-selected empty vector and TsUBE2L3-HA C2C12 myoblast cell lines were differentiated into myotubes. Half the cells were induced using doxycycline for 24 hours (+ DOX) and half were not induced. Cells were then analysed by immuno-fluorescence (IFA). B. Myotubes were processed for IFA and incubated with DAPI (to stain nuclei) and anti-HA antibodies (with Alexa-488 secondary antibodies). Samples were visualised at 20x magnification using confocal microscopy. Figure shows the brightfield image (BF), the DAPI (blue) the Alexa-488 (green), the overlay and the overlay plus the BF image.
Characterisation of the function of the secreted *T. spiralis* E2 enzyme, UBE2L3:

5.5. The effect of TsUBE2L3-HA on C2C12 skeletal muscle cell morphology

In the presence of TsUBE2L3, both myoblasts and myotubes appeared (by light microscopy) to have a slightly modified morphology. The effect on myoblasts was less predominant (data not shown), but myotubes often appeared less organised and uniform, forming more lateral fusions with neighbouring myotubes (Figure 5.9 A). In order to visualise the differences in morphology with more clarity, immuno-fluorescence experiments were carried out using anti-tubulin antibodies, in order to visualise the cytoskeletal network and therefore the general structure of the cells.

Both C2C12 myoblast cell lines, one carrying the pLVX TsUBE2L3-HA expression vector and the other carrying the pLVX empty expression vector, were counted and seeded at the same density. Half of each cell line was differentiated into myotubes, the other half of each cell line passaged as normal and maintained as undifferentiated myoblasts. Cells were induced with doxycycline for 24 hours and processed for immuno-fluorescence analysis. Cells were incubated with anti-tubulin antibodies and anti-HA antibodies and fluorophore-conjugated secondary antibodies (Figure 5.9 B). Minimal background staining was observed in control samples prepared with the secondary antibodies only (no primary antibodies, data not shown). In the presence of TsUBE2L3-HA, undifferentiated myoblasts often contained individual cells that were much larger than the majority of the other cells. These enlarged cells often contained enlarged nuclei. Although enlarged cells were observed in the empty vector myoblasts samples as well, the frequency of enlarged cells was greater in the presence of TsUBE2L3-HA. Differentiated myotubes expressing TsUBE2L3-HA demonstrated an overall weaker anti-tubulin signal than the empty vector control myotubes, however this observation was not reflected in immuno-blot analyses of tubulin in cell lysates, which, after normalisation of protein concentrations, remained equivalent throughout the study. Myotubes appeared to be less well organised in orientation with one-another. In the presence of TsUBE2L3-HA, myotubes appeared wider and contained more lateral fusion events.
Figure 5.9 A  Analysis of the effect of TsUBE2L3 on C2C12 morphology

A. C2C12 myotubes were induced using doxycycline (+ DOX) for 22 hours. Using light microscopy, the morphology of myotube cells expressing TsUBE2L3-HA were compared to those carrying empty vector DNA. Red arrows indicate areas of orientation between myotubes. Black arrows indicate areas of extensive branching.
Figure 5.9 B Analysis of the effect of TsUBE2L3 on C2C12 morphology

B. C2C12 myoblast and myotube cell lines (empty vector and TsUBE2L3-HA) were harvested and processed for immuno-fluorescence (IFA) 24 hours after induction using doxycycline. Nuclei were stained using DAPI, and cells were incubated with anti-tubulin antibodies (with red Alexa-555 secondary antibodies) and anti-HA antibodies (with green Alexa-488 secondary antibodies). Samples were visualised using confocal microscopy at 63x magnification. Figure shows the brightfield image (BF), the DAPI, tubulin and HA signals, the overlay and the overlay plus the BF image.
Analysis of the subcellular localisation of TsUBE2L3-HA in C2C12 cells

C2C12 myoblast and myotube TsUBE2L3-HA cell lines were induced using doxycycline for 24 hours and harvested and processed for immuno-fluorescence (IFA). Nuclei were stained using DAPI, and cells were incubated with anti-HA antibodies (with green Alexa-488 secondary antibodies). Samples were visualised using confocal microscopy at 63x magnification and where indicated the manual zoom function was used. Figure shows the brightfield image (BF), the DAPI, tubulin and HA signals, the overlay and the overlay plus the BF image. Images were taken of (A) a myoblast showing predominant cytoplasmic localisation of TsUBE2L3-HA, (B) a dividing myoblast, showing the condensed localisation of TsUBE2L3-HA signal inside the nucleus (arrows) and (C) a mature myotube showing exclusive cytoplasmic localisation of TsUBE2L3-HA.
Unfortunately this data was not quantitative. Due to the variations in phenotype observed within individuals of a large population of cells, I was unable to design an accurate method of quantification of the differences in morphology. The next aim was to try and explain the underlying cause for the changes in morphology observed. Since the *T. spiralis* SP-induced dedifferentiation of muscle cells during infection causes a characteristic change in morphology of the host cell, I decided to investigate the effect of TsUBE2L3-HA expression on the differentiation state of C2C12 myotubes. To do so the effect of TsUBE2L3 on the protein levels of C2C12 myotube differentiation factors was analysed.

### 5.6. The effect of TsUBE2L3-HA on C2C12 skeletal muscle cell differentiation factors

Following the invasion of skeletal muscle by *T. spiralis* during infection, host cell expression of muscle-specific proteins such as myosin heavy chain (MHC), tropomyosin and myogenin is lost, [131,156,163]. In addition, Bai *et al.* showed that the secreted proteins of *T. spiralis* muscle larvae are also able to inhibit the differentiation process of C2C12 cells when added to them *in vitro* [160,161]. This study also showed a specific inhibition of the early differentiation marker myogenin and the late differentiation marker MHC 2x. These proteins were therefore chosen as appropriate markers of differentiation in C2C12 cell lines. C2C12 cell lines carrying either the pLVX empty expression vector DNA or the pLVX TsUBE2L3-HA expression vector DNA were counted and seeded at the same density. Half of each cell line was differentiated into myotubes, the other half was passaged as normal and maintained as undifferentiated myoblasts. All cells were then induced using doxycycline and harvested at the indicated time-points post-induction (Figure 5.11). The protein concentrations of the cell lysates were normalised and immuno-blot analyses were carried out using anti-MHC and anti-myogenin antibodies. Anti-tubulin antibodies were reacted with the same samples as a loading control. Figure 5.11 B is representative of 3 biological repeats of this experiment.

In the myoblast sample, because the cells are rapidly dividing, the cell confluency significantly increased over the 48-hour period of induction, allowing more points of contact to form between myoblasts. These points of contact promote spontaneous differentiation, which involves the up-regulation of tubulin as the cells begin to fuse in the early stages of myotube formation. I propose that this is why the tubulin loading control signal is observed to increase over time in the myoblast samples. Despite this,
no reactivity with anti-myogenin antibodies was observed in myoblast lysates at the time of doxycycline induction (T0) or at 18, 22 and 48 hours post-induction (Figure 5.11 A). In myotubes no changes were observed in the levels of myogenin or MHC expression in the presence of TsUBE2L3 until 48 hours post-induction. Here, a reduction in myogenin was observed in the TsUBE2L3-HA myotubes when compared to the empty vector myotubes (Figure 5.11 B). Although no change was observed between 0 and 22 hours, at 48 hours post-induction a slight reduction in MHC was also observed in the TsUBE2L3-HA myotubes when compared to the empty vector myotubes. This result was not significant, and may have been due to small differences in the amount of protein loaded for SDS-PAGE that could not be visually detected using the anti-tubulin loading control. In order to draw any conclusions about the effect of TsUBE2L3-HA on MHC and myogenin in C2C12 myotubes, this experiment would need to be repeated. Given more time, it would be of interest to repeat the experiment with an extended time-course.

During this experiment, it was observed that each time a cell sample was lysed, the protein concentration (before normalisation) was always higher in the cell line carrying the TsUBE2L3-HA vector DNA than in the cell line carrying the empty vector DNA. If the total number of cells and the rates of proliferation in each cell line were equal, an increase in protein concentration could be caused by an increased rate of protein synthesis. Alternatively, an increase in protein concentration could be caused by an increased number of cells caused by an increased rate of proliferation. Both of these were investigated, starting with a measurement of protein synthesis.
Chapter 5: Results

Figure 5.11 Immuno-blot analysis of differentiation markers in C2C12 cells

C2C12 pLVX stable cell lines $\rightarrow$ 24 h doxycycline induction $\rightarrow$ immuno-blot (IB)

<table>
<thead>
<tr>
<th>A</th>
<th>IB: myogenin</th>
<th>IB: tubulin</th>
</tr>
</thead>
<tbody>
<tr>
<td>empty vector</td>
<td>TsUBE2L3-HA</td>
<td>C2C12 myoblasts</td>
</tr>
<tr>
<td>IB: myogenin</td>
<td>empty vector</td>
<td>TsUBE2L3-HA</td>
</tr>
<tr>
<td>IB: MHC</td>
<td>empty vector</td>
<td>TsUBE2L3-HA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>IB: myogenin</th>
<th>IB: tubulin</th>
</tr>
</thead>
<tbody>
<tr>
<td>empty vector</td>
<td>TsUBE2L3-HA</td>
<td>C2C12 myotubes</td>
</tr>
<tr>
<td>IB: myogenin</td>
<td>empty vector</td>
<td>TsUBE2L3-HA</td>
</tr>
<tr>
<td>IB: MHC</td>
<td>empty vector</td>
<td>TsUBE2L3-HA</td>
</tr>
</tbody>
</table>

Figure 5.11 Analysis of the effect of TsUBE2L3-HA on muscle differentiation markers in C2C12 myoblasts and myotubes

Empty vector and TsUBE2L3-HA myoblast and myotube cell lines were induced using doxycycline for 0, 18, 22, 24 and 48 hours before analysis by immuno-blot (IB) using antibodies raised against muscle differentiation markers: myogenin and MHC (myosin heavy chain). A. Myoblast cell lines were analysed, and (B) myotube cell lines were analysed. All samples were also analysed using anti-tubulin antibodies as a protein loading control.
5.7. The effect of TsUBE2L3-HA on protein synthesis in C2C12 skeletal muscle cells

Radiolabelled proteins can be detected at very low amounts. The quantification of the radioactivity of radiolabelled proteins accurately represents relative protein amount and, when measured over time, rates of protein synthesis. To quantify the difference observed in protein amounts between empty vector myotubes and TsUBE2L3-HA myotubes, both cell lines were counted, seeded at the same density and grown to confluency. They were then differentiated for 3 days, before being induced using doxycycline. At 20 hours post-induction, the cell culture media was replaced with cysteine and methionine-free media for 1 hour. Cysteine and methionine amino acids that had been synthesised using the radioactive isotope of sulphur, $^{35}$S, were then added to the media for 3 hours. Cells were harvested (24 hours post-induction) and lysed in equal volumes of lysis buffer and aliquoted in triplicate for each sample. The $^{35}$S radiolabelled protein amounts were analysed by measuring scintillation counts per minute (CPM). The experiment was repeated so that 3 biological replicates were carried out.

The amount of $^{35}$S radiolabelled cysteine and methionine incorporated into proteins during protein synthesis in the 3-hour incubation time was compared between the 2 cell lines (Figure 5.12 A). The amount of $^{35}$S radiolabelled cysteine and methionine incorporated into the empty vector cells was consistently lower than in the TsUBE2L3 cells. The mean scintillation count of all 3 biological replicates was found to be significant by a student's t-test (Figure 5.12 B). Data suggests that protein synthesis is significantly increased in C2C12 myotubes expressing the *T. spiralis* E2, TsUBE2L3-HA. This experiment however, does not account for proliferating myoblasts. A differentiated skeletal muscle cell culture, is composed of a mixed population of myotubes, myoblasts and satellite cells. The life-span of the nuclei of myotubes is finite, and as they mature they are replenished with new nuclei from surrounding myoblasts that fuse with the myotube. I therefore wanted to investigate whether or not the source of the increased protein expression in the C2C12 cells was due to an increase in proliferation of the surrounding C2C12 myoblasts, rather than an up-regulation of protein synthesis of either the myoblasts and/or the myotubes. To do this, the effect of TsUBE2L3 on a panel of cell cycle factors was analysed.
Figure 5.12 Analysis of the effect of TsUBE2L3-HA on the rate of protein synthesis in C2C12 myotubes

Empty vector and TsUBE2L3-HA myotube cell lines were induced using doxycycline for 21 hours before being labeled using $^{35}$S cysteine and methionine amino acids for a further 3 hours. Cells were then harvested and the radioactivity of the cell lysate was measured as CPM (counts per minute) in triplicate for each protein sample. **A.** Three biological repeats were carried out, and the mean CPM with standard error of the mean (SEM) is represented as error bars. **B.** The mean of the 3 biological replicates was taken (error bars = SEM) showing a 21% increase in the rate of protein synthesis in myotubes expressing TsUBE2L3-HA compared to the empty vector control.
5.8. The effect of TsUBE2L3-HA on C2C12 skeletal muscle cell cycle factors

In human cells, UBE2L3 plays an important role in cell cycle regulation, pushing cells through S to G2 phase, thus promoting proliferation [320]. Terminally differentiated cells such as myotubes exit the cell cycle after a final M phase division and exist in an arrested state known as G0 phase where they no longer divide (Figure 5.13 A). During nurse cell formation in *T. spiralis* infection, terminally differentiated skeletal muscle cells are induced to re-enter the cell cycle [163]. Here, DNA synthesis occurs, indicative of the transition through G1 to S phase. *In vitro*, it was demonstrated that adding the secreted proteins of *T. spiralis* muscle larvae to C2C12 myoblasts in culture up-regulates the expression of the G1-specific cyclin D1 and promotes proliferation [160,161]. I therefore decided to look at the effect of the secreted *T. spiralis* E2 TsUBE2L3-HA on the cell cycle state of C2C12 myoblasts and myotubes as represented by a panel of cell cycle markers. The protein levels of 4 cell cycle factors that are differentially expressed throughout the different stages of the cell cycle were investigated. The aim was to observe how each particular stage might be affected by the *T. spiralis* E2. Throughout the cell cycle, cyclin D1 is up-regulated during G1 phase, cyclin E is required for G1 to S transition, cyclin B1 is required for G2 to M phase transition and the serine 10 residue of the chromatin protein histone H3 becomes phosphorylated during chromosome condensation in M phase [333-336](Figure 5.13 A).

Both myoblast cell lines, one carrying the pLVX TsUBE2L3-HA expression vector DNA and the other carrying the pLVX empty expression vector DNA, were counted and seeded at the same density. Half of each cell line was differentiated into myotubes, the other half was passaged as normal and maintained as undifferentiated myoblasts. Doxycycline was then added to all the cells to induce the expression of TsUBE2L3-HA. Cells were harvested after at 0, 18, 22 and 48 hours post-induction. Cell samples were lysed and the concentrations of the cell lysates were normalised before being analysed by immuno-blots using antibodies raised against cyclin D1, cyclin E, cyclin B1 and phospho-histone 3 (Figure 5.13 B and C). The same samples were analysed by immuno-blots using anti-tubulin antibodies, providing a loading control. Figure 5.13 B and C are representative of 3 biological repeats of these experiments.
Figure 5.13 Analysis of the effect of TsUBE2L3-HA on cell cycle factors in C2C12 myoblasts and myotubes

A. Graph showing variations in cyclin concentrations throughout the cell cycle. Histone 3 (H3) is phosphorylated (pH3) during chromosome condensation in M phase. Empty vector and TsUBE2L3-HA myoblast (B) and myotube (C) cell lines were induced using doxycycline for 0, 18, 22, 24 and 48 hours before analysis by immuno-blot (IB) using antibodies raised against cell cycle factors: cyclin D1, cyclin E, cyclin B1 and pH3. All samples were reacted with anti-tubulin antibodies as a protein loading control.
Cyclin D1 protein levels were unaffected by the presence of TsUBE2L3 in myoblasts and myotubes. In myoblasts, cyclin E appeared as 2 bands, possibly representative of the full length 50 kDa protein and a lower molecular weight isoform, although this would need to be confirmed [337]. Although no difference was observed between the 2 cell lines, at 18 hours post-induction, levels of phospho-histone 3 were specifically up-regulated in myoblasts. This may indicate that the 18-hour time-point captured a view of a synchronous M-phase event of cell division. The peak expression of cyclin B1 levels in myoblasts, observed at 18 hours post induction, mirrored this pattern of pH3 expression. No Phosphorylated histone 3 was detected in myotubes of either the TsUBE2L3-HA cell line or the empty vector cell line. Since myotube cultures mainly contain differentiated cells that have exited the cell cycle, the absence of phospho-histone 3 was unsurprising. Finally, cyclin B1, the G2/M-phase-specific cyclin, was also unaffected in myoblasts, however it was significantly reduced in the presence of TsUBE2L3-HA in myotubes at all time-points. A reduction in cyclin B1 in mammalian cells is associated with G2/M arrest and this is one of the mechanisms by which the tumour suppressor protein p53 induces cell cycle arrest at the G2/M checkpoint [338]. Data therefore suggests a reduction in the proportion of cells in G2/M phase (in the presence of TsUBE2L3) that is myotube-specific. Unexpectedly, an inhibitory effect on cyclin B1 in myotubes was observed at T0, before doxycycline-mediated induction. Despite no anti-HA signal at T0, this effect may have been caused by leaky expression of TsUBE2L3-HA in myotubes as observed earlier by immunofluorescence.

In order to try and verify differences in the cell cycle states between the 2 cell lines, the effect of TsUBE2L3 on C2C12 proliferation was examined.

5.9. The effect of TsUBE2L3-HA on C2C12 skeletal muscle proliferation

Proliferation of C2C12 cells was analysed by monitoring DNA replication using the Click-it EdU Cell Proliferation Assay (Life Technologies). The modified nucleoside base EdU is a thymidine analogue that contains an alkyne modification. EdU is cell permeable and can be incorporated into the DNA (in the place of thymidine) of dividing cells during DNA replication. The alkyne of EdU can form a covalent bond with a fluorophore that contains an azide modification. This method of conjugation of chemically modified biological compounds is a form of "Click chemistry" [339].

Both myoblast cell lines, one carrying the pLVX TsUBE2L3-HA expression vector DNA and the other carrying the pLVX empty expression vector DNA, were counted
and seeded at the same density. Half of each cell line was differentiated into myotubes, the other half was passaged as normal and maintained as undifferentiated myoblasts. Doxycycline was then added to all cells to induce expression of TsUBE2L3-HA. At the 18-hour time-point post-induction, EdU was added to the cells and they were incubated for a further 6 hours before being harvested and processed for immuno-fluorescence analysis (24 hours post-induction). The fluorophore Alexa-647-azide was incubated with the cells as instructed by Life Technologies (Click-iT EdU Cell Proliferation Assays). This was to allow nuclei that had undergone a DNA replication event, and had therefore incorporated EdU into their DNA, to bind to the Alexa Fluor-647-azide fluorophore (red). Cells were then incubated with anti-HA antibodies and Alexa-488 fluorophore-conjugated (green) secondary antibodies. All nuclei of cells were stained with 4',6-diamidino-2-phenylindole (DAPI), a fluorescent compound that binds to A-T rich regions of DNA (Figure 5.14 A). Negative control samples were also prepared: 1. Cells from both cell lines that were stained with the Alexa-488 fluorophore-conjugated secondary antibodies only (without the anti-HA antibodies). 2. Cells from both cell lines that were stained with the Alexa-647-azide only (without EdU incubation). 3. Cells from both cell lines that had not been induced with doxycycline were also processed with all staining combinations and minimal background staining was observed (data not shown). Samples were then visualised by confocal microscopy.

No anti-HA signal was observed in the empty vector myoblast or myotube cells after 24 hours of induction with doxycycline. Anti-HA signal was observed in the TsUBE2L3-HA myoblast and myotube cells after 24 hours of induction with doxycycline. Alexa-647 positive nuclei (red) indicated that DNA replication had occurred during the 6-hour incubation with EdU. Red nuclei were observed in all samples. The presence or absence of red nuclei did not significantly or exclusively correlate with TsUBE2L3-HA expression (Alexa-488 signal, green). In other words, the red nuclei were evenly distributed amongst cells with and without anti-HA signal. However, nuclei that were red, often demonstrated a higher level of HA signal, consistent with the hypothesis that the TsUBE2L3 is recruited to the nucleus for replication. Myotubes are in the G0 phase of the cell cycle and would only replicate their DNA if pushed into the G1/S phase. In all samples, no red nuclei were observed in the myotubes of the empty vector cells or the TsUBE2L3 cells (red nuclei were only observed in myoblasts), confirming that TsUBE2L3 was not able to induce DNA replication.
Six different areas of the empty vector myoblast sample and 6 different areas of the TsUBE2L3-HA myoblast sample (at 63x magnification) were chosen at random and images were taken. The total number of DAPI-stained nuclei per area (same for each image) were counted, and then the number of Alexa-647-stained (EdU positive) nuclei per area were counted. The mean number of DAPI-positive nuclei and the mean number of Alexa-647-positive nuclei in each sample is presented in Figure 5.14 B and C. A small reduction in the total number of cells counted was observed in the cells expressing TsUBE2L3-HA compared to the empty vector control. The percentage of Alexa-647-positive nuclei out of the number of DAPI-positive nuclei was calculated. A small reduction in the proportion of Alexa-647 nuclei was observed in the cells expressing TsUBE2L3. Data was not calculated to be significant when analysed by a student’s t test and more images would have to be counted before a conclusion could be drawn.

To try and verify and more accurately quantify any possible difference in the frequency of DNA replication in dividing C2C12 myoblasts in the presence of the T. spiralis E2, flow cytometry analyses were carried out. Flow cytometry is the process by which single cells are counted and each cell is analysed for its fluorescent profile. Molecules of interest within a cell or on the surface of a cell are bound to fluorescent markers that can be detected and quantified. In order to compare signals between single cells, they must be of a similar (uniform) size and granularity. Differentiated myotubes are long thin multinucleated, asymmetrical cells that vary in shape and size. They therefore cannot be used for flow cytometry. The effect of TsUBE2L3 on proliferation could therefore only be investigated in C2C12 myoblast cell lines.

Both myoblast cell lines, one carrying the pLVX TsUBE2L3-HA expression vector DNA and the other carrying the pLVX empty expression vector DNA, were counted and seeded at the same density. Doxycycline was then added to all the cells to induce expression of TsUBE2L3-HA. Six hours post-induction, EdU was added to the cells and they were incubated for a further 18 hours before being harvested and processed for flow cytometry analysis (24 hours post-induction). To observe any non-specific binding of the Alexa-647-azide fluorophore to the cells, samples that were not incubated with EdU were also prepared as a control. Samples were then incubated with the Alexa-647-azide fluorophore. Cells were also prepared that had been processed for flow cytometry but not incubated with EdU or Alexa-647 azide as a guide for cell size and granularity for flow cytometry. Once collected, flow cytometry data was analysed using FlowJo software.
Figure 5.14 A

Analysis of the effect of TsUBE2L3-HA on DNA replication in C2C12 cells

<table>
<thead>
<tr>
<th>A</th>
<th>+ DOX 24 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF</td>
<td><img src="image" alt="BF" /></td>
</tr>
<tr>
<td>DAPI</td>
<td><img src="image" alt="DAPI" /></td>
</tr>
<tr>
<td>HA</td>
<td><img src="image" alt="HA" /></td>
</tr>
<tr>
<td>EdU</td>
<td><img src="image" alt="EdU" /></td>
</tr>
<tr>
<td>overlay</td>
<td><img src="image" alt="overlay" /></td>
</tr>
<tr>
<td>overlay + BF</td>
<td><img src="image" alt="overlay + BF" /></td>
</tr>
</tbody>
</table>

myoblast empty vector  myoblast TsUBE2L3-HA  myotube empty vector  myotube TsUBE2L3

Figure 5.14 A. Immuno-fluorescence analysis of the effect of TsUBE2L3-HA on DNA replication in C2C12 myoblasts and myotubes

A. Immuno-fluorescence of C2C12 myoblast and myotube empty vector and TsUBE2L3-HA cell lines after 24 hours of doxycycline induction. Cells were incubated with EdU for 6 hours before harvesting and processing for IFA and Click chemistry. Cells were incubated with Alexa-647-alkyne, DAPI to stain the nuclei, and anti-HA antibodies with Alexa-488 secondary antibodies. Figure shows the Brightfield (BF) the DAPI signal (blue) the Alexa-647 signal (red), the Alexa-488 signal (green), the overlay and the overlay plus the BF image.
Figure 5.14 B and C

Analysis of the effect of TsUBE2L3-HA on DNA replication in C2C12 cells

B. The number of blue (DAPI-positive) and red (Alexa-647-positive, representing DNA replication) nuclei were counted in 6x images (63x magnification). The mean number of nuclei is presented with standard error of the mean (SEM) as error bars. C. The number of DAPI-positive cells was taken as the total number of cells (100%) and the proportion of Alexa-647-positive cells was calculated for each cell line. The mean percentage with SEM as error bars is presented.
Figure 5.15 A and B show the Alexa-647 signal of the empty vector and TsUBE2L3 myoblasts presented as flow histograms. In the samples that had been incubated with Alexa-647 but not EdU, a small amount of non-specific binding by the fluorophore was observed in both cell lines. This was represented as slight increase in the percentage of cells with a positive Alexa-647 signal. EdU incorporation during DNA synthesis was observed as a significant increase in the percentage of cells with a positive Alexa-647 signal. In total, 90.04% of the empty vector cells had an Alexa-647 positive signal and 88.33% of the TsUBE2L3 cells had an Alexa-647 positive signal (Figure 5.15 C). The signal however was split into 3 distinct peaks. Upon further analysis of each individual peak, it was observed that the fluorescence intensity of peak 2 was almost double that of peak 1, and the fluorescence intensity of peak 3 was almost double that of peak 2 Figure 5.15 E. There were 3 different cell populations, each representing an exponential increase in the incorporation of EdU. The peaks may therefore represent the number of DNA replication rounds and therefore the number of cell cycle events through which an individual cell has passed in the 18-hour (EdU) incubation period. Cells in peak 1 had undergone 1 round of the cell cycle, cells in peak 2 had undergone 2 rounds of the cell cycle and cells in peak 3 had undergone 3 rounds of the cell cycle. There was not however, a significant difference between the empty vector cells and the TsUBE2L3 cells, neither in the percentage of cells in each peak, nor in the fluorescence intensity of each peak (Figure 5.15 C and D). This data, along with the immuno-fluorescence analysis suggests that despite having an inhibitory effect on cyclin B1, TsUBE2L3 does not significantly affect DNA replication in C2C12 myoblasts, and does not induce DNA replication in C2C12 myotubes.

As mentioned previously, p53 plays a pivotal role in the terminal differentiation of skeletal muscle cells [132,340]. P53 also plays a well-characterised role in cell cycle arrest, and at the G2/M checkpoint down-regulates cyclin B1 transcription to this end [338]. The mammalian E2, UBE2L3, in collaboration with the E3, E6-AP (E6-Associated Protein), polyubiquitinates p53 and facilitates the proteasomal degradation of the protein thus down-regulating p53 at the protein level [66,341,342]. Experiments were therefore designed to analyze the effect of the *T. spiralis* UBE2L3 on p53 in C2C12 cells.
Figure 5.15 Analysis of the effect of TsUBE2L3-HA on DNA replication of C2C12 cells

C2C12 empty vector and TsUBE2L3-HA myoblasts were induced with doxycycline for 24 hours and incubated for 18 hours with EdU before being processed for Click chemistry using Alexa-647-azide. Flow histogram showing a shift in florescence intensity at 640-670 nm representing Alexa-647-positive nuclei in (A) empty vector cells and (B) TsUBE2L3-HA cells. Cells were gated on a forward and side scatter pattern that appeared typical for myoblasts. Control cells were not incubated with EdU or Alexa-647. Baseline signal was taken from cells that were not incubated with EdU but were treated with Alexa-647-azide. C. The flow histograms for empty vector cells and TsUBE2L3-HA cells incubated with EdU and Alexa-647-azide were overlaid. D. The percentage of the maximum number of cells with a particular fluorescence intensity was determined for three distinct peaks within the main peak. E. The fluorescence intensity of each peak (1,2 and 3) was calculated for each cell line.
5.10. The effect of TsUBE2L3-HA on p53 in C2C12 skeletal muscle cells

In healthy cells, p53-induced cell cycle arrest allows DNA damage or abnormality to be repaired during a cell cycle check point before the continuation of the cell cycle. If the damage irreparable, p53 induces programmed death (apoptosis) of the cell, preventing its replication and thus preventing the inheritance of the damage or abnormality. Mutations in the p53 gene are found in various diseases, many of which are cancers and cell cycle-related disorders [343-345].

The effect of TsUBE2L3-HA on wild-type p53 protein levels in C2C12 cells was examined. The TsUBE2L3-HA and empty vector cell lines were seeded at equal densities. Half of each cell line was grown to confluency and differentiated and the other half was passaged and allowed to remain as myoblasts. After 3 days of differentiation TsUBE2L3-HA expression was induced with doxycycline. Cells were harvested after at 0, 22 and 48 hours post-induction. Cells were lysed, the protein concentrations of the cell lysates were normalised and the samples were analysed by immuno-blot using anti-p53 antibodies (Figure 5.16 A). The same samples were reacted with anti-tubulin antibodies as a protein loading control. Figure 5.16 is representative of 3 biological repeats of this experiment. Very low levels of p53 protein were observed in C2C12 myoblasts throughout the time-course. This is expected since the cells have a rapid rate of turnover and few cells will be growth-arrested. After 48 hours, more p53 was observed, consistent with the fact that as the cells become more confluent over time, they begin to spontaneously differentiate. No differences were observed between the myoblasts expressing TsUBE2L3 and those carrying the empty vector.

Compared to the myoblasts, the overall p53 expression was significantly up-regulated in the differentiated myotubes. In the presence of TsUBE2L3-HA however, p53 protein was significantly reduced when compared to the empty vector control. This down-regulation was observed at T0, suggesting some leaky expression of TsUBE2L3-HA before induction with doxycycline. After induction, the expression of p53 continued to be significantly reduced in the presence of TsUBE2L3. Minimal difference in the reactivity with anti-tubulin antibodies was observed between samples, suggesting equal protein loading. These results are representative of more than 3 biological replicates of this experiment.

Since immuno-blot analyses are not accurately quantitative, a radio-immuno-precipitation of p53 was carried out to verify and quantify the reduction of p53 protein...
in myotubes expressing TsUBE2L3-HA. Both cell lines were seeded at the same density and grown to confluency. They were then differentiated for 3 days, before being induced using doxycycline. At 20 hours post-induction, the cell culture media was replaced with cysteine and methionine-free media for 1 hour. $^{35}$S radiolabelled cysteine and methionine amino acids and the proteasome inhibitor MG132 were then added to the media for 3 hours. MG132 blocks the core of the proteasome, causing a build up of ubiquitinated proteins that would otherwise be degraded. Since p53 is highly regulated by ubiquitination and proteasomal degradation, MG132 was used to allow the accumulation of all ubiquitinated forms of p53 as well as native p53, thus increasing the p53 signal. Cells were harvested (24 hours post-induction), lysed and total $^{35}$S radiolabelled protein concentrations were analysed by measuring scintillation counts per minute (CPM). Radiolabelled lysate concentrations were normalised based on their $^{35}$S CPM. Protein G suspended anti-p53 antibodies were then used to immuno-precipitate p53 from each of the lysates. The input (lysate before immuno-precipitation), the immuno-precipitated proteins and the unbound (proteins that did not immuno-precipitate), were all separated by SDS-PAGE. Proteins were visualised by auto-radiography (Figure 5.16 B).

Figure 5.16 B represents a short exposure of the auto-radiogram showing that the amounts of protein loaded into (input) and recovered from (unbound) the immuno-precipitation were well normalised between the 2 cell lines. A protein corresponding to the expected size of native p53 (53 kDa, Figure 5.16 B red arrow) was immuno-precipitated from the empty vector lysate. The amount of this protein immuno-precipitated from the TsUBE2L3-HA lysate was significantly reduced. A slightly larger protein was also immuno-precipitated in equal measures from both cell line lysates, at 55 kDa. If this band is p53-specific, it may represent mono-ubiquitinated p53. Native p53 is actually 43 kDa in size, but runs at 53 kDa when separated by SDS-PAGE. Therefore, mono-Ub p53 runs very close to native p53. Alternatively, this band may represent a modified form of p53 other than ubiquitinated p53, because as well as by ubiquitin, p53 is highly regulated by phosphorylation and acetylation [346].

A very large protein of approximately 250 kDa was also immuno-precipitated from the empty vector lysate that was present but greatly reduced in the TsUBE2L3-HA lysate sample. When the auto-radiogram was over-exposed, 2 more bands were observed to be immuno-precipitated only in the empty vector sample (Figure 5.16 C, black arrows). These were between 15 and 35 kDa. The experiment was repeated twice and the same pattern of bands was observed (data not shown). The area of the 3 autoradiograms corresponding to the size of native p53 (53 kDa) were measured
for particle gray value (number of pixels with gray value) using gel analysis software (Fiji). This allowed for the quantification of the intensity of the band in each immuno-precipitation sample (Figure 5.16 C). By calculating the mean of the 3 measurements from each experiment, native p53 immuno-precipitated from C2C12 myotubes expressing TsUBE2L3-HA was significantly less than that immuno-precipitated from cells carrying the empty vector DNA. Results strongly agreed with the immuno-blot data using anti-p53 antibodies presented earlier.
Figure 5.16 Analysis of the effect of TsUBE2L3-HA on p53 in C2C12 cells

**A.** Immuno-blot analysis of C2C12 empty vector and TsUBE2L3-HA cell lines after 0, 22 and 48 hours of induction using doxycycline using anti-p53 antibodies. Myotubes were also analysed for TsUBE2L3-HA expression using anti-HA antibodies, and all samples were analysed using anti-tubulin antibodies as a protein loading control. **B.** Myotube C2C12 empty vector and TsUBE2L3-HA cells were induced for 24 hours, labeled with ^35^S cysteine and methionine amino acids and treated with the proteasome inhibitor MG132. Anti-p53 antibodies were used to immuno-precipitate p53 from concentration-normalised cell lysates. The input, immuno-precipitation products and the proteins that did not immuno-precipitate (unbound) were separated by SDS-PAGE and analysed by autoradiography. **C.** The autoradiograph was overexposed to reveal more signals affected by TsUBE2L3-HA. Black arrows indicate protein bands affected by TsUBE2L3-HA. The red arrow (1) indicates the expected size of mono-ubiquitinated p53 and (2) the expected size of native p53. The experiment was repeated. **D.** The gray value (pixels) of the band representing native p53 was measured in the empty vector and TsUBE2L3-HA immuno-precipitation products using Fiji image analysis software. The mean of 3 gray values from biological replicates are presented with standard error of the mean represented as error bars. A 34% reduction in native p53 was observed in the presence of TsUBE2L3-HA.
The reduction of a protein observed may have arisen from a post-translational inhibitory effect, for example ubiquitin-proteasome-mediated degradation, or from an inhibitory effect at the transcriptional level (mRNA). To measure levels of p53 mRNA, a quantitative reverse transcriptase PCR (qRT-PCR) was carried out. During qRT-PCR the amplification by PCR of a reverse transcribed transcript is quantified. The levels of specific mRNA transcripts can therefore be measured and compared between cell lines. The transcript level of a ‘housekeeping’ gene that is unaffected by the experiment variable is measured as a baseline level for transcription.

Three C2C12 myotube cell lines were cultured: the empty vector, TsUBE2L3-HA and an additional inducible cell line that was set up in the same way to express another *T. spiralis* secreted protein called Ts6496. This protein was identified by LC/MS/MS of *T. spiralis* SP as an abundant uncharacterised (*T. spiralis*-specific) secreted protein (supplementary appendix 4). This cell line was assayed as an additional control to ensure that the effect on p53 was not simple an artefact of the over-expression of any *T. spiralis* protein in muscle cells, since the empty vector does not contain a gene for protein expression. Total RNA was extracted from each myotube cell line after the induction of TsUBE2L3 expression using doxycycline for 0, 2, 4, 8, 12 and 24 hours. RNA was used to synthesise cDNA using the Takara reverse transcriptase kit with oligo(dT)₁₂₋₁₈ and random primers, according to the manufacturer’s instructions (Clontech). Transcript levels of p53 (target gene), myosin heavy chain (MHC) and the ‘housekeeping gene’ glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were analysed (Figure 5.11). MHC was chosen as an alternative (additional) control to GAPDH as no changes in MHC protein levels were observed between the 2 cell lines during immuno-blot analysis between 0 and 22 hours post-induction. For consistency, the gene that encodes the MHC 2x protein (analysed previously by immuno-blot), MYH1 was analysed.

Each sample was quantitatively analysed by real-time PCR (in triplicate) using specific TaqMan Gene Expression Assays from Life Technologies. Each sample was analysed in triplicate and the mean CT (cycle threshold) value was calculated with standard error (SEM). The raw data suggested the expression of GAPDH was not affected by TsUBE2L3-HA or Ts6496 expression and the expression of all 3 genes was found to be relatively efficient. This allowed for the direct comparison of the expression of MHC and p53 to GAPDH using the Comparative CT Method (ΔΔCT) as described in the Applied Biosystems ABI PRISM User Bulletin #2. Therefore, the expression was calculated relative to GAPDH, and the fold change of each gene at
each time-point was calculated relative to the empty vector control cell line sample taken at T0 (Figure 5.17).

The expression of MHC and GAPDH in C2C12 myoblasts was not significantly affected by the expression of TsUBE2L3-HA or the expression of Ts6496 at any post-induction time-points evaluated. A very small reduction in MHC transcript was observed in the TsUBE2L3 samples after 12 and 24 hours. This was in line with a slight reduction observed earlier by immuno-blot, however this experiment would need to be repeated with an extended time-course to confirm an effect of TsUBE3L3 on MHC. The expression of Ts6496 or TsUBE2L3-HA in C2C12 myotubes did not significantly affect p53 transcript levels at any post-induction time-point that was evaluated. This would suggest that the down-regulation of p53 by TsUBE2L3 does not occur at the transcript level, rather at the protein level. This reinforced the hypothesis that the T. spiralis E2, like its mammalian orthologue UBE2L3 (UbcH7), can mediate the ubiquitination and subsequent proteasomal degradation of p53.
Figure 5.17 qRT-PCR analysis of the effect of TsUBE2L3-HA on p53 and MHC transcription

A  p53 transcription

B  MHC transcription

Figure 5.17 Analysis of the effect of TsUBE2L3-HA on the transcription of p53 in C2C12 myotubes

Qualitative real-time PCR (qRT-PCR) was carried out on myotube cDNA to assess the effect of TsUBE2L3-HA on (A) p53 transcription and (B) myosin heavy chain (MHC) transcription normalised to the transcription of the housekeeping gene GAPDH over a 24-hour time-course post induction using doxycycline. This was compared to the effect in empty vector cells and in cells expressing another T. spiralis unrelated protein, named 6496. The fold change in transcription relative to the empty vector time 0 sample is presented. Measurements were taken in triplicate and error bars indicate standard error of the mean.
Summary of results

5.1. The secretion of TsUBE2L3 by *T. spiralis* muscle larvae was confirmed by immuno-blot. The secretion of the putative *T. spiralis* E3 enzyme, TsRNF13 could not be confirmed.

5.2. RACE-PCR confirmed the full coding sequence for TsUBE2L3. The coding sequence for TsRNF13 could not be confirmed.

5.5. The C2C12 mouse skeletal muscle cell line was successfully cultured and differentiated into myotubes.

5.4. Two lentiviral vector-mediated expression systems were established that allowed expression and detection of TsUBE2L3-HA in C2C12 myotubes.

5.5. The morphology of C2C12 myotubes was affected by the presence of TsUBE2L3. More enlarged myoblasts and laterally fused myotubes were observed.

5.6. TsUBE2L3-HA expression in C2C12 myotubes did not significantly affect the skeletal muscle differentiation markers myogenin and myosin heavy chain (MHC).

5.7. Protein synthesis in C2C12 myotubes was up-regulated by TsUBE2L3.

5.8. TsUBE2L3-HA expression in C2C12 myotubes did not affect the cell cycle factors cyclin D1, cyclin E or phospho-histone 3, but did cause a significant reduction in cyclin B1.

5.9. TsUBE2L3-HA expression in C2C12 myotubes did not affect the rate of proliferation as measured by DNA replication in dividing C2C12 myoblasts.

5.10. TsUBE2L3-HA expression in C2C12 myotubes significantly down-regulated the tumour suppressor protein p53 at the protein level, but not at the transcript level.
Discussion

Experiments in chapter 3 showed that the muscle stage larvae of the parasite *T. spiralis* express components of a functional ubiquitin proteasome system, components of which have been highly conserved throughout evolution. To ascertain if these proteins might be specifically host-targeted, chapter 4 investigated whether or not ubiquitin pathway enzymes are secreted by *T. spiralis* muscle larvae. Experiments could not detect deubiquitinating or deNeddylating activity in the secreted proteins of *T. spiralis* muscle larvae, however ubiquitin conjugation (E2) and ligation (E3) activity was detected. Proteomic analyses of all of the secreted proteins identified a single putative *T. spiralis* E2 enzyme and a single putative *T. spiralis* E3 enzyme. The purpose of experiments presented in chapter 5 was to confirm the identity of the proteins responsible for the ubiquitin conjugation activity, and to characterise their possible function in host-parasite interactions during *T. spiralis* infection of muscle cells.

Immuno-blot evidence confirmed the identity of the secreted E2 enzyme as the *T. spiralis* UBE2L3 orthologue. Given 62.3% identity between the *T. spiralis* and human protein sequences, it is not surprising that an anti-human antibody was able to cross-react with proteins of an expected size in both *T. spiralis* lysate and the secreted proteins. Bands of reactivity however, were multiple in both samples. The mammalian UBE2L3 can be mono and polyubiquitinated, and this may explain some of the additional bands observed [320]. To confirm this, a pan deubiquitinating enzyme such as the murine cytomegalovirus M48, that removes all ubiquitin modifications from proteins, could be added to the lysate or secreted proteins before immuno-blot with anti-UBE2L3 antibodies [94]. Indeed adding the proteasome inhibitor MG132 to C2C12 myotubes expressing TsUBE2L3-HA, caused an enrichment of a larger form of HA-tagged protein that would correspond to the size of monoubiquitinated TsUBE2L3 (appendix 4).

The human protein has 3 isoform coding sequences (GI: 373432685, 17.9 kDa, GI: 373432681, 14.1 kDa, GI: 373432683, 24 kDa) [347,348]. The most abundantly expressed form is the 18 kDa isoform 1 [324]. The most predominant anti-UBE2L3-reactive band in the secreted proteins corresponded to approximately 18 kDa. The full sequence for TsUBE2L3 was confirmed by RACE-PCR experiments, and this sequence was found to translate to a protein of 18 kDa. This sequence was also identified by AUGUSTUS gene prediction software, but only when using gene
arrangement information based on the genome of the filarial nematode *Brugia malayi* as a template. During experiments carried out in chapters 3 and 4, evidence of the mis-annotation of the *T. spiralis* draft genome was observed multiple times. This suggests that the draft annotation is not a reliable source of data for gene prediction in *T. spiralis* and may explain why another nematode, *Brugia malayi*, was required for a gene prediction that was in agreement with the experimental data. The start codon and the beginning portion of the RACE-PCR-confirmed sequence aligned well with orthologous sequences from multiple other species including the human sequence that has been confirmed experimentally [347]. This sequence was therefore used as the basis for cloning the gene encoding the *T. spiralis* E2 enzyme. The same experiments were unable to confirm the identity and secretion of the putative *T. spiralis* E3 (RNF13 orthologue) and studies on this protein were discontinued.

To try and understand the role the E2 enzyme might play during nurse cell formation, I investigated the effect of TsUBE2L3 on muscle cells. *Trichinella spiralis* infects all mammals indiscriminately and the disease phenotype is the same for all species. I therefore decided to work with mouse rather than human muscle cells as protocols and reagents for the culture and differentiation of C2C12 mouse skeletal muscle cells are well established. During infection, *T. spiralis* newborn larvae invade terminally differentiated myofibres, not immature myoblasts [157]. It was therefore important to express TsUBE2L3 in already differentiated myotubes. To achieve this, the coding sequence of TsUBE2L3-HA was cloned into a lentiviral vector system. The efficiency of transduction from the first vector used, pCSGW, did not exceed 50% and methods affected the health of the myotubes. Instead, an inducible system allowed for the generation of a stable myoblast cell line that could be induced to express the *T. spiralis* E2 after differentiation.

Immunofluorescence analysis showed that the location of TsUBE2L3-HA within myoblasts was both cytoplasmic and nuclear, and even more nuclear in cells that were dividing. Being able to determine, by observational microscopy, that a cell is actively dividing suggests that the cell is already in a later stage of the cell cycle, such as G2 or M. The localisation of TsUBE2L3-HA appeared to be consistently more nuclear in cells that were undergoing a division event at the time of harvest. In mammalian cells, UBE2L3 is involved in cell cycle progression from S phase to G2 [320]. Whitcomb *et al.* show that UBE2L3 is down-regulated in S phase, but becomes up-regulated during G2 and M phase. Taken together, data suggests that the up-regulation and recruitment of TsUBE2L3 to the nucleus may be required for latter stages of the cell cycle. In differentiated (non-dividing) C2C12 myotubes, the majority
of TsUBE2L3-HA was located cytoplasmically, re-enforcing the hypothesis that nuclear recruitment of TsUBE2L3 is a predominant requirement of dividing cells. Due to time constraints, this was not confirmed experimentally. Experiments could be carried out by blocking the cell cycle at specific stages and analysing the effect on the localisation of TsUBE32L3.

Using lentiviral expression systems, variations in the number of DNA integration events and the sites of integration may occur from cell to cell. Levels of regulation of either the modified CMV promoter of the pLVX control vector or the pTight promoter of the pLVX expression vector may also vary. Finally, differentiated C2C12 cultures still contain a mixed population of cells. Myoblasts are present, which can differentiate into myocytes and then myotubes. Because cultures were transduced and selected before differentiation, cell populations may fluctuate between cell lines, and each cell type (myoblast or myotube) may observe different transduction efficiencies or regulate integration or promoters differently. This may explain the variation in TsUBE2L3-HA expression observed by IFA between individual cells of the same line.

C2C12 cells expressing TsUBE2L3-HA appeared morphologically different to cells carrying the empty vector. Differences were subtle and it was difficult to observe patterns or quantify changes. A degree of ‘leaky’ expression of TsUBE2L3-HA was observed in myotubes, but not myoblasts by immunofluorescence but not by immuno-blot. This may explain why some changes in morphology were already observed after the differentiation of the cells, but before their induction with doxycycline (data not shown). The differentiation process itself, however, did not appear to be affected by the leaky TsUBE2L3 expression, and myoblasts fused and formed myotubes, albeit with a slightly modified morphology. Immuno-fluorescence using anti-tubulin antibodies showed that dividing C2C12 myoblasts expressing TsUBE2L3 were sometimes enlarged with enlarged nuclei. Tubulin expression in myoblasts was not affected as measured by IB, however by IFA, in the cases where cells were enlarged, the tubulin network was larger in order to fill the increased cell area. Inside myotubes, the anti-tubulin signal was less intense, suggesting reduced tubulin expression. However this was contradicted by immuno-blot experiments where tubulin levels were equal between cell lines when protein concentrations were normalised. Myotubes expressing TsUBE2L3 appeared as less regular tube shapes, and had more lateral fusions, leading them to often appear wider than myotubes carrying the empty vector. It is worth noting that all structural abnormalities discussed were also observed in the cells carrying the empty vector, but to a much lower
frequency. Morphological effects of TsUBE2L3 were not caused by changes in the levels of the muscle differentiation markers myogenin and myosin heavy chain (MHC) and the cause of the modified morphology remains to be determined.

The next unexpected observation was that TsUBE2L3-HA expression affected the protein concentrations of C2C12 lysates. Each time a cell lysate concentration was measured, the protein concentration was higher in the presence of TsUBE2L3. To confirm and quantify this observation, time-dependent protein labelling was carried out. Results indicated that there was consistently 21% more protein synthesis occurring in cells expressing TsUBE2L3. It is unlikely that the over-expression of the recombinant TsUBE2L3-HA protein itself was the cause of this, since immuno-blot analysis of the HA-tag showed levels of protein comparable to the tubulin protein that was analysed as a loading control (although antibody affinities will vary). It was hypothesised that the differences in protein concentration was due to a difference in the rates of proliferation, leading to differences in population density. Bai et al. showed that adding total *T. spiralis* secreted proteins to C2C12 myoblasts in culture, increased their rate of proliferation and in parallel, increased cyclin D1 and proliferating cell nuclear antigen levels [160,161]. I therefore decided to test the effect of TsUBE2L3 expression on the cell cycle of C2C12 cells.

Differentiated C2C12 myotubes are in G0 phase, having exited the cell cycle after M phase. They are therefore not expected to express high levels of cyclin proteins that are specifically up-regulated during the cell cycle. No effect of TsUBE2L3 was observed on cyclin D1 (G1 phase), cyclin E (G1/S phase) or phospho-histone 3 (chromosome condensation during M phase). Cyclin B1 however, was significantly reduced in the presence of TsUBE2L3. Cyclin B1 is expressed during the G2/M phase and its transcriptional down-regulation by the tumour suppressor protein p53 causes G2 cell cycle arrest [338,349]. Down regulation of cyclin B1 can lead to an inhibition of proliferation and tumour growth [350,351]. In order to investigate if the down-regulation of cyclin B1 in the presence of TsUBE2L3, was able to affect the cell cycle of C2C12 cells, a proliferation assay was carried out. By IFA, no significant differences were observed between cells carrying the empty vector DNA and cells expressing TsUBE2L3-HA. In addition, no significant or exclusive correlation was observed between cells that had undergone a DNA replication event and those that were HA- (TsUBE2L3-HA) positive. In other words, the proportion of cells that had undergone a DNA replication event was no higher in HA-positive cells than in HA-negative cells. As expected, no terminally differentiated myotube cells had replicated their DNA in either cell line. The data was quantified and the proportion of replicating
cells was not significantly different in the presence of TsUBE2L3. Preliminary flow cytometry experiments on myoblasts confirmed this result, showing that the proportion of cycling cells was almost identical in the 2 cell lines, and suggesting that over 18 hours some cells had undergone 3 rounds of replication. To increase the confidence of this data, flow cytometry analyses would need to be repeated using optimised parameters. It would be preferable to synchronise the myoblast cell cycle by drug treatment first, and use different DNA replication labels such as propidium iodide (PI). Rather than labelling only newly synthesised strands of DNA, which is the principle of thymidine analogues such as EdU, PI is an intercalating agent that labels both DNA strands during replication. Flow cytometry-based PI assays therefore allow for a more accurate quantification of the number of replication events of a cell. I would also measure the cell cycle over a shorter period of time, or during a time-course, to try and capture only single cell cycle events, and to more accurately compare DNA replication between the 2 cell lines. Due to their shape, asymmetry and the fact that they are multinucleated, flow cytometry analysis could not be carried out using myotubes. To observe changes other than DNA replication events, in the cell cycle state of myotubes, another method of cell cycle phase detection would have to be used. Immuno-blot or microarray analyses on a greater panel of cell cycle markers would be informative.

Data suggests that the increased rate of protein synthesis in the presence of TsUBE2L3, is therefore not due to an increased rate of proliferation of myoblasts in the culture. However, this data does not represent what is happening in cell cycle arrested myotubes in the culture. Myotubes exist in a mixed population of cells that contains myoblasts. This consistently limits the myotube-specific signal from any assay. It would be ideal if myotubes could be efficiently isolated (en mass), so that signals would not be masked by those from the myoblast population of a culture. Preliminary efforts using the compound cytosine arabinoside (AraC, which kills replicating cells) were not effective because myotubes began to die before the experiment was finished (data not shown). It makes sense that the health of the myotube depends on the presence of the surrounding satellite cells and myoblasts, as was also discussed by Hinterberger et al. [352].

Another cell cycle factor, p53, has been previously shown to be a specific substrate for ubiquitination of the mammalian UBE2L3 [66,353]. Polyubiquitination of p53 by UBE2L3 leads to its degradation by the proteasome [65]. It was therefore hypothesised the highly conserved T. spiralis UBE2L3 would also be able to ubiquitinate p53. Initial immuno-blot analyses showed that p53 was significantly
down-regulated by TsUBE2L3 in myotubes and this was confirmed by radio-immunoprecipitation, showing a significant p53 reduction in the presence of the *T. spiralis* E2 enzyme. This reduction however, was not caused by a reduction in the transcript levels of p53, which remained largely unaffected by TsUBE2L3-HA expression over a 24-hour period.

It was therefore proposed that the reduction in p53 must occur at the protein level, and may be mediated by ubiquitination and proteasomal degradation. Monoubiquitination of p53 by the p53-specific E3 ligase MDM2, is usually a signal for nuclear export [354,355]. Nuclear accumulation of p53 above a certain level can lead to polyubiquitination by MDM2 and degradation of p53 by the nuclear proteasome [356]. Monoubiquitin-mediated nuclear export and cytoplasmic accumulation of p53 can also lead to cytoplasmic polyubiquitination of p53, which in turn targets the protein for degradation by the cytoplasmic proteasome [357]. Further analysis of ubiquitinated forms of p53 in the presence and absence of the *T. spiralis* E2 would allow clarification of the mechanism of TsUBE2L3-mediated p53 depression.

During the G2/M checkpoint, p53 is activated in response to DNA damage, leading to a transcriptional reduction in cyclin B1 that contributes to G2 arrest [338,349]. TsUBE2L3 expression also led to a reduction of cyclin B1 in myotubes. A TsUBE2L3-mediated reduction in p53 would be expected to lead to a rise in cyclin B1, which is the opposite of what was observed. Results therefore suggested that the effect of TsUBE2L3 on p53 is not linked to the effect on cyclin B1. The cyclinB1-cdk1 complex is essential for mitosis, and so an inhibition of cyclin B1 can lead to cell cycle arrest before M phase [358]. When it is no longer needed after mitosis, the anaphase promoting complex E3 ligase is responsible for the ubiquitination of cyclin B1, leading to its degradation. The question therefore remains: how can TsUBE2L3 down-regulate the tumour suppressor p53 and at the same time the promoter of proliferation cyclin B1. It appears likely that the mechanism of cyclin B1 suppression by TsUBE2L3 in C2C12 myotubes is a p53-independent process. Considering the complexity of cell signalling, it is also possible that the effect of the *T. spiralis* E2 on both proteins is indirect. For example, TsUBE2L3 may specifically ubiquitinate other proteins, having an indirect downstream effect on p53 or cyclin B1.


Chapter 6: Final Discussion

Significance of study

The disruption of the ubiquitin pathway is associated with various human diseases and has been directly implicated in the process of infection by viruses and bacteria [93, 96, 97, 99, 100, 228, 272-275, 359]. Since these prokaryotic organisms do not express their own ubiquitin pathway, viral and bacterial-derived ubiquitin pathway proteins are thought to be specifically host targeted and indeed this has been previously demonstrated [97, 98]. However, the importance of the ubiquitin pathway in infection extends beyond viruses and bacteria. Considering the importance of this pathway in eukaryotic cell biology, and its conservation throughout evolution, this study began with the hypothesis that eukaryotic parasites also manipulate the ubiquitin pathway to their advantage during infection. Because eukaryotic parasites express their own ubiquitin pathway, for this study the proposed contribution of this pathway to the process of parasitic infection was divided into 2 categories:

1. The role of the parasite ubiquitin pathway in parasite biology.

2. The role of the ubiquitin pathway in direct host-parasite interactions.

The purpose of studying the parasite ubiquitin pathway during infection (thus addressing point 1) is two-fold. Firstly research can shed light on the evolution of the function of particular ubiquitin pathway components from non-parasitic to parasitic organisms and from lower eukaryotes to higher eukaryotes. This can help us understand the role of these components in important biological processes and how they are involved when these processes are disrupted, for example in disease. Secondly, it can help us identify proteins that are essential for parasite survival, thus identifying potential drug targets for the development of therapeutics aimed at blocking infection [112, 230, 360, 361]. The importance of the ubiquitin pathway in a vast number of cellular processes means that certain ubiquitin pathway components must be essential for parasite survival during all life cycle stages. If so, targeting these proteins would lead to broad treatments capable of disrupting all stages during infection. In addition, the high conservation of ubiquitin pathway proteins amongst different species of the same genus, could result in therapeutics with even farther reaching efficacy.
The purpose of studying the role of the ubiquitin pathway in direct host-parasite interactions (thus addressing point 2) is to identify ubiquitin pathway components that are parasite-derived but specifically host-targeted. In the case of helminths, these will likely be the proteins that come into direct contact with the host during infection; the parasite surface or secreted proteins [2,115]. If these proteins are found to have antigenic potency, they may serve as novel candidates for the development of vaccines aimed at preventing infection. In addition, studying helminth-derived host-targeted ubiquitin components may identify proteins that have an interesting modulatory effect on the biology of the host cell. These studies can be used as model systems to help us to dissect the mechanisms underlying the particular biology of the host cell that is affected by the helminth protein. This will contribute to our greater understanding of mammalian cell biology, its regulation and deregulation and provide ideas on how it may be deliberately manipulated for therapeutic benefit.

During the infective stages, helminths require their host for survival. They therefore do not usually intend to kill the host, rather to evade immune recognition by down-regulating host immunity until transmission occurs. These helminth-induced immunomodulatory effects can therefore, in some instances, be beneficial to the host as well. These beneficial effects predominantly stem from the enhanced and often sustained anti-inflammatory T helper 2 (T\textsubscript{H}2) immune response that gives rise to the release of protective cytokines [362]. This protective immune response can reverse elevated and pathogenic immune responses that are characteristic of allergy and autoimmune disease [363,364]. Various studies now show that whole parasitic nematodes can be administered to treat ulcerative colitis, asthma, rheumatoid arthritis, multiple sclerosis and type 1 diabetes amongst others [365-369]. Furthermore, human clinical trials have shown that infection with whole parasitic nematodes can provide relief and even remission from autoimmune disorders (reviewed by [370] and [371]). For example, ingesting pig whipworm (\textit{Trichuris suis}) eggs can significantly reduce the symptoms of Chrohn’s disease [372,373] and the symptoms of allergic rhinitis [374]. All of these examples use whole parasites to stimulate a therapeutic effect. Instead, this study aimed to address the idea of isolated helminth effectors, with a view to the exploitation of helminth-derived proteins, rather than whole parasites, for therapeutic purposes. In contrast to the widely studied immuno-modulation by intestinal helminths, the purpose of this project was to look at muscle cell modulation by proteins of the intramuscular parasite \textit{T. spiralis}, with a particular focus on the worm's ubiquitin pathway.
To date, no parasite-derived ubiquitin pathway proteins had been determined to be specifically host targeted. The main reason for this is technical. For many parasites, especially those that exist inside the host cell, it is extremely challenging to uncouple the parasite’s own ubiquitin pathway proteins from those that are host targeted. Helminths provide a solution for this problem since they secrete an abundance of proteins that modulate their immediate niches in the host. In some lucky cases, such as with *T. spiralis*, these secretions can be isolated from parasites in culture at high purity.

I initially investigated the role of the *T. spiralis* ubiquitin pathway, focusing on the function of the deubiquitinating enzyme TsUCH37. Following this, I investigated the role of the ubiquitin pathway in direct host-parasite interactions during infection by *T. spiralis*, with a particular interest in identifying parasite-derived ubiquitin pathway modulators and dissecting their mechanism of modulation. Because *T. spiralis* doesn’t kill the host muscle cell, this worm was ideal for using as a model to study the mechanisms underlying the reprogramming of muscle cells.

**Key findings and future work**

The first significant finding was that the function of the *T. spiralis* DUB, TsUCH37, has been highly conserved throughout evolution, from nematodes to humans [1]. In mammals, this ubiquitin hydrolase enzyme is the only ubiquitin-C-terminal hydrolase (UCH) domain DUB to be associated with the proteasome [32,33,233]. It therefore plays a unique and crucial role in the removal of Ub from substrate proteins that have been targeted to the proteasome for degradation. The orthologue of this protein has been shown to be essential for the survival of mice and Drosophila [36,264]. Although I was unable to show if TsUCH37 alone is essential for the survival of *T. spiralis* in culture, this study does present evidence that the inhibition of the UCH domain family of enzymes leads to the death of cultured larvae. This data contributes to on-going work that sets precedence for the further investigation of UCH37 as a drug target in disease. This includes work by other members of our group that are studying the human malaria parasite (*Plasmodium falciparum*) orthologue, PfUCH54 [102] whose knockout in blood-stage *P. falciparum*, is lethal (unpublished data, Sachiko Miyata). It has also been previously shown that despite their level of conservation, UCH family enzymes can be individually targeted by drugs [246-248,267]. Further analyses of the crystal structure of TsUCH37, solved in collaboration with Chittaranjan Das’ laboratory at Purdue University, will contribute to
our understanding of the structural differences between UCH DUBs and their orthologues that may be exploited in drug discovery [3].

Having proven that biological tools based on human ubiquitin could be successfully used to study the ubiquitin pathway of *T. spiralis*, I wanted to use these tools to study the *T. spiralis* components that are directly involved in host-parasite interaction during infection, the secreted proteins. Specifically, due to their fascinating ability to remodel terminally differentiated skeletal muscle cells, I wanted to investigate the proteins involved in nurse cell formation during the invasion of muscle tissue by *T. spiralis* larvae. The *T. spiralis*-induced remodelling of the host cell involves changes in host cell transcription, differentiation and the cell cycle, biological processes for which the ubiquitin pathway is a major player [5,52]. It was therefore hypothesised that *T. spiralis* may secrete ubiquitin pathway components into the muscle cell that target and modulate the host ubiquitin pathway, contributing to nurse cell formation.

Initially, it was found that secreted proteins of *T. spiralis* muscle larvae possess *in vitro* ubiquitin conjugation (E2) and ubiquitin ligation (E3) activity that was only activated when supplied with the human recombinant E1 and human E3 and E2 partners, namely the E2, UBE2L3 and the E3, parkin. Efforts were then made to identify the proteins responsible for this activity, and after many methodologies were unsuccessfully explored, a proteomic analysis of all the *T. spiralis* secreted proteins was carried out. During this process, an entire putative secretome for *T. spiralis* was compiled (supplementary appendix 4). Of the 333 annotated proteins identified, some are redundant sequences. Data therefore suggests that in total there are less than 333 total proteins secreted by *T. spiralis* muscle larvae. Some of these proteins contained conserved domains and could be categorised into the various known processes in which they function in other organisms. Some of these functions are relevant to nurse cell formation, such as inflammation, angiogenesis, cytoskeletal arrangement, collagen production and differentiation and some are also related to processes that are highly regulated by the ubiquitin pathway such as the cell cycle.

During this analysis, only 2 ubiquitin-related enzymes were identified in the *T. spiralis* secreted proteins. Interestingly, the putative *T. spiralis* UBE2L3, was the only E2 identified, the orthologue of the human UBE2L3 that was able to activate the ubiquitination activity of *T. spiralis* secreted proteins *in vitro*. Unfortunately the protein initially thought to be responsible for E3 activity could not be characterised, leaving an interesting avenue open for further investigation. Since the E3 activity was confirmed but not identified using techniques based on known protein orthology,
there is always a chance that the source of the E3 activity may be *T. spiralis* specific. In other words, *T. spiralis* may express a novel domain with ubiquitin ligation activity that is thus far uncharacterised.

It is worth noting that C2C12 myotubes express the mouse UBE2L3 orthologue, which is 68.4% identical in sequence to the *T. spiralis* protein (appendix 5). If *T. spiralis* secrete the E2 into a host cell that already expresses a conserved version of the E2, three hypotheses are viable. 1. that the function of the *T. spiralis* E2 is to collaborate with the host E1 and E3 to ubiquitinate other *T. spiralis* secreted proteins that are not recognised by the host E2, 2. that the increased concentration of UBE2L3 due to the addition of TsUBE2L3 enhances the function of the host E2 that is already present, or 3. that the *T. spiralis* E2 possesses novel functional activity and is able to recognise and ubiquitinate host protein substrates that are not recognised by the host orthologue. Although the evidence presented here for the secretion of TsUBE2L3 by muscle larvae in culture is robust, its secretion into the nurse cell was not confirmed. I designed 2 specific anti-TsUBE2L3 antibodies that were produced by Davids Biotechnology (Germany). I also prepared histological sections of nurse cells in rat muscle tissue for immuno-histochemical analyses [204] (appendix 6). However due to time constraints I was unable to complete these experiments and confirm the secretion of TsUBE2L3 into the nurse cell complex. This would therefore require completion before a role of TsUBE2L3 in nurse cell formation could be assigned. Nonetheless, in the interest of finding parasite-derived muscle cell modulators that may serve a therapeutic purpose, this detail is irrelevant. In other words, *T. spiralis* may express and secrete proteins that are not directly involved in parasite biology, the infection of muscle cells or nurse cell formation, but are nonetheless able to beneficially modulate a muscle cell phenotype. This effect could then be harnessed for its therapeutic potential.

As mentioned earlier, parasite ubiquitin pathway enzymes have been previously characterised, however to date none have been proven to be specifically host-targeted [102-104,113,231,375]. This study therefore presents evidence for the first parasite-derived host-targeted ubiquitin pathway enzyme, the *T. spiralis* UBE2L3. This ubiquitin conjugating enzyme (E2) is highly conserved and plays an important role in the regulation of the cell cycle in mammalian cells [319,320]. UBE2L3 is involved in both the passage of a cell through S phase, and the regulation by ubiquitination of the tumour suppressor protein p53. P53 is required for the terminal differentiation of skeletal muscle cells and for the cell cycle arrest that occurs during the G2/M checkpoint in response to DNA damage [134,338,340]. Interestingly, during
nurse cell formation. *T. spiralis* muscle larvae induce cell cycle re-entry followed by a sustained state of G2/M arrest of infected myofibres via an unknown mechanism.

To determine the potential involvement of TsUBE2L3 in these processes, I set out to develop a system for the characterisation of the *T. spiralis* secreted E2 in muscle cells. Many helminths including *T. spiralis* are refractory to established methods of genetic manipulation. It was therefore not possible to characterise the function of TsUBE2L3 by manipulating its expression *in vivo* and looking at the subsequent effect on nurse cell formation in the host. Instead, I developed an *in vitro* system for the (over) expression of the *T. spiralis* E2 enzyme in a skeletal muscle cell line C2C12. A lentiviral-mediated expression system was adopted and stable myoblast cell lines were generated that could be differentiated and then induced to express the *T. spiralis* protein in a timely manner. This system could potentially be employed in the future for characterising other *T. spiralis* secreted proteins of interest, such as those identified by the proteomic study presented here.

C2C12 muscle cells expressing TsUBE2L3 were observed to differentiate and fuse forming myotubes at the same rate as the control cells. Their morphology however was altered. This alteration was not a result of a loss of skeletal muscle differentiation markers and did not affect the life span of the cells. Whitcomb *et al.* showed that the human UBE2L3 plays an important role in the progression of cells through S phase, but despite an unexplained rise in the rate of protein synthesis, the *T. spiralis* E2 did not enhance the rate of proliferation of C2C12 myoblasts [319,320]. The reason for the enhanced nuclear localisation of TsUBE2L3 in actively dividing myoblasts could therefore not be determined and I propose that a more robust investigation into the effect of the *T. spiralis* E2 on proliferation would be required before a role could be completely ruled out.

During infection, *T. spiralis*-induced cell cycle re-entry of host myofibres involves DNA replication, and myofibre nuclei remain thereafter as 4n [162]. I speculate that the process of forced DNA replication, in cells that would ordinarily never again replicate their DNA, causes DNA damage and mutation that may activate the G2/M checkpoint which induces the following arrest. Inducing DNA replication in terminally differentiated cells is a major feat. I therefore propose that during infection, this process requires a complex of *T. spiralis* secreted proteins acting in synergy. It is therefore unsurprising that TsUBE2L3 alone was unable to induce DNA replication in terminally differentiated myotubes. It was however able to down-regulate the G2/M
specific cell cycle protein cyclin B1. The mechanism and consequence of this effect on cyclin B1 remains to be determined.

The final finding was that TsUBE2L3 significantly down-regulates native p53 in C2C12 myotubes. This was surprising, since cyclin B1 down-regulation is usually concomitant with p53 activation and subsequent G2/M arrest, a phenotype that is induced by *T. spiralis*. It is therefore likely that either the mechanisms of p53 and cyclin B1 down-regulation in response to TsUBE2L3 occur independently of each other, or that they occur in a highly regulated time-dependent fashion. For example p53-induced down-regulation of cyclin B1 occurs first and is sustained, and then p53 itself is down-regulated. This study was not able to determine the mechanism of either. It did conclude that the p53 depression did not occur at the transcript level, and therefore must occur at the protein level. It was therefore plausible that, since UBE2L3 is a p53-specific E2 enzyme in mammalian cells, the mechanism of p53 down-regulation at the protein level was via proteasomal degradation.

Some of the final experiments presented in chapter 5 would need expanding before a firm conclusion on the function of the *T. spiralis* E2 in skeletal myotubes could be confirmed. Future studies will be able to expand on the techniques presented here to further characterise TsUBE2L3 and its effect on p53. For example, the methods presented in this thesis could be used to develop a C2C12 cell line that expresses HA-tagged Ub upon induction. This could be used to analyse the entire ubiquitome of the muscle cell and how it may be perturbed by the *T. spiralis* E2. Following this with proteomics analyses may identify ubiquitination substrates that are specific to the *T. spiralis* E2, including p53. Using methods such as microarray and co-precipitation could highlight altered gene expression patterns and interaction partners and therefore pathways in which the protein may be involved in muscle cells.

In summary, this thesis demonstrates viable methods for the characterisation of *T. spiralis* ubiquitin pathway proteins. Furthermore, this thesis demonstrates the successful development of a system for the characterisation of *T. spiralis* secreted proteins in terminally differentiated skeletal muscle cells *in vitro*, with a multitude of possible experimental read-outs.
Implications and future work

*T. spiralis* infection in humans is not considered a public health emergency because nowadays infection and mortality are rare [118]. *T. spiralis* however does infect mammals indiscriminately, posing an agricultural problem that affects livestock, and especially due to the nature of transmission, animals reared for meat production [376,377]. The treatment of muscle stage *T. spiralis* larvae during chronic infection is extremely challenging, due to the resilience of the nurse cell complex [179]. The initial intestinal phase of the life cycle however, is amenable to drug treatment. An intestinal phase of development is common to many parasitic nematodes, and in many cases this is the ultimate target tissue. Considering the high level of conservation of the ubiquitin pathway, prospective ubiquitin pathway drug targets in *T. spiralis* may also be expressed by other helminths. This could lead to the development of novel anthelmintics also active against parasites that pose a more widespread risk to human life [378].

The ubiquitin pathway is emerging as effective target for the treatment of diseases such as cancer. Due to its predominant role in the regulation of the cell cycle, its potential as a drug target for treating cancer has been considered for some time. Proteasome inhibitors have reached clinical trials, DUB inhibitors have shown anti-cancer potential in animal models, and ubiquitin conjugation machinery is being characterised with a view to identifying additional ubiquitin pathway targets [82,379]. Although thus far the ubiquitin pathway as a drug target in infectious disease has been largely overlooked, these studies contribute to the theory that it shows great potential.

The development of the nurse cell during the infection of muscle cells by *T. spiralis* involves the manipulation of many processes that require a functional ubiquitin pathway. It was therefore hypothesised that some of the *T. spiralis* host-targeted proteins might be themselves ubiquitin pathway enzymes. Studying the secreted proteins of *T. spiralis* muscle larvae allowed for the distinction to be made between the host-targeted and the ‘parasite’s own’ proteins. It is likely that many other parasitic nematodes express and secrete orthologues of these proteins, and the full proteomic analysis of *T. spiralis* secreted proteins will therefore be informative to the helminthology community. Likewise, by identifying proteins that are involved in the processes required for the reprogramming of skeletal muscle cells, this study contributes to our understanding of the role of the ubiquitin pathway in muscle cell
biology, and how it may be disrupted. This information is relevant in the fields of regenerative medicine and muscle cell cycle disorders. Understanding how *T. spiralis* is able to re-program terminally differentiated mammalian skeletal muscle cells will provide insight into potential methods of manipulation of muscle cells for medically beneficial purposes.

Firstly, as shown by Bai *et al.* the secreted proteins of *T. spiralis* muscle larvae are able to promote the proliferation of C2C12 myoblasts in culture [160,161]. Although the *T. spiralis* E2 alone did not mimic this effect, this demonstrates promise that *T. spiralis* may be used as a model for skeletal muscle-specific degenerative disorders such as muscular dystrophy [380]. In healthy muscle tissue, when myofibres die either due to age or physical injury to the tissue, satellite cells become activated to differentiate into myoblasts and myocytes which fuse into multinucleated, terminally differentiated myofibres [137,381]. One important feature of Duchenne muscular dystrophy is the failure of satellite cells to continue to self-renew, a feature important for maintaining the satellite cell pool. Over time the myoblast capacity to differentiate when required for the repair and regeneration of the myofibres diminishes [382]. Being able to isolate proteins that have the capacity to induce the proliferation of skeletal myoblasts may help identify pathways and mechanisms involved in stabilising the regenerative capacity of skeletal muscle tissue.

Secondly, during infection by *T. spiralis* the secreted proteins are able to force terminally differentiated skeletal muscle cells to re-enter the cell cycle, indicating that the secreted proteins have the capacity to reverse the differentiation status of myofibres. This also proves that the myofibre has the molecular capacity to re-enter the cell cycle. During infection by SV40 and Polyoma viruses, myofibres re-enter the cell cycle and also undergo mitosis [172]. This shows that the myofibre also has the capacity to divide. Although in the case of *T. spiralis*, the process of cell cycle re-entry terminates in G2/M arrest, it may be possible to separate the *T. spiralis* secreted proteins that induce the initial stages of cell cycle re-entry from those that induce the subsequent G2/M cell cycle arrest.

Dissecting the mechanisms involved in the deregulation of the myofibre will help us understand the molecular basis of terminal differentiation, how it is sustained and how it may be manipulated for the purposes of tissue regeneration and repair [152]. This may even be through the use of *T. spiralis*-derived muscle cell modulators. During infection, helminths such as *T. spiralis* occupy specialised niches within the host, often eliciting minimal pathogenicity. The *T. spiralis* muscle tissue niche is a
perfect example of how a helminth can strike a delicate balance between host manipulation and host protection. The parasite modulates cell cycle differentiation state, angiogenesis, collagen production and many other interesting biological processes. Understanding how these changes occur, may lead to the discovery or development of parasite-derived or parasite-inspired therapeutics, with the aim of treating disease. This may not be restricted to muscle or cell cycle-related disorders, nor to the *T. spiralis* niche as a model. This thesis therefore ends with the proposal that host-targeted proteins of parasites may one day become parasite-derived therapeutics.


## Appendices

### Appendix 1. Co-precipitation of 6His-TsUCH37 with *T. spiralis* lysate

<table>
<thead>
<tr>
<th>#</th>
<th>Identified Proteins (109)</th>
<th>Accession Number (alternative no.)</th>
<th>Predicted molecular weight</th>
<th>No. of unique peptides</th>
<th>Band</th>
<th>Conserved proteasome-related domain</th>
<th>Proteasome subunit</th>
<th>BLAST human homologue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>putative proteasome activator complex subunit 3[<em>Trichinella spiralis</em>](gi</td>
<td>316965573 (+1))</td>
<td>28 kDa</td>
<td>13</td>
<td>4, 5, 6, 10, 12, 16</td>
<td>PA28</td>
<td>immuno-proteasome (prosome)</td>
<td>PA28</td>
</tr>
<tr>
<td>2</td>
<td>26S proteasome non-ATPase regulatory subunit 2[<em>Trichinella spiralis</em>](gi</td>
<td>316977340 (+1))</td>
<td>102 kDa</td>
<td>28</td>
<td>2</td>
<td>RPN1</td>
<td>19S base</td>
<td>PC1</td>
</tr>
<tr>
<td>3</td>
<td>26S proteasome non-ATPase regulatory subunit 1[<em>Trichinella spiralis</em>](gi</td>
<td>316971878 (+1))</td>
<td>156 kDa</td>
<td>18</td>
<td>1</td>
<td>RPN2</td>
<td>19S base</td>
<td>PC1</td>
</tr>
<tr>
<td>4</td>
<td>26S protease regulatory subunit 4[<em>Trichinella spiralis</em>](gi</td>
<td>316975801 (+1))</td>
<td>74 kDa</td>
<td>18</td>
<td>8</td>
<td>RPT1</td>
<td>19S base</td>
<td>PC1</td>
</tr>
<tr>
<td>5</td>
<td>putative 26S proteasome non-ATPase regulatory subunit 8[<em>Trichinella spiralis</em>](gi</td>
<td>316974557 (+1))</td>
<td>31 kDa</td>
<td>7</td>
<td>7</td>
<td>RPT1</td>
<td>19S base</td>
<td>PC1</td>
</tr>
<tr>
<td>6</td>
<td>26S proteasome non-ATPase regulatory subunit 4[<em>Trichinella spiralis</em>](gi</td>
<td>316973936 (+1))</td>
<td>45 kDa</td>
<td>4</td>
<td>4, 9</td>
<td>RPN10</td>
<td>19S base</td>
<td>PC1</td>
</tr>
<tr>
<td>7</td>
<td>26S protease regulatory subunit 6A[<em>Trichinella spiralis</em>](gi</td>
<td>316979457 (+1))</td>
<td>51 kDa</td>
<td>8</td>
<td>7</td>
<td>RPT1</td>
<td>19S base</td>
<td>PC1</td>
</tr>
<tr>
<td>8</td>
<td>proteasome subunit alpha type-1[<em>Trichinella spiralis</em>](gi</td>
<td>316971831 (+1))</td>
<td>218 kDa</td>
<td>4</td>
<td>8</td>
<td>Ntn hydrolase</td>
<td>20S</td>
<td>proteasome subunit alpha type-1</td>
</tr>
<tr>
<td>9</td>
<td>proteasome subunit beta type-1[<em>Trichinella spiralis</em>](gi</td>
<td>316971688 (+1))</td>
<td>34 kDa</td>
<td>2</td>
<td>8</td>
<td>Ntn hydrolase</td>
<td>20S</td>
<td>proteasome beta 1 subunit variant</td>
</tr>
<tr>
<td>10</td>
<td>proteasome activator complex subunit 3[<em>Trichinella spiralis</em>](gi</td>
<td>316969768 (+1))</td>
<td>16 kDa</td>
<td>3</td>
<td>7, 8</td>
<td>PA28</td>
<td>immuno-proteasome (prosome)</td>
<td>PC1</td>
</tr>
<tr>
<td>11</td>
<td>proteasome subunit alpha type-6[<em>Trichinella spiralis</em>](gi</td>
<td>316962618 (+1))</td>
<td>13 kDa</td>
<td>2</td>
<td>8</td>
<td>Ntn hydrolase</td>
<td>20S</td>
<td>proteasome subunit alpha type-6</td>
</tr>
<tr>
<td>12</td>
<td>proteasome subunit beta type-6[<em>Trichinella spiralis</em>](gi</td>
<td>316975433 (+1))</td>
<td>62 kDa</td>
<td>3</td>
<td>12, 13</td>
<td>Ntn hydrolase</td>
<td>20S</td>
<td>proteasome subunit beta type-6</td>
</tr>
<tr>
<td>13</td>
<td>proteasome subunit alpha type-2 A[<em>Trichinella spiralis</em>](gi</td>
<td>316973439 (+1))</td>
<td>26 kDa</td>
<td>4</td>
<td>10</td>
<td>Ntn hydrolase</td>
<td>20S</td>
<td>proteasome subunit alpha type-2</td>
</tr>
<tr>
<td>14</td>
<td>26S proteasome non-ATPase regulatory subunit 11[<em>Trichinella spiralis</em>](gi</td>
<td>316975446 (+1))</td>
<td>50 kDa</td>
<td>4</td>
<td>4</td>
<td>RPN6</td>
<td>19S lid</td>
<td>PC1</td>
</tr>
<tr>
<td>15</td>
<td>protein ADRM1[<em>Trichinella spiralis</em>](gi</td>
<td>316977949 (+1))</td>
<td>48 kDa</td>
<td>3</td>
<td>9</td>
<td>Rpn13, UCH37, Ntn</td>
<td>19S base</td>
<td>ADRM1</td>
</tr>
<tr>
<td>16</td>
<td>26S proteasome non-ATPase regulatory subunit 3[<em>Trichinella spiralis</em>](gi</td>
<td>316974493 (+1))</td>
<td>59 kDa</td>
<td>6</td>
<td>7</td>
<td>PCI, Rpn3, Ntn</td>
<td>19S lid</td>
<td>26S proteasome non-ATPase regulatory subunit 3</td>
</tr>
<tr>
<td>17</td>
<td>26S proteasome non-ATPase regulatory subunit 4[<em>Trichinella spiralis</em>](gi</td>
<td>316973936 (+1))</td>
<td>45 kDa</td>
<td>6</td>
<td>9</td>
<td>RPN10</td>
<td>19S base</td>
<td>PC1</td>
</tr>
<tr>
<td>18</td>
<td>26S protease regulatory subunit 6B[<em>Trichinella spiralis</em>](gi</td>
<td>316968151 (+1))</td>
<td>39 kDa</td>
<td>4</td>
<td>11</td>
<td>RPT1</td>
<td>19S base</td>
<td>PC1</td>
</tr>
<tr>
<td>19</td>
<td>26S protease regulatory subunit 8 [<em>Trichinella spiralis</em>](gi</td>
<td>316972496 (+1))</td>
<td>39 kDa</td>
<td>4</td>
<td>12</td>
<td>RPT1</td>
<td>19S base</td>
<td>PC1</td>
</tr>
</tbody>
</table>

### Appendix 1. Co-precipitation of 6His-TsUCH37 with *T. spiralis* lysate

Table of all putative *T. spiralis* proteasome subunits identified by LC/MS/MS of the co-precipitation of 6His-TsUCH37 with lysate of *T. spiralis* muscle larvae.
### Appendix 2. Bioinformatics prediction of *T. spiralis* secreted proteins

<table>
<thead>
<tr>
<th>Programme</th>
<th>GI accession #</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>iPSORT, wolfPSORT and signalP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>339242525</td>
<td></td>
<td>E3 ubiquitin-protein ligase</td>
</tr>
<tr>
<td>339238271</td>
<td></td>
<td>E3 ubiquitin-protein ligase RNF34</td>
</tr>
<tr>
<td>33927319</td>
<td></td>
<td>RING finger protein 13</td>
</tr>
<tr>
<td>339245151</td>
<td></td>
<td>RWD domain protein</td>
</tr>
<tr>
<td><strong>iPSORT only</strong></td>
<td>339261774</td>
<td>BTB/POZ domain-containing protein KCTD17</td>
</tr>
<tr>
<td><strong>iPSORT and wolfPSORT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>339235881</td>
<td></td>
<td>E3 ubiquitin-protein ligase synoviolin-A</td>
</tr>
<tr>
<td>339262132</td>
<td></td>
<td>E3 ubiquitin-protein ligase sia-1</td>
</tr>
<tr>
<td>339237381</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 2</td>
</tr>
<tr>
<td>339239888</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2</td>
</tr>
<tr>
<td>339264723</td>
<td></td>
<td>RING-H2 finger protein ATL1D</td>
</tr>
<tr>
<td>339245539</td>
<td></td>
<td>polycomb group RING finger protein 3</td>
</tr>
<tr>
<td>339236113</td>
<td></td>
<td>polycomb group RING finger protein 1</td>
</tr>
<tr>
<td>339264545</td>
<td></td>
<td>bromodomain and WD repeat-containing protein 2</td>
</tr>
<tr>
<td>339237985</td>
<td></td>
<td>WD domain, G-beta repeat-containing domain protein</td>
</tr>
<tr>
<td>339248085</td>
<td></td>
<td>putative WD repeat-containing protein 85</td>
</tr>
<tr>
<td>339239445</td>
<td></td>
<td>putative WD repeat-containing protein 3</td>
</tr>
<tr>
<td>339238113</td>
<td></td>
<td>WD domain, G-beta repeat-containing domain protein(WD48)</td>
</tr>
<tr>
<td>339242135</td>
<td></td>
<td>bromodomain and WD repeat-containing protein 2</td>
</tr>
<tr>
<td>339253088</td>
<td></td>
<td>putative HECT-domain protein (NEDD4-L)</td>
</tr>
<tr>
<td><strong>iPSORT and signalP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>339238337</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2-21 protein</td>
</tr>
<tr>
<td>33923478</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 K</td>
</tr>
<tr>
<td>339236213</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 D</td>
</tr>
<tr>
<td>339238007</td>
<td></td>
<td>ubiquitin carboxy- hydrolase, family 1</td>
</tr>
<tr>
<td>339248837</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 variant 2</td>
</tr>
<tr>
<td>339244557</td>
<td></td>
<td>ubiquitin carboxyl- hydrolase 7</td>
</tr>
<tr>
<td>339236793</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 Q1</td>
</tr>
<tr>
<td>339250668</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 G2</td>
</tr>
<tr>
<td>339252782</td>
<td></td>
<td>ubiquitin-conjugating enzyme</td>
</tr>
<tr>
<td>339242301</td>
<td></td>
<td>ubiquitin-fold modifier 1</td>
</tr>
<tr>
<td>339233154</td>
<td></td>
<td>putative ubiquitin associated and SH3 domain-containing protein A</td>
</tr>
<tr>
<td>339253040</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 1</td>
</tr>
<tr>
<td>339252718</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 1</td>
</tr>
<tr>
<td>339243657</td>
<td></td>
<td>putative ubiquitin associated and SH3 domain-containing protein B</td>
</tr>
<tr>
<td>339257344</td>
<td></td>
<td>ubiquitin-conjugating enzyme E22</td>
</tr>
<tr>
<td>339232770</td>
<td></td>
<td>E3 ubiquitin-protein ligase UBR3</td>
</tr>
<tr>
<td>339246159</td>
<td></td>
<td>putative ubiquitin carboxyl- hydrolase CYLD</td>
</tr>
<tr>
<td>339260542</td>
<td></td>
<td>putative ubiquitin associated and SH3 domain-containing protein B</td>
</tr>
<tr>
<td>339261298</td>
<td></td>
<td>putative ubiquitin family protein</td>
</tr>
<tr>
<td>339239021</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2-18 kDa</td>
</tr>
<tr>
<td>339238085</td>
<td></td>
<td>E3 ubiquitin-protein ligase synoviolin-A</td>
</tr>
<tr>
<td>339240047</td>
<td></td>
<td>ubiquitin-conjugating enzyme E2 L3</td>
</tr>
<tr>
<td>339241693</td>
<td></td>
<td>ubiquitin-related modifier 1</td>
</tr>
<tr>
<td>ID</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>339262330</td>
<td>putative ubiquitin family protein</td>
<td></td>
</tr>
<tr>
<td>339253028</td>
<td>ubiquitin-conjugating enzyme E2 1</td>
<td></td>
</tr>
<tr>
<td>339257342</td>
<td>ubiquitin-conjugating enzyme E22</td>
<td></td>
</tr>
<tr>
<td>339259260</td>
<td>ubiquitin carboxyl-hydrolase 4</td>
<td></td>
</tr>
<tr>
<td>339247039</td>
<td>ubiquitin carboxyl-hydrolase 20</td>
<td></td>
</tr>
<tr>
<td>339265467</td>
<td>ubiquitin conjugation factor E4 B</td>
<td></td>
</tr>
<tr>
<td>339257112</td>
<td>ubiquitin-conjugating enzyme E22</td>
<td></td>
</tr>
<tr>
<td>339259776</td>
<td>ubiquitin carboxyl-hydrolase superfamily</td>
<td></td>
</tr>
<tr>
<td>339252976</td>
<td>ubiquitin-conjugating enzyme E2 1</td>
<td></td>
</tr>
<tr>
<td>339232982</td>
<td>ubiquitin family protein</td>
<td></td>
</tr>
<tr>
<td>339260657</td>
<td>putative ubiquitin associated and SH3 domain-containing protein B</td>
<td></td>
</tr>
<tr>
<td>339259274</td>
<td>E3 ubiquitin-protein ligase UBR5</td>
<td></td>
</tr>
<tr>
<td>339265187</td>
<td>putative ubiquitin-conjugating enzyme E2 9</td>
<td></td>
</tr>
<tr>
<td>339252692</td>
<td>ubiquitin-conjugating enzyme E2 1</td>
<td></td>
</tr>
<tr>
<td>339252686</td>
<td>ubiquitin-conjugating enzyme E2 1</td>
<td></td>
</tr>
<tr>
<td>339239576</td>
<td>ubiquitin-carboxyl-hydrolase superfamily</td>
<td></td>
</tr>
<tr>
<td>339252898</td>
<td>ubiquitin associated and SH3 domain-containing protein B</td>
<td></td>
</tr>
<tr>
<td>339233958</td>
<td>ubiquitin family protein</td>
<td></td>
</tr>
<tr>
<td>339263743</td>
<td>ubiquitin associated and SH3 domain-containing protein B</td>
<td></td>
</tr>
<tr>
<td>339239901</td>
<td>putative ubiquitin associated and SH3 domain-containing protein B</td>
<td></td>
</tr>
<tr>
<td>339252980</td>
<td>ubiquitin-conjugating enzyme E2 1</td>
<td></td>
</tr>
<tr>
<td>339232996</td>
<td>ubiquitin family protein</td>
<td></td>
</tr>
<tr>
<td>339246865</td>
<td>ubiquitin carboxyl-hydrolase 14</td>
<td></td>
</tr>
<tr>
<td>339263786</td>
<td>putative ubiquitin family protein</td>
<td></td>
</tr>
<tr>
<td>339258236</td>
<td>putative ubiquitin associated and SH3 domain-containing protein B</td>
<td></td>
</tr>
<tr>
<td>339242727</td>
<td>ubiquitin carboxyl-hydrolase BAP1</td>
<td></td>
</tr>
<tr>
<td>339257336</td>
<td>ubiquitin-conjugating enzyme E2 2</td>
<td></td>
</tr>
<tr>
<td>339252978</td>
<td>ubiquitin-conjugating enzyme E2 1</td>
<td></td>
</tr>
<tr>
<td>339256894</td>
<td>ubiquitin-related modifier 1</td>
<td></td>
</tr>
<tr>
<td>339258230</td>
<td>putative ubiquitin associated and SH3 domain-containing protein B</td>
<td></td>
</tr>
<tr>
<td>339258130</td>
<td>putative ubiquitin associated and SH3 domain-containing protein B</td>
<td></td>
</tr>
<tr>
<td>339253044</td>
<td>ubiquitin-conjugating enzyme E2 1</td>
<td></td>
</tr>
<tr>
<td>339257338</td>
<td>ubiquitin-conjugating enzyme E2 2</td>
<td></td>
</tr>
<tr>
<td>339236215</td>
<td>putative ubiquitin-conjugating enzyme protein 17</td>
<td></td>
</tr>
<tr>
<td>339252880</td>
<td>E3 ubiquitin-protein ligase</td>
<td></td>
</tr>
<tr>
<td>339250174</td>
<td>BTB/POZ domain-containing protein KCTD12</td>
<td></td>
</tr>
<tr>
<td>339236281</td>
<td>putative BTB/POZ domain protein</td>
<td></td>
</tr>
<tr>
<td>339259936</td>
<td>BTB/POZ domain-containing protein TNFAIP1</td>
<td></td>
</tr>
<tr>
<td>339246047</td>
<td>BTB/POZ domain-containing protein KCTD2</td>
<td></td>
</tr>
<tr>
<td>339250176</td>
<td>BTB/POZ domain-containing protein KCTD8</td>
<td></td>
</tr>
<tr>
<td>339246107</td>
<td>BTB/POZ domain-containing protein KCTD3</td>
<td></td>
</tr>
<tr>
<td>339243199</td>
<td>putative BTB/POZ domain protein</td>
<td></td>
</tr>
<tr>
<td>339250172</td>
<td>BTB/POZ domain-containing protein KCTD16</td>
<td></td>
</tr>
<tr>
<td>339249469</td>
<td>putative UBA/TS-N domain protein</td>
<td></td>
</tr>
<tr>
<td>339265151</td>
<td>cullin-4B</td>
<td></td>
</tr>
<tr>
<td>339240427</td>
<td>cullin family protein</td>
<td></td>
</tr>
<tr>
<td>Accession Number</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>339250980</td>
<td>RING finger protein 44</td>
<td></td>
</tr>
<tr>
<td>339243635</td>
<td>putative MIZ/SP-RING zinc finger</td>
<td></td>
</tr>
<tr>
<td>339252988</td>
<td>RING-H2 finger protein ATL2A</td>
<td></td>
</tr>
<tr>
<td>339253520</td>
<td>SUMO-conjugating enzyme UBC9</td>
<td></td>
</tr>
<tr>
<td>339235373</td>
<td>WD repeat-containing protein 68</td>
<td></td>
</tr>
<tr>
<td>339261602</td>
<td>putative WD repeat-containing protein 54</td>
<td></td>
</tr>
<tr>
<td>339248885</td>
<td>WD repeat-containing protein 82</td>
<td></td>
</tr>
<tr>
<td>339249903</td>
<td>putative WD repeat-containing protein 43</td>
<td></td>
</tr>
<tr>
<td>339247185</td>
<td>WD repeat-containing protein 20</td>
<td></td>
</tr>
<tr>
<td>339247269</td>
<td>WD repeat-containing protein 74</td>
<td></td>
</tr>
<tr>
<td>339248763</td>
<td>katanin p80 WD40-containing subunit B1</td>
<td></td>
</tr>
<tr>
<td>339234581</td>
<td>WD repeat-containing protein 51A</td>
<td></td>
</tr>
<tr>
<td>339261652</td>
<td>putative bromodomain and WD repeat-containing protein 2</td>
<td></td>
</tr>
<tr>
<td>339262962</td>
<td>WD repeat-containing protein 51A</td>
<td></td>
</tr>
<tr>
<td>339237779</td>
<td>WD domain, G-beta repeat-containing domain protein</td>
<td></td>
</tr>
<tr>
<td>339238527</td>
<td>WD domain, G-beta repeat-containing domain protein</td>
<td></td>
</tr>
<tr>
<td>339265355</td>
<td>WD repeat-containing protein 47</td>
<td></td>
</tr>
<tr>
<td>339264228</td>
<td>bromodomain and WD repeat-containing protein 2</td>
<td></td>
</tr>
<tr>
<td>339238381</td>
<td>WD domain, G-beta repeat-containing domain protein</td>
<td></td>
</tr>
<tr>
<td>339242147</td>
<td>bromodomain and WD repeat-containing protein 2 (WDR11)</td>
<td></td>
</tr>
<tr>
<td>339246265</td>
<td>WD domain, G-beta repeat-containing domain protein</td>
<td></td>
</tr>
<tr>
<td>339242167</td>
<td>putative bromodomain and WD repeat-containing protein 2</td>
<td></td>
</tr>
<tr>
<td>339240803</td>
<td>WD repeat-containing protein 19</td>
<td></td>
</tr>
<tr>
<td>339237047</td>
<td>putative WD repeat-containing protein 55</td>
<td></td>
</tr>
<tr>
<td>339262234</td>
<td>WD domain, G-beta repeat-containing domain protein</td>
<td></td>
</tr>
<tr>
<td>339264901</td>
<td>WD repeat-containing protein 82</td>
<td></td>
</tr>
<tr>
<td>339242171</td>
<td>putative bromodomain and WD repeat-containing protein 2</td>
<td></td>
</tr>
<tr>
<td>339241089</td>
<td>WD repeat-containing protein 24</td>
<td></td>
</tr>
<tr>
<td>339241997</td>
<td>WD repeat-containing protein 24</td>
<td></td>
</tr>
<tr>
<td>339261604</td>
<td>putative ThiF family protein</td>
<td></td>
</tr>
<tr>
<td>339253356</td>
<td>putative HECT-domain protein</td>
<td></td>
</tr>
<tr>
<td>339250272</td>
<td>putative HECT-domain protein</td>
<td></td>
</tr>
<tr>
<td>339253450</td>
<td>putative HECT-domain protein</td>
<td></td>
</tr>
<tr>
<td>339261350</td>
<td>E3 SUMO-protein ligase SIZ2</td>
<td></td>
</tr>
<tr>
<td>339253520</td>
<td>SUMO-conjugating enzyme UBC9</td>
<td></td>
</tr>
<tr>
<td>339264292</td>
<td>E3 SUMO-protein ligase PIAS4</td>
<td></td>
</tr>
<tr>
<td>wolfPSORT and signalP</td>
<td>E3 ubiquitin-protein ligase synoviolin-A</td>
<td></td>
</tr>
<tr>
<td>signalP only</td>
<td>putative ubiquitin--protein ligase</td>
<td></td>
</tr>
</tbody>
</table>

Appendix 2. Bioinformatics prediction of *T. spiralis* secreted proteins

All ubiquitin-related putative *T. spiralis* proteins predicted by iPSORT, signalP and wolfPSORT to be secreted
### Appendix 3. Proteomic analysis of anti-NEDD4-L reactive *T. spiralis* SP

<table>
<thead>
<tr>
<th>#</th>
<th>Identified Proteins (29)</th>
<th>Accession Number</th>
<th>Molecular Weight</th>
<th>Band excised (numbered)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>serine protease [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>3169805931</td>
<td>35 kDa</td>
</tr>
<tr>
<td>2</td>
<td>serine protease [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316980533  (+2)</td>
<td>48 kDa</td>
</tr>
<tr>
<td>3</td>
<td>multi cystatin-like domain protein precursor [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>3111183175  (+3)</td>
<td>46 kDa</td>
</tr>
<tr>
<td>4</td>
<td>secreted S-nucleotidase [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>22656349</td>
<td>62 kDa</td>
</tr>
<tr>
<td>5</td>
<td>conserved hypothetical protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316972936  (+1)</td>
<td>59 kDa</td>
</tr>
<tr>
<td>6</td>
<td>conserved hypothetical protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316974176  (+1)</td>
<td>41 kDa</td>
</tr>
<tr>
<td>7</td>
<td>RWD domain protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316972581  (+3)</td>
<td>77 kDa</td>
</tr>
<tr>
<td>8</td>
<td>putative sulfhydryl oxidase 1 [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316977388  (+1)</td>
<td>73 kDa</td>
</tr>
<tr>
<td>9</td>
<td>large tegument protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316971025  (+1)</td>
<td>84 kDa</td>
</tr>
<tr>
<td>10</td>
<td>putative Low-density lipoprotein receptor repeat class B [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316971109  (+1)</td>
<td>168 kDa</td>
</tr>
<tr>
<td>11</td>
<td>conserved hypothetical protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316966482  (+1)</td>
<td>58 kDa</td>
</tr>
<tr>
<td>12</td>
<td>putative sodium/potassium-transporting ATPase subunit beta-1 [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316971027  (+1)</td>
<td>91 kDa</td>
</tr>
<tr>
<td>13</td>
<td>conserved hypothetical protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>339258164</td>
<td>29 kDa</td>
</tr>
<tr>
<td>14</td>
<td>conserved hypothetical protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316972588  (+1)</td>
<td>51 kDa</td>
</tr>
<tr>
<td>15</td>
<td>putative fasciclin domain protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316965855  (+1)</td>
<td>66 kDa</td>
</tr>
<tr>
<td>16</td>
<td>putative ISDe13, transposase OrfA [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316975125  (+1)</td>
<td>50 kDa</td>
</tr>
<tr>
<td>17</td>
<td>DNase II [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>317581955  (+3)</td>
<td>38 kDa</td>
</tr>
<tr>
<td>18</td>
<td>48 kDa poly-cysteine and histidine tagged protein precursor [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>3179170388  (+5)</td>
<td>48 kDa</td>
</tr>
<tr>
<td>19</td>
<td>putative calcium binding EGF domain protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316975135  (+1)</td>
<td>280 kDa</td>
</tr>
<tr>
<td>20</td>
<td>conserved hypothetical protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316974192  (+1)</td>
<td>57 kDa</td>
</tr>
<tr>
<td>21</td>
<td>putative nudix hydrolase [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>164521920</td>
<td>46 kDa</td>
</tr>
<tr>
<td>22</td>
<td>putative Fe-S protein assembly co-chaperone Hsc8 [48 kDa poly-cysteine and histidine tagged protein precursor [<em>Trichinella spiralis</em>]]</td>
<td>gi</td>
<td>316965621  (+1)</td>
<td>119 kDa</td>
</tr>
<tr>
<td>23</td>
<td>adenyls/cysteine synthase [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316966393  (+1)</td>
<td>47 kDa</td>
</tr>
<tr>
<td>24</td>
<td>enolase [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>139373125  (+3)</td>
<td>48 kDa</td>
</tr>
<tr>
<td>25</td>
<td>DNA polymerase (Pol2) superfamily [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316969308  (+1)</td>
<td>134 kDa</td>
</tr>
<tr>
<td>26</td>
<td>UNC45 protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316971886  (+1)</td>
<td>138 kDa</td>
</tr>
<tr>
<td>27</td>
<td>papilin [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316976630  (+1)</td>
<td>231 kDa</td>
</tr>
<tr>
<td>28</td>
<td>conserved hypothetical protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316964581  (+1)</td>
<td>30 kDa</td>
</tr>
<tr>
<td>29</td>
<td>conserved hypothetical protein [<em>Trichinella spiralis</em>]</td>
<td>gi</td>
<td>316978836  (+1)</td>
<td>52 kDa</td>
</tr>
</tbody>
</table>

**Appendix 3. Proteomic analysis of anti-NEDD4-L reactive *T. spiralis* SP**

LC/MS/MS results of 6 anti-NEDD4-L reactive protein bands manually excised for analysis

268
Appendix 4. Analysis of the effect of MG132 on TsUBE2L3-HA expression in C2C12 myotubes.

Myotube empty vector and TsUBE2L3-HA cell lines were grown in the presence and absence of MG132 before analysis by immuno-blot (IB) of the effect of MG132 on TsUBE2L3-HA using anti-HA antibodies. The same samples were analysed using anti-tubulin antibodies as a protein loading control.
Appendix 5. Analysis of UBE2L3 expression in C2C12 myotube empty vector and TsUBE2L3-HA cell lines.

Myotube empty vector and TsUBE2L3-HA cell lines were induced using doxycycline for 24 hours before analysis by immuno-blot (IB) using anti-human UBE2L3 antibodies. The same samples were analysed by coomassie staining as a protein loading control.
Appendix 6. Paraffin sections of *T. spiralis* infected rat muscle tissue

*T. spiralis* infected rat tissue (triceps, tongue and diaphragm) was collected 1 month after infection and 12 months after infection. Tissues were dissected and fixed in 10% neutral buffered formalin overnight at 4°C. Tissues were embedded in paraffin and 4 \( \mu \)m sections were cut and stained using Hematoxylin and eosin, all using standard protocols [194]. Arrows indicate the outer collagen capsule.
Appendix 7. Structural analysis of the catalytic domain of TsUCH37

Superimposition of human UCH37 and TsUCH37\textsuperscript{cat}-Ub-VME crystal structures

A. Full structure

B. Active site

C. Electrostatic charges of active site cleft

Appendix 7. Structural analysis of the catalytic domain of TsUCH37

The crystal structure of TsUCH37\textsuperscript{cat}-Ub-VME (Morrow et al. Biochemistry 2013 [3]). A. The crystal structure of TsUCH37\textsuperscript{cat} (blue) in complex with Ub-VME (orange) with the structure of human UCH37 (olive, PDB 3RIS) superimposed. The crossover loop location of each DUB is boxed in red. B. A close-up of the active-site residues of the TsUCH37\textsuperscript{cat}-Ub-VME and human UCH37 superimposition. Positions of the active-site residues Gin, Cys, His, and Asp and their bond distances are shown. C. Electrostatic charge map of the active-site clefts of human UCH37 (left) with Ub modelled into the binding position (grey) and TsUCH37\textsuperscript{cat} (right) bound to Ub-VME (orange). The catalytic cleft is circled and the active-site cysteine is highlighted in yellow. Positive charges are red, negative charges are blue and neutral charged are white.

All images modified from Morrow et al, Biochemistry, 2013.
End