Mechanisms of Prion Disease: Initial Misfolding of the Prion Protein and Metabolomic Changes in the Brain

by

Zelin Fu

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

Doctor of Philosophy

Department of Biochemistry
University of Alberta

© Zelin Fu, 2020
Abstract

Prion diseases are neurodegenerative diseases that are caused by the misfolding and aggregation of cellular prion protein (PrP\textsuperscript{C}) into the pathogenic form termed PrP\textsuperscript{Sc}. To contribute to a better understanding of the mechanism of prion disease, in this thesis I present my work on the structural characterization of a mutant of PrP\textsuperscript{C} (HRdup), which led to genetic prion disease and the investigation of the brain metabolomic changes during prion disease progression. In Chapter 2, the structural stability of HRdup was examined as well as the structure of its canonical antiparallel β-sheet. Detailed structural study of HRdup such as its backbone rigidity, is covered in Chapter 3. Combining the results from these two chapters, we learned the extra residues introduced by an insert to the central hydrophobic region of HRdup do not disrupt its canonical PrP\textsuperscript{C} structure or have major effect on its structural stability, but a few of these residues in the insert displayed high β-propensity which is indicative of an additional β-turn in this otherwise disordered region. In Chapter 4, employing a novel metabolite extraction method, the brain metabolome of the mice inoculated with mice prion was analyzed along the progression of the disease. The brain metabolome profile consisted of 50 metabolites displayed continuous changes and the key metabolites were identified. These results contribute to a better understanding of the toxicity mechanism of prion disease, suggesting that tracking of neurochemical profiles could be effective in monitoring the progression of prion diseases as well as other neurodegenerative diseases. Chapter 5 summarized both research projects and expressed some of my thoughts about the future of both studies.
Preface

Chapter 1, the introduction chapter as well as Chapter 5, the conclusion chapter are my own original work.

The content of Chapter 2 has been published in PLoS Pathogen as Robert C. C. Mercer, Nathalie Daude, Lyudmyla Dorosh, Ze-Lin Fu, Charles E. May, Hristina Gapeshina, Serene L. Wohlgemuth, Claudia Y. Acevedo-Morantes, Jing Yang, Neil R. Cashman, Michael B. Coulthart, Dawn M. Pearson, Jeffrey T. Joseph, Holger Wille, Jiri G. Safar, Gerard H. Jansen, Maria Stepanova, Brian D. Sykes, David Westaway (2018) *A novel Gerstmann-Sträussler-Scheinker disease mutation defines a precursor for amyloidogenic 8 kDa PrP fragments and reveals N-terminal structural changes shared by other GSS alleles.* For this work, ZLF synthesized the recombinant protein material used in the NMR experiments in this study. ZLF and BDS designed and performed the NMR experiments as well as analyzing the data; ZLF made Figure 4A, 4B and 4D and wrote the “Stability and β-sheet structure of HRdup PrP” portion of the result section and NMR portion in material and method section with BDS and DW.

The content of Chapter 3 has been published in Journal of Molecular Biology as Ze-Lin Fu, Peter C. Holmes, David Westaway, Brian D. Sykes (2019) *Nascent β structure in the elongated hydrophobic region of a Gerstmann-Sträussler-Scheinker PrP allele.* ZLF, DW and BDS designed the experiments. ZLF prepared all the recombinant proteins. ZLF and BDS performed NMR experiments. ZLF and PCH analyzed the NMR data. ZLF wrote the manuscript with BDS, DW and PCH.

A modified version of Chapter 4 has been re-submitted to Metabolomics journal as Ze-Lin Fu, Pascal Mercier, Ghazaleh Eskandary-Sedighi, Jing Yang, David Westaway, Brian D. Sykes (2020) *Metabolomic study of disease progression in scrapie prion infected mice; validation of a novel method for brain metabolite extraction.* DW, BDS and ZLF conceived and designed the research. JY and GES collected the samples. GES
performed western blot analysis. ZLF conducted the metabolite extraction, NMR experiments and analyzed the data with PM. ZLF wrote the manuscript with DW and BDS.

All the animal studies were conducted in accordance with Canadian Council on Animal Care (CCAC) guidelines and the protocols were approved by the animal care use committee for Health Sciences laboratory Animal Services at the University of Alberta. The protocols used are AUP00000356 and AUP00000357.
“Honestly I didn’t think I’d make it this far.”

-Po, Kung Fu Panda
Acknowledgements

First and foremost, I would like to thank both of my supervisors, Dr. Brian Sykes and Dr. David Westaway. It was such great fortune of mine that I had not only one, but two outstanding scientists and excellent mentors to guide me through the program (potentially I just needed that much extra help) and I cannot thank you both enough. For the past years, I did not need to be concerned about anything else other than my own science and becoming a better scientific researcher. Without any hesitation, I could always simply come to either of you for help and advice (as well as rescues from time to time). I have learned from both of you way more than just science; you have influenced me on how to learn, research, and think critically. In the future, whenever I play golf, call other people “bozos” and have a pint of Guinness, I will certainly be thinking of you two and the great times we spent together.

Come with two great mentors are two wonderful labs. I would like to thank all the former and current lab members for making my graduate school experience so much more accommodating. Big thanks to Dr. John Paul Glaves, who helped me kick-start my graduate studies. Without him walking me through the initial lab training and working in the containment area, it would have been a lot more difficult for me to start off my program. At the beginning, I struggled to understand Dr. Robert Mercer, who speaks rather fast and has a sort of jumpy sense of humor; from Robert I received a lot of useful information and material about the HRdup mutant I ended up working with and I really appreciate that. Dr. Peter Holmes and Dr. Pascal Mercier played such important roles in my research, without whose generous help the projects I have worked on would be nowhere near where they are right now. As for my current lab members, in no particular sequence, I would like to thank: Fangze Cai, Brittney Klein, Kieran Cockburn, Dr. Serene Wohlgemuth, Dr. Sang-Gyun Kang, Dr. Nathalie Daude, Dr. Andrew Castle, Dr. Ghazaleh Eskandari-Sedighi, Jing Yang, Luis Arce, George Han and Klinton Shmeit. With these people I immensely enjoyed my experiences of all the golfing events, Christmas parties, potluck picnics, Friday lunches and beers.

Friday beer is an event I hold near and dear to my heart. Other than a few Westaway lab members there are also Dr. Camilo Velasquez, Dr. Richard Kanyo, Dr. Allen Herbst and Dr. Alicia Otero as part of the drinking crew, with whom I have shared so much laughter as well as numerous green onion cakes and fries. I am very grateful for their company. I would also
like to thank Kelsey at Sherlock’s pub on campus for taking good care of our table and the brewery of Arthur Guinness at St. James’s Gate for making Guinness, my go-to drink for Fridays (and many other days).

I would also like to thank Dr. Holger Wille for being on my supervisory committee and for many insightful and helpful discussions throughout my program. Many thanks to Philip Liu from Hwang lab for helping me out in the lab as well as recommending me many delicious recipes, and I apologize for still not having used my slow cooker four or five years after purchasing it.

Thanks to all my friends who kept me sane. In no specific order, very likely to be missing a few here: Michael Lajeunesse, Nicolai Nodilo, Dr. Matthias Lipfert, Dr. Angelique Faramus, Dr. Matthias Heger, Xin Wang, Giselle Uzcategui and Daniel Moran. I look forward to the future reunions and cannot wait to spend more time with all of you.

I have always been and always will be grateful for my dear parents. It could not have been easy for you to have your only child live half-an-earth away, but you have been nothing but supportive of my decisions every step of the way. I love you both deeply and we will enjoy our ordinary yet extraordinary life together as a tight unit as we always have. Thank you Jasmine, for showing up in my life at the perfect time and showering me with all your love and support. We have a wonderful journey ahead of us and I am very excited for what is waiting for the two of us to explore and enjoy.

Finally, I would like to express my gratitude to the Department of Biochemistry for granting me the opportunity to study here at University of Alberta. And I would like to thank the funding agencies, CIHR and APRI for supporting my research work directly and indirectly.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSE</td>
<td>Bovine spongiform encephalopathy</td>
</tr>
<tr>
<td>CJD</td>
<td>Creutzfeldt-Jakob disease</td>
</tr>
<tr>
<td>CSF</td>
<td>Cerebrospinal fluid</td>
</tr>
<tr>
<td>CWD</td>
<td>Chronic wasting disease</td>
</tr>
<tr>
<td>HR</td>
<td>Hydrophobic region</td>
</tr>
<tr>
<td>HRdup</td>
<td>Hydrophobic region partially duplicated prion protein mutant</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear single-quantum coherence</td>
</tr>
<tr>
<td>LDA</td>
<td>Linear Discriminant Analysis</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamic</td>
</tr>
<tr>
<td>MRS</td>
<td>Magnetic Resonance Spectroscopy</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>MVA</td>
<td>Multivariate analysis</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>nOe</td>
<td>Nuclear Overhauser effect</td>
</tr>
<tr>
<td>NOESY</td>
<td>Nuclear Overhauser effect spectroscopy</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PIRIBS</td>
<td>Parallel in-register intermolecular β-sheets</td>
</tr>
<tr>
<td>PLS-DA</td>
<td>Partial least squares discriminant analysis</td>
</tr>
<tr>
<td>PMI</td>
<td>Precision medical initiative</td>
</tr>
<tr>
<td>PRNP</td>
<td>Prion protein gene</td>
</tr>
<tr>
<td>PrP</td>
<td>Prion protein</td>
</tr>
<tr>
<td>PrPC</td>
<td>Cellular prion protein</td>
</tr>
<tr>
<td>PrPSc</td>
<td>Scrapie prion protein</td>
</tr>
<tr>
<td>RML</td>
<td>Rocky Mountain Lab adapted mouse prion</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>R₁</td>
<td>Longitudinal relaxation rate (=1/T₁)</td>
</tr>
<tr>
<td>R₂</td>
<td>Transverse relaxation rate (=1/T₂)</td>
</tr>
<tr>
<td>T₁</td>
<td>Longitudinal relaxation time</td>
</tr>
<tr>
<td>T₂</td>
<td>Transverse relaxation time</td>
</tr>
<tr>
<td>TSE</td>
<td>Transmissible spongiform encephalopathy</td>
</tr>
</tbody>
</table>
Table of contents

CHAPTER 1 – Introduction ........................................................................................................... 1

SECTION I: General introduction ......................................................................................... 1

Prions in Alberta ...................................................................................................................... 1

Prion hypothesis and where it is at now ................................................................................. 3

The evil prion: PrPSc .............................................................................................................. 6

Two topics of this thesis ......................................................................................................... 8

SECTION II: Structural study of the initial misfolding mechanism of PrP\(^{C}\) .................. 9

Brief history of the development of NMR spectroscopy ....................................................... 9

NMR application in prion studies .......................................................................................... 10

PrP\(^{C}\): back to the origin ..................................................................................................... 11

GSS disease .......................................................................................................................... 16

SECTION III: Metabolomics study of neurotoxicity during prion disease ..................... 18

Brief history of metabolomics ............................................................................................... 18

Experimental process of metabolomics ............................................................................... 19

NMR vs MS as metabolomics platform ................................................................................ 21

Applications of metabolomics .............................................................................................. 22

Metabolomics: Where we are at and where do we go ......................................................... 25

FIGURES AND TABLES ........................................................................................................ 27

REFERENCES ......................................................................................................................... 37

CHAPTER 2 – Canonical structure and structural stability of HRdup ........................... 58

OVERVIEW ............................................................................................................................. 59
INTRODUCTION .............................................................................................................. 59

MATERIAL AND METHODS ...................................................................................... 61

GSS patients and material classification, brain samples, and PRNP gene sequencing ................................................................. 61
Animal husbandry and inoculations ............................................................................. 62
Western blotting and limited proteolysis .................................................................... 62
Histopathological analysis ......................................................................................... 63
Sucrose Gradient Ultracentrifugation ......................................................................... 63
NMR spectroscopy and analysis .................................................................................. 64
Modeling of the insert structure ................................................................................ 65
Molecular dynamics simulations ................................................................................. 66
Analysis tools ............................................................................................................... 67
Essential collective dynamics ....................................................................................... 67
Protein docking ............................................................................................................ 68

RESULTS ................................................................................................................... 68

Spontaneous and transmissible prion disease in transgenic HRdup PrP mice ........... 68
Histopathology in the index case and induced by HRdup PrP transgenes ....... 70
Proteinase-K resistant signature fragments in HRdup PrP transgenic mice ......... 71
Assembly of the HRdup PrP into high molecular weight complexes ................. 72
Stability and β-sheet structure of HRdup PrP .............................................................. 73
Molecular dynamics simulations of M128V and HRdup PrP ............................ 74
A thermolysin-resistant PrP signature fragment in GSS ........................................ 76
Origins and properties of the thermolysin-resistant PrP fragment ..................... 77
List of Tables

Table 1-1. Metabolites changed in different neurodegenerative diseases................. 36

Table 2-1. Expression level, disease onset and signature PrP fragments in transgenic mice............................................................................................................................ 93

Table 4-1. Comparison of different extraction methods .......................................... 157

Table A-1. Pathology in mouse Tg.HRdup lines ..................................................... 224
Table A-2. m values Table for 5 PrP residues in a urea denaturation experiment.. 225
Table A-3. The results of SWISS-MODEL template library search with BLAST and HHBlits..................................................................................................................... 226
Table A-4. Time-averaged β-content occupancies in four MD trajectories............. 227
Table A-5. Per residue solvent accessible surface area............................................ 228
Table A-6. Total SASA for two trajectories of M128V and HRdup PrP alleles ....... 229
Table A-7. Residue pairs with robust inter-correlations for Cα atoms.................... 231
Table A-8. Residues in M128V and HRdup monomers predicted to be involved in dimer contacts........................................................................................................... 231

Table C-1. Experimental design of the number of brains harvested at each timepoint. .................................................................................................................................. 247
Table C-2. In total of 50 metabolites are identified on the spectra of all the samples .................................................................................................................................. 248
Table C-3. Two-tail t-test.......................................................................................... 249
List of Figures

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Schematic of PRNP disease-associated variants</td>
<td>27</td>
</tr>
<tr>
<td>1-2</td>
<td>The proposed four-rung β-solenoid architecture of PrP&lt;sub&gt;Sc&lt;/sub&gt;</td>
<td>28</td>
</tr>
<tr>
<td>1-3</td>
<td>The PIRIBS model for the fold of the PrP amyloid core</td>
<td>29</td>
</tr>
<tr>
<td>1-4</td>
<td>Schematic representation of the proteolytic processing of PrP&lt;sup&gt;C&lt;/sup&gt;</td>
<td>30</td>
</tr>
<tr>
<td>1-5</td>
<td>Structure of mouse prion protein (121-231)</td>
<td>31</td>
</tr>
<tr>
<td>1-6</td>
<td>Diagram indicating the octarepeat region of PrP&lt;sup&gt;C&lt;/sup&gt;</td>
<td>32</td>
</tr>
<tr>
<td>1-7</td>
<td>The three distinct binding components of the OR</td>
<td>33</td>
</tr>
<tr>
<td>1-8</td>
<td>Schematic representation of metabolomics flow</td>
<td>34</td>
</tr>
<tr>
<td>2-1</td>
<td>HRdup PrP causes GSS in transgenic mice</td>
<td>84</td>
</tr>
<tr>
<td>2-2</td>
<td>PK-resistant PrP in the index case, spontaneously sick and inoculated Tg</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>mice</td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>In vivo and in vitro properties of HRdup prion</td>
<td>86</td>
</tr>
<tr>
<td>2-4</td>
<td>NMR spectra and beta-sheet signatures of recombinant proteins</td>
<td>87</td>
</tr>
<tr>
<td>2-5</td>
<td>Molecular dynamics assessment of PrP HRdup</td>
<td>88</td>
</tr>
<tr>
<td>2-6</td>
<td>A 16 kDa thermolysin-resistant PrP signature in human brain material</td>
<td>89</td>
</tr>
<tr>
<td>2-7</td>
<td>A 16 kDa thermolysin-resistant PrP signature in mouse brain material</td>
<td>90</td>
</tr>
<tr>
<td>2-8</td>
<td>Ontogeny and properties of abnormal PrP species in Tg.HRdup mice</td>
<td>92</td>
</tr>
<tr>
<td>3-1</td>
<td>Superimposition of &lt;sup&gt;1&lt;/sup&gt;H,&lt;sup&gt;15&lt;/sup&gt;N 2D HSQC NMR spectra of V128 and HRdup</td>
<td>121</td>
</tr>
<tr>
<td>3-2</td>
<td>Methyl region of 2D &lt;sup&gt;1&lt;/sup&gt;H, &lt;sup&gt;13&lt;/sup&gt;C-HSQC NMR spectra of V128 and HRdup</td>
<td>122</td>
</tr>
<tr>
<td>3-3</td>
<td>Secondary structure of HRdup calculated by TALOS+</td>
<td>123</td>
</tr>
<tr>
<td>3-4</td>
<td>Snapshot of MD simulation trajectory of HRdup</td>
<td>124</td>
</tr>
<tr>
<td>3-5</td>
<td>&lt;sup&gt;15&lt;/sup&gt;N R&lt;sub&gt;2&lt;/sub&gt; relaxation data of HRdup and controls</td>
<td>125</td>
</tr>
<tr>
<td>3-6</td>
<td>NMR spectra of G130</td>
<td>126</td>
</tr>
<tr>
<td>3-7</td>
<td>Proposed misfolding mechanism of HRdup</td>
<td>127</td>
</tr>
<tr>
<td>4-1</td>
<td>Model of Projection</td>
<td>149</td>
</tr>
<tr>
<td>4-2</td>
<td>Geometric representation of PCA matrices</td>
<td>150</td>
</tr>
<tr>
<td>4-3</td>
<td>Number of latent variables and goodness of fit and prediction</td>
<td>151</td>
</tr>
<tr>
<td>4-4</td>
<td>The aliphatic region of the overlay of 30 high resolution 500 MHz 1D &lt;sup&gt;1&lt;/sup&gt;H-</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>NMR spectra</td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td>Unsupervised PCA model plots</td>
<td>153</td>
</tr>
<tr>
<td>4-6</td>
<td>PCA loading plot of all samples</td>
<td>154</td>
</tr>
<tr>
<td>4-7</td>
<td>The concentrations of the metabolites of different timepoints</td>
<td>155</td>
</tr>
<tr>
<td>4-8</td>
<td>Unsupervised PCA on all 34 NBH control samples</td>
<td>156</td>
</tr>
</tbody>
</table>
Fig. A-1. The cellular location of HRdup is comparable to that of M128V PrP. .......... 212
Fig. A-2. Histopathological examination of Tg.HRdup-10 and Tg.HRdup-32 animals. ................................................................. 214
Fig. A-3. Histopathological examination of Tg.M128V-39 animals. ..................... 215
Fig. A-4. Lack of myopathy and neuropathy in Tg.HRdup-26 mice. .................... 215
Fig. A-5. Brain homogenate from sick Tg.HRdup animals induces early pathological changes in mice expressing homologous PrP. .......................................................... 216
Fig. A-6. 1D NMR spectra for three PrP alleles.................................................. 217
Fig. A-7. 1D NMR spectra of urea titrations for three PrP alleles. ...................... 218
Fig. A-8. Complete 2D NMR spectrum. ............................................................ 219
Fig. A-9. Secondary structure evolution in MD simulations............................... 220
Fig. A-10. ECD correlation maps......................................................................... 221
Fig. A-11. Docking analysis.................................................................................. 222
Fig. A-12. A 16 kDa thermolysin-resistant PrP fragment is present in neonatal Tg.HRdup-26 mice. .............................................................. 223

Fig. B-1. 2D $^1$H, $^{15}$N-HSQC NMR spectrum of WT mPrP........................................ 233
Fig. B-2. 2D $^1$H, $^{15}$N-HSQC NMR spectrum of V128 mPrP.......................... 234
Fig. B-3. 2D $^1$H, $^{15}$N-HSQC NMR spectrum of HRdup mPrP ..................... 235
Fig. B-4. A strip-plot of 3D NMR experiment...................................................... 236
Fig. B-5. Diagrams of the backbone structure of the proposed type 2 $\beta$-turn ........ 237
Fig. B-6. Two ways the $\beta$1 strand can be formed............................................. 238
Fig. B-7. 1D $^1$H-NMR spectra of V128 at 500, 600 and 700MHz....................... 239

Fig. C-1. The downfield region of the overlay of 30 high resolution 500 MHz 1D $^1$H-NMR spectra.............................................................................................................. 240
Fig. C-2. The explained variance by each component in the PCA analysis........... 241
Fig. C-3. Spectra of outliers compared with other samples. A)The upper field of the spectra of sample 82806 aligned with 82805...................................................... 242
Fig. C-4. More concentrations of metabolites at different timepoints ............... 246
CHAPTER 1 – Introduction

SECTION I: General introduction

Prions in Alberta

In May 2003, a case of bovine spongiform encephalopathy (BSE), a form of transmissible spongiform encephalopathy (TSE) affecting cattle was confirmed in an Alberta herd [1]. The consequences of this BSE incident were devastating for all the beef and cattle industries throughout Canada. International trade borders were shut to Canadian beef and cattle; the total losses to the Canadian cattle industry due to the BSE crisis was conservatively estimated at about $4.2 billion [2]. With the wounds still fresh from the disastrous BSE episode, chronic wasting disease (CWD), another form of TSEs that can affect deer, elk and moose was first documented in a wild deer in Alberta in September 2005 [3]. CWD in Alberta is believed to have been spilled from the neighbouring province Saskatchewan by infected deer and it continues to make its way westward into eastern and east central Alberta. At the time of writing this thesis, the evidence of CWD transmitting into domestic livestock or humans has yet to be identified; nonetheless, as the management of CWD in the free-ranging populations of deer and elk remains challenging and the spreading continues, TSE is still a potential threat of the livelihood in Alberta.

TSEs are infectious and lethal neurodegenerative diseases with clinical symptoms and signs such as cognitive and motor dysfunction. The neuropathological features of TSEs include neuronal loss, astrocytic activation and spongiform change. The most striking feature, however, is the amyloid plaques formed by aggregates of scrapie prion protein (PrPSc). PrPSc was found out to be the misfolded isoform of the endogenous cellular prion protein (PrPC). The word “prion” comes from “proteinaceous infectious particle”; as prion plays the lead role for the infectivity and propagation of TSEs, this category of diseases is also called prion diseases.
Transmissible and non-transmissible prion diseases

TSEs were first described in sheep as early as 1732 and the animals affected exhibited abnormal behaviors such as ataxia, involuntary movements and excessive scratching, earning the disease the name “scrapie” [4]. Kuru is a form of human TSE that was first reported in the 1950s. Among members of the Fore tribe in Papua New Guinea where funerary cannibalism was practiced, there existed an endemic disease with clinical signs of cerebellar ataxia and shivering [5]. In 1959, Bill Hadlow made an important observation of the clinical and neuropathological similarities between Kuru and scrapie [6]. In later years, Kuru was transmitted to primates such as chimpanzees [7] and various laboratory rodents.

The 1980s and early 1990s witnessed the infamous BSE incident in the United Kingdom. Due to the use of meat-and-bone meals, up until year 2003, around 180,000 cows [8] were confirmed to have been infected with BSE [9] during the two decades. A total of 4.4 million cows were slaughtered during the outbreak in order to stop the spreading. This outbreak triggered a wave of concerns of BSE transmitting to humans. The concerns, unfortunately, were verified when the first cases of variant Creutzfeldt-Jakob disease (vCJD) began rising in year 1994. Numerous studies in neuropathology and biochemistry all pointed to the fact that BSE had broken the species barrier and transmitted into humans [10-12]. As of today, vCJD has claimed more than 230 victims globally [13].

Nonetheless, the human prion disease caused by exposure to prions are rare. Other than the aforementioned kuru and vCJD, the only cases of transmitted human prion diseases are iatrogenic CJD. Iatrogenic CJD is transmitted during medical procedures; it is caused by individuals receiving corneal transplant [14], dura mater implants [15], human growth hormone [16] or blood transfusions from CJD patients [17]. After care was taken to screen through the donors and including prion-reduction steps in sterilization of penetrating instruments tools, iatrogenic CJD transmission is effectively under control [18].
The most common human TSE cases, however, are sporadic and CJD accounts for more than 90% of all cases of sporadic prion diseases. The earliest recording of CJD dates back to the 1920s [13] and now CJD cases appear at a rate of approximately 1 case per million people per year [19]. Around 10% of the human prion disease cases are genetic. The patients of these cases possess an autosomal dominant mutation of prion protein gene (PRNP) which encodes PrPC on chromosome 20. While single base pair exchanges are the most frequent forms of mutations, sequence repetition or deletion also exists and these sequence alterations are associated with different phenotypes (Fig. 1-1). Based on the clinical features and the pathologic manifestations, genetic human prion diseases are broadly separated into three principal phenotypes: familial Creutzfeldt-Jacob disease (fCJD, also known as gCJD with g stands for genetic); Gerstmann-Sträussler-Scheinker (GSS) syndrome; and familial fatal insomnia (FFI) [20].

**Prion hypothesis and where it is at now**

The cause of prion disease was under debate for a long time. In 1967, Griffith first speculated that protein could act as the infectious agent that causes scrapie [21]. Prusiner and coworkers pioneered the work for the prion hypothesis by showing that the pathogen for prion disease is composed of proteinaceous material [22]; in 1982 the protease-resistant prion protein was isolated from infectious material [23] and the concentration of this protein was found to be proportional to the infectivity titer [24]. One of the strongest pieces of evidence was the finding that PrPC knock out mice were resistant to prion inoculation [25]. This crucial finding was a strong argument for the prion hypothesis, for which Prusiner was subsequently awarded the Nobel Prize in Physiology or Medicine in year 1997.

The most compelling and fundamental evidence for proving the proteinaceous nature of the prion pathogen would be to replicate mammalian prions *in vitro*. The initial success of the *in vitro* synthesis of PrP amyloid fibrils using recPrP (recombinant prion protein) was claimed in year 2004 [26] and the injection of these fibrils into transgenic mice overexpressing PrPC (truncated) seemed to have induced prion-like symptoms
after extended incubation time. However, the transgenic mice overexpressing PrP<sup>C</sup> were known to develop prion disease spontaneously [27, 28] and the fact that the same fibrils failed to induce prion disease in wild-type animals raised many concerns.

It was not until 2010 that *de novo* generation of *bona fide* infectious prions *in vitro* was achieved with recPrP [29]. The technique involved is called protein-misfolding cyclic amplification (PMCA) and it was invented as early as 2001 [30]. Most often, PMCA is performed with diluted healthy brain homogenate as substrate. However, in this work by Wang *et al.* only recPrP was used as well as RNA and lipids functioning as cofactors. PMCA was proved to be a very successful technique in generating infectious prions. One of the most interesting features with PMCA is that different prion strains in humans and mice can be replicated and maintain their features after multiple passages *in vivo* and *in vitro* [31], which had long been considered as a phenomenon difficult to explain with proteinaceous infection agent.

Using recPrP as a substrate to carry out PMCA as an ultrasensitive method for detecting PrP<sup>Sc</sup> was first achieved in year 2007 [32]. By replacing the sonication with automated tube shaking, quaking-induced conversion (QUIC) method using recPrP as substrate was invented [33]. By incorporating the thioflavin T (ThT) fluorescence into QUIC, real-time quaking-induced conversion (RT-QuIC) was created in year 2011 [34] and it is now used as the primary diagnostic tool for CJD, having replaced the conventional 14-3-3 detecting method. In year 2018, successful synthesis of sCJD with recPrP as substrate and the presence of cofactor ganglioside GM1 using QUIC was reported and these synthetic human prions are infectious to transgenic mice expressing non-glycosylated human PrP [35]. This defining piece of evidence is another important step for removing all doubts about the proteinaceous pathogen hypothesis for prion diseases.

Nonetheless, some key issues surrounding PrP<sup>Sc</sup> still remain unanswered. One of the hot topics is whether or not to consider the molecules (cofactors), which improve the propagation efficiency and infectivity of PrP<sup>Sc</sup> demonstrated by numerous studies, as an additional component of prion pathogen. In fact, both the aforementioned cases
where infectious PrPSc were achieved have employed cofactors during the synthesizing processes. There are a few reasons why the involvement of co-factors does not undermine the “protein only” hypothesis, in my opinion. To start off, there are a broad band of molecules, such as RNA, lipids and synthetic polyanion, that seem to be able to assist the PrPC to PrPSc conversion. The treatments eliminating nucleic acids, protein or lipids do not prevent prion replication in vitro [36], meaning they are not involved directly during the structural conversion. One interesting phenomenon to mention is that the same type of molecules do not seem to play the same assisting role for PrPSc amplification in different species; RNA for instance, promotes the prion replication for hamster PrP but does not seem to be effective for mouse PrP [37, 38]. Moreover, as much as cofactors catalyze the conversion, the withdrawal of cofactor does not put a definite stop to the propagation of protease-resistant PrPSc [39], which demonstrates the autocatalytic property of PrPSc propagation. The same study also reported that PrPSc conformation changes upon the withdrawal of cofactor, which potentially is the cause of the diminished infectivity of “protein-only” PrPSc compared with “protein-PE” PrPSc. The change of the conformation could be caused by the change of the environment; instead of being intertwined into the structure of infectious PrPSc, the cofactors could be simply providing an environment that is preferred for the propagation of infectious PrPSc. Most significantly, it is not unprecedented that bona fide prions were successfully generated in the absence of any cofactors; Choi et al. were able to infect mice expressing C-terminally truncated PrP (moPrP23-144) with amyloid fibrils formed with recombinant PrP (23-144) only [40].

In conclusion, the fact that some cofactors can greatly improve the infectivity and propagation of PrPSc in some studies does not contradict with the protein only prion hypothesis. The mechanism of the involvement of cofactors still needs to be further elucidated, and from that we are likely to gain some useful insight into the properties of PrPSc.
The evil prion: PrPSc

The most important piece of the puzzle in the prion study field, by some has always been considered as the ‘holy grail’ of the prion disease research field, is the structure of PrPSc. This misfolded and aggregated form of prion protein is insoluble and self-perpetuating, which means PrPSc serves as a template to conform PrPC into its own isoform to continue propagating [41]. PrPSc is proteinase resistant; treatment with proteinase K (PK) removes the N-terminal end of PrPSc and forms a molecule termed PrP27-30 due to its molecular weight and the infectious titer remains [42, 43]. Despite knowing that PrPSc is rich in β-sheet structure, due to its insolubility and high propensity towards aggregation a detailed structure of PrPSc remained unavailable until recently.

Decades of effort has been contributed to the quest of revealing the structure of PrPSc. Numerous models of the molecular architecture of PrPSc and PrP 27-30 were proposed based on experimental data of Fourier-transform infrared spectroscopy (FTIR), circular dichroism spectroscopy (CD), electron microscopy and X-ray fiber diffraction [44]. However, most of these structures fail to accommodate all experimental constraints and were eventually discarded. In recent years, we came to realize that in vitro generated PrP prions can range from being biologically inert to pathogenic, infectious even transmissible in subsequent passages, and these differences can only be correlated with conformational differences; hence, different conformation and basic multimer architecture should be expected in multiple PrPSc structures [45]. As this thesis is being written, there are two models that seem the most promising; they are 4-rung β-solenoid model and the parallel in-register intermolecular β-sheet (PIRIBS) model.

β-solenoid structure at the core of PrPSc was proposed based on electron crystallography studies of N-terminally truncated PrP 27-30 2D crystals [46]. This model was enhanced when comparing the conventional PrP 27-30 2D crystals and isomorphic 2D crystals from a sequence-edited ‘mini-prion’ (PrPSc106) [47] for the reason that traditional β-strand architectures are not compatible with the tight packing
presented in the 2D crystals. The early structural models of β-solenoid hypothesis still retained some α-helices to accommodate residual α-structures once strongly supported by the FTIR data [48, 49]; however, as the FTIR data was re-interpreted by later studies [50] α structures are no longer a requisite component in the proposed structures since.

X-ray fiber diffraction experiments on brain-derived PrP 27-30 and PrPSc yielded meridional diffraction patterns containing a series of reflection at 9.6, 6.4 and 4.8 Å, which would correspond to the second, third and fourth order signals of a 19.2 Å repeating unit [46]. This piece of evidence strongly supports a 4-rung β-solenoid fold at the core of PrPSc. Moreover, fragments of PrP can also form β-solenoid structure; the aforementioned ‘mini-prion’ was proved to have a 4-rung β-solenoid core [51]. Most recently, cryo-electron microscopy was employed to decipher the structure of PrPSc [52]. Three-dimensional reconstructions of individual PrPSc amyloid fibrils revealed that the average molecular height of each PrPSc molecule along the fibril axis is about 17-19 Å, which is consistent with the proposed model (Fig. 1-2); unfortunately the resolution of the cryo-electron microscopy is not high enough for revealing the atomic features of the fibrils. Using β-solenoid structure hypothesis, the ‘seeding activity’ of PrPSc can be explained: the upper and lowermost rungs contain ‘unpaired’ β-strands that can imprint their hydrogen-bonding pattern onto the next molecule to propagate [53] and this mechanism as well as the 4-rung β-solenoid atomic resolution model was demonstrated using molecular dynamics (MD) simulation last year [54].

The PIRIBS model was first introduced by Cobb et al. when studying the structure of amyloid composed of recombinant human PrP(90-231) using site-directed spin labeling (SDSL) and electron paramagnetic resonance (EPR) spectroscopy [55]. In this model, the residues of the C-terminal part of the protein form single molecule layer and stack on top of each other to construct a parallel, in-register alignment of β-strands (Fig. 1-3). This model has been doubted for its inability to accommodate the glycans on residue Asn 181 and Asn 197 (human PrPSc numbering) in PrPSc molecules [56]; however, it would be unwise to completely overlook this model at the current stage. In
a recent study, the structure of amyloid fibrils formed by Y145stop, a mutant of PrP\textsuperscript{C} that causes GSS disease, were shown to have a PIRIBS core structure \cite{57} and the infectivity of these amyloid fibrils were demonstrated in an aforementioned study \cite{40}. The structure of these fibrils consist of two protofilaments and the \(\beta\)-sheet region runs parallel to the long fibril axis. The structure of the propagated prions in the Y145stop amyloid that infected animals still remains unknown; there is not yet evidence that the PrP\textsuperscript{Sc} in those animals would still hold the PIRIBS structure or further adapt to \(\beta\)-solenoid structure during the propagation. It is possible that different types of infectious PrP amyloids can propagate and co-exist in the brain; and it is also very likely that PIRIBS structure and 4-rung \(\beta\)-solenoid structure can template off of each other. But just the fact that these PIRIBS amyloids are infectious suggests we should not get on a “one for all” pathway just yet, especially taking into consideration about species barriers and strain effects. Furthermore, perhaps, this is also a gentle reminder that prion-disease-causing PrP assemblies and PrP\textsuperscript{Sc} may not end up having the same meaning after all.

**Two topics of this thesis**

In this thesis, the two major components of research work are dedicated to two research topics. The first section is about the structural study of a novel GSS-causing PrP mutant revealed to have an N-terminus with increased \(\beta\)-propensity is described in Chapter 2 and 3. This is an attempt to understand the molecular mechanism of PrP\textsuperscript{C} to PrP\textsuperscript{Sc} conversion, which could provide important evidence for coming up with prevention of intervention treatments for the individuals carrying \textit{PRNP} gene mutations. Chapter 4 covers the metabolomic changes in the brains during prion disease progression. This latter section is about the attempt to understand the mechanism of the neurotoxicity of PrP\textsuperscript{Sc} using metabolomics approach. The research in this area is likely to provide some inspiration for finding an effective therapy for prion disease treatment as currently the validated pharmacological targets are extremely sparse for prion diseases \cite{58}. The discussion of experimental results, future directions of these projects as well as my interpretations are all included in Chapter 5.
SECTION II: Structural study of the initial misfolding mechanism of PrP<sup>C</sup>

**Brief history of the development of NMR spectroscopy**

Nuclear magnetic resonance (NMR) spectroscopy is one of the most powerful tools for investigating structure, molecular dynamics and interactions of biological molecules [59]. It is applied in various scientific disciplines such as chemistry, biology and medicine. Almost 70 years ago that two scientific groups discovered nuclear magnetic resonance independently [60, 61]; Felix Bloch and Edward Purcell were jointly awarded the Nobel Prize “for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith” in year 1952. Chemical shift phenomenon was first observed by Proctor and Yu when they observed that the two nitrogens in NH₄NO₃ produced two atoms [62]. In year 1953, Overhauser made the monumental discovery that the saturation of electrons in metals caused the increase of the nuclear polarization, termed “nuclear Overhauser effect” (nOe) [63]. This discovery together with the spin-spin coupling effect discovered by Freeman and Whiffen in year 1961 [64] laid out the ground work for the application of NMR in solving the structure of biomolecules.

The arrival of superconducting magnets in early 1960s revolutionized the development of NMR technique. In addition, the application of Fourier transformation into NMR technique also greatly enhanced NMR; Fourier transformed (FT) NMR [64] greatly improved the detection sensitivity by enabling the simultaneous excitation of a broadband of the nuclei of a certain isotope, if not all and converting the signals of different frequency into a spectrum. Numerous pulse sequences were invented after the advent of FT-NMR, and multi-dimensional NMR experiments were subsequently developed by Ernst et al. and became essential tools for resonance assignment [65]. Ernst was awarded the Nobel Prize in Chemistry in 1991 “for his contributions to development of the methodology of high resolution NMR spectroscopy.” The first biomolecule structure solved with NMR was in year 1985, before 3D NMR was first introduced [66]. It was the structure of bovine pancreatic trypsin inhibitor (BPTI) and the work was done in the Wüthrich lab [67]. Wüthrich was awarded the Nobel Prize in year 2002 for his contribution in protein structure determination.
NMR spectroscopy further evolved as a result of a series of technological achievements such as new superconducting material, which enabled stronger magnetic field, cryoprobe and most importantly, computer science. As of this thesis being written, a total of 12,934 biological macromolecular structures solved by NMR are deposited in Protein Data Bank (http://www.rcsb.org/pdb/). Compared with other structure solving strategies such as X-ray and cryo-EM, NMR has its unique advantages in being able to study the dynamics of molecules. The molecules are dissolved into solutions when studied by NMR, which is a state closest to the native environment and NMR permits the investigation of dynamics over a wide range of time scales. This includes protein-protein interaction, internal and external motions of molecules, backbone rigidity and pKa measurement of exchangeable groups. All this makes NMR spectroscopy an irreplaceable approach in studying biological molecules.

**NMR application in prion studies**

The great majority of structural work in the prion field was done using NMR. This preference is partially due to the fact that the first few PrP\(^C\) structures solved using NMR [68, 69], which has set the ground work for the structural study of prion protein. Later studies using NMR depicted the dynamics of PrP\(^C\) molecules in great detail and revealed various time-scale motions of the main-chain as well as the rigidity of the structured portions of PrP\(^C\) [70, 71].

NMR is also the most suitable tool for studying the structures of disease-causing PrP\(^C\) mutations, as the great majority of these mutants still possess the same overall PrP\(^C\) canonical structure yet some subtle structural and molecular dynamics changes can be caused by the substitution of certain residues. For instance, in a GSS disease related mutant Q212P, NMR revealed that in this PrP\(^C\) mutant the \(\beta_2-\alpha_2\) loop region exposes its hydrophobic side more to the solvent and that the flexibility is increased [72]. Meanwhile, the orientation of \(\alpha_2\) and \(\alpha_3\) helices is also different from the structure of WT protein; a turn of \(\alpha_3\) helix around Pro212 is observed possibly to accommodate the steric interactions of proline. Similar effects were noticed in mutant V210I [73]. In the
case of the CJD-related mutant E200K, while the canonical structure of PrP\textsuperscript{C} is largely retained, the loss of salt-bridge interaction between side chain of Glu200 and Lys204 induced the perturbation of surface electrostatic potential [74].

Other than conventional structural studies, there are other NMR-based methodologies that offered valuable information on the attributes of PrP\textsuperscript{C} molecules. One good example is the differential stability study of bPrP [75]. Using urea as denaturant, by monitoring the resonances on the spectra corresponding to the residues situated at various locations of the C-terminal globular domain, it was possible to monitor the entire unfolding process of PrP\textsuperscript{C} during urea denaturation and this study yielded accurate regional stability information of the structured portion of PrP\textsuperscript{C}. In recent years, with the advancement of technology the resolution solid-state NMR (ssNMR) has been significantly improved [76] and it has become a useful tool for elucidating the structure of PrP\textsuperscript{Sc} [57] in complementary to cryo-EM.

**PrP\textsuperscript{C}: back to the origin**

Human PrP\textsuperscript{C} (hPrP\textsuperscript{C}) is encoded by *PRNP* gene on chromosome 20. It is most highly expressed in the central nervous system (CNS) and is synthesised as a precursor protein of 253 amino acids. At the N-terminus of PrP\textsuperscript{C}, there is a signal peptide for the entry into ER, where the C-terminal signal peptide is cleaved and a glycophosphatidylinositol (GPI) anchor is added. The removal of both signal sequences produces a mature PrP\textsuperscript{C} of 208 amino acids which spans from residue 23 to 230 of the precursor protein. In Golgi apparatus, the N-linked glycans mature at Asn 181 and Asn 197 and the matured PrP\textsuperscript{C} is sent to cell surface where it is attached to the extracytoplasmic face of cell membrane by the GPI (glycosylphosphatidylinositol) anchor.

**Proteolytic processing of PrP\textsuperscript{C}**

PrP\textsuperscript{C} can be proteolytically processed in a few different ways (Fig. 1-4). The \( \alpha \) form of cleavage is an enzymatic cleavage that mostly cleave the peptidyl bond between residue
110 and 111. The two fragments produced by α cleavage are called N1 and C1 for the N-terminal fragment and the C-terminal fragment, respectively. While N1 fragment is released from the cell, the C1 fragment was shown to be trafficked to the cell membrane [77, 78]. Some studies have shown that members of the disintegrin and metalloproteinase domain-containing protein (ADAM) family are involved in the process of the α cleavage [79]. In contrast to α cleavage, which takes place inside the cell, β cleavage takes place at cell surface. The cleavage site has been described as being between the adjacent His and Gly residues of each octapeptide repeat region at the N-terminus of PrP^C, which are His 61/ Gly 62, His 69/ Gly 70, His 77/ Gly 78 and His 85/ Gly 86 in human numbering [79]. After the cleavage, the C-terminal fragment C2 retains on the cell membrane whilst the N-terminal fragment N2 is released. The β cleavage can be affected by the presence of Cu^{2+}, reactive oxygen species (ROS) and enzymatic processing by ADAM8 [79, 80]. In addition to α and β cleavage forms, PrP^C can be cleaved at 227 and 228 peptidyl bond by ADAM10, and this cleavage form cuts off the GPI anchor leading to the release the remaining PrP^C into the extracellular medium. This form of cleave is also referred to as “shedding” [81] (Fig. 1-4).

Among all the fragments generated by the physiologic processing of PrP^C, C1 is generally considered to play an inhibitory role against the formation of PrP^Sc and its shorter length determines it is resistant to misfolding [82]. The longer fragment C2, however, retains the ability to misfold and may also be resistant to α cleavage [83]. The α cleavage seems to be the most prevalent form of processing, based on the fact that C1 fragments accounts for about 50% of PrP molecules in sheep cerebral cortex [84]. As much as β-cleavage is associated with ROS indicating it can be a response to oxidative stress [85], small amount of C2 fragment is found in healthy brain tissues from various species which demonstrates β-cleavage also happens physiologically [84].

**Physiological function of PrP^C**

A full understanding of the physiological function of PrP^C is still developing. The PrP^C-knockout mice were first generated in the early 1990s, using gene targeting methods.
These animals were resistant to scrapie transmission; however, they did not present any prominent phenotypes. Given how conserved the sequence of PrP\textsuperscript{C} is among different mammalian species, this result was very surprising. Over the years, evidence showed that PrP\textsuperscript{C} has a big number of interacting partners and is involved in a big variety of cellular processes [13]. Other than the rather-well established functions of PrP\textsuperscript{C} such as copper homeostasis, stress-protection and neuronal excitability, its roles in myelination of peripheral nervous system (PNS) regulation and the modulation of the proliferation, differentiation, adhesion and morphology of cells are reported in recent studies [87]. Few studies also pointed out that PrP\textsuperscript{C}-knockout mice can present altered sleep activities, indicating the potential role of PrP\textsuperscript{C} in maintaining circadian rhythm [88]. PrP\textsuperscript{C} is proposed to be a scaffolding protein at the cell surface that assists the formation of various complexes [87]. During the progression of prion disease, however, the level of PrP\textsuperscript{C} decreases, indicating the potential “loss of function” mechanism of prion diseases [89]; this piece of evidence itself implies the importance of having a clear and specific description of the physiological function of PrP\textsuperscript{C}. A lot more experimental evidence is still needed to reach that state.

**C-terminal globular domain**

The structure of PrP\textsuperscript{C} was first determined by NMR in year 1996 [68]. The structure consists of a structured C-terminal globular domain and an intrinsically disordered N-terminus. The C-terminal globular domain of PrP\textsuperscript{C} molecule (128-231, human numbering) is comprised of a two-strand anti-parallel $\beta$-sheet and three $\alpha$-helices [69], and between residue A and B there is a disulfide bond holding the $\alpha$-helix two and three together (Fig. 1-5). The structure of C-terminal globular domain is highly conserved in different vertebrate species [90, 91]. This scaffold structure consisting one $\beta$-sheet and three $\alpha$-helices is highly consistent even among the disease-causing point mutations [92]. It has been proposed that the misfolding of PrP\textsuperscript{C} starts with the dissociation of the $\beta$-sheet [93]. This proposal was supported by bovine PrP structural stability study using NMR spectroscopy [75]. Using the same methodology, Julien *et al.* discovered that different species of PrP\textsuperscript{C} have the same unfolding mechanism when
denatured with urea; however different species have different susceptibility towards the denaturation [94]. Malevanets et al. reported similar observation that hamster and rabbit PrP displayed different susceptibility towards pH-induced aggregation [95], which is consistent with the differences in disease susceptibility of the two species. Additionally, they also discovered that the protonation of a buried histidine destabilizes both variants, more hamster than rabbit.

**N-terminus of PrP^C: more important than expected**

The N-terminus of PrP^C is intrinsically disordered and it contains an octarepeat region (OR) and a hydrophobic core (Fig. 1-6). The OR sequence, interestingly, is among the most highly conserved regions of PrP^C [96]. While the inherited prion diseases are typically caused by point mutations in the C-terminus, altered number of the octapeptide repeat sequences can also cause diseases. Normal PRNP allele contains four octapeptide repeat sequences spans from PrP 60-91 (human numbering), both fewer and higher number of the repeats would cause prion diseases [20]. The diseases caused by varied number of octapeptide repeats cover a big phenotypic range, even histopathology can vary from minimal features to widespread spongiform degeneration and PrP plaque deposits [97, 98]. Typically, the higher number of the repeats, the earlier the disease onset and the higher likelihood of observing the deposition of PrP amyloid [98].

It is well established that OR is responsible for copper binding *in vivo* [99, 100]. The PrP OR sequence is highly selective for Cu^{2+} with the binding affinity for the other metal-ion species being very weak, if not nonexistent, including Cu^{+}. The early measurements of PrP^C’s binding affinity for Cu^{2+} varied as much as eight orders of magnitude [101, 102]. Electron paramagnetic resonance (EPR) spectroscopy later revealed that the OR can rearrange itself into different configurations that have different binding affinity for Cu^{2+} in order to capture the Cu^{2+} at different concentration levels (Fig. 1-7) [103]. As a result, PrP’s binding affinity for Cu^{2+} at about 0.12 nM at low occupancy and 7-10 µM at high occupancy [104]. This affinity is compatible with
the known Cu$^{2+}$ concentrations in the synapses [105]. The role of PrP$^C$ as copper transporter or reductase seems promising when combined with the fact that PrP$^C$ is concentrated at presynaptic membranes where copper localizes and fluxes [106].

The hydrophobic core region of PrP$^C$ roughly spans from residue 110 to 135 and has been reported to adopt alternative membrane topologies associated with neurodegeneration [107]. Among the point mutations of PrP$^C$, the ones located in this region have led to some of the most drastically different downstream outcomes with some interesting twists. Within the highly conserved palindromic region (113-AGAAAGA-120), G114V mutant led to a CJD subtype with some unusual features such as only 2 glycoforms existed when subjected to pK treatment instead of the usual 3, and that the brain homogenate failed to transmit prion disease to transgenic mice overexpressing hPrP that are typically susceptible to CJD transmission [108]. A117V mutation would lead to GSS disease when valine is present at the polymorphism site 129 [109]. Most interestingly, the PrP variant G127V, exclusively conjugated with M129 appeared to play strong protective role during the epidemic of kuru in Papua New Guinea [110]; it was later demonstrated to confer strong protection against prion disease in the heterozygous state [111]. In the same study, the animals expressing only PrP V127 were completely resistant to all prion strains.

One thing to notice here is the matching effect of these mutations and the polymorphism at site 129. This is a high-frequency polymorphism site in human PRNP gene, the genotypic composition (M/M versus M/V versus V/V) [112] has profound impact on the progression and phenotype of prion diseases. This affects both the genetic diseases caused by germline mutations and infectious diseases (vCJD and iatrogenic prion diseases). If heterozygosity is at this polymorphism site (M/V), there is substantial resistance toward prion disease and the disease onset is delayed significantly [113]. The most interesting example is that D178N germline mutation can lead to genetic CJD when valine is present at site 129 on the same chromosomal allele, whilst methionine at site 129 would lead to fatal familial insomnia [114]. Collinge and coworkers reported that hPrP with either M129 or V129 appeared to possess the same
canonical PrP$^C$ structure, and they seemed to have the same molecular dynamics attribute and stability [115].

As compared with the C-terminus, the N-terminus of PrP$^C$ is not studied as extensively. This is largely due to the fact that PrP 27-30 is composed of PrP (90-231), and such evidence that animals expressing PrP depleted of the N-terminus are still susceptible to infection also had strong implications that the N-terminus seemed uncorrelated with prion diseases [47]. In addition, as intrinsically disordered molecules usually tend to fail to capture the attention of structural biologists, the unstructured portion of PrP N-terminus flew under the radar for a long time. Nonetheless, some recent studies demonstrated the role of the N-terminus in PrP trafficking and response to oxidative stress [116, 117] as well as its toxic effect during the disease progression [118]; moreover, the N-terminus also plays an important role in the interaction between prion protein and other protein oligomers such as amyloid-β and this is important for understanding the mechanism of neurodegenerative diseases [119]. In addition, the quintessential role of N-terminus in GSS disease perfectly demonstrates the importance of studying this portion of the protein.

**GSS disease**

GSS disease, a form of genetic prion diseases has a few intriguing features such as its propensity to form amyloid which can be accompanied by the formation of neurofibrillary tangles (NFTs) [120, 121]. Interestingly, these features are considered as hallmarks of Alzheimer’s Disease. What also distinguishes GSS disease from the rest of the prion diseases is its limited transmission capacity; in fact the overall experimental transmissibility of GSS to other species is low. Among all strains of GSS-associated mutations, the transmissibility of P102L strain is the most-well studied and the transmission was only successful in about one third of the cases [122]. The transmission of GSS can be improved, however, by employing animal models engineered to express the PrP allele that matches the PrP allele in the inoculum [123]. Also bank voles also seem to be a good host for GSS disease transmission [124].
One other remarkable feature of GSS disease lies in the electrophoretic profile of its pK digestion product, where a signature 6-8 kDa fragment is present in addition to the typical bands of approximately 19-21 kDa and 27-30 kDa. This signature fragment is usually comprised of the residues in the 70-150 region of PrP\textsuperscript{C} [109, 125, 126], which encompasses both the N-terminus and the C-terminus of PrP\textsuperscript{C}. Previous study in Westaway lab suggested that a 16 kDa thermolysin-resistant band present in the most common GSS strains after thermolysin digestion is very likely to be the precursor of this 6-8 kDa pK-resistant band, as the 16 kDa thermolysin-resistant signature were detected prior to the 6-8 kDa pK-resistant feature in some strains and that an \textit{in vitro} double-digested experiment produced the 8 kDa species after the thermolysin digestion [127]. In a recent study by Cracco \textit{et al.}, using mass spectrometry they discovered that the higher molecular weight pK-resistant bands in electrophoretic profiles are the covalently-linked multimers of the 6-8 kDa bands, which are exclusively composed of the central region of PrP between residues 78 and 152 [128]. With current knowledge, it remains unclear if the results of these two studies contradict with each other. The MS study indicated that the 8 kDa fragment is the building block of the pK-resistant core of the amyloids, which does not exclude the possibility that the 16 kDa thermolysin band observed could well be the immature state of the 8 kDa fragment in a dimer form.

The important message from these two pieces of evidence is that the involvement of the N-terminal and central region of PrP is critical in GSS disease. Furthermore, the elegant study by Kraus et al. revealed that individual mutations of proline 102 and proline 105 to noncyclic aliphatic residues (such as leucine, which is a GSS-associated mutant) can promote the formation of extended protease-resistant cores of PrP amyloid generated \textit{in vitro} [129]. They also demonstrated that the central lysine cluster encompassing PrP 101-110 region impedes the amyloid formation of PrP; mutating these lysine residues into alanine allowed more rapid and spontaneous formation of amyloids [129].

Combining all the above studies, the take-home message is that the N-terminal and central regions of PrP molecule seem very likely to play important role in the
initialization and the progression of GSS disease. This intrinsically disordered portion of PrP \(^\text{C}\) remains understudied and it promises to offer a new perspective for understanding the mechanism of prion diseases.

SECTION III: Metabolomics study of neurotoxicity during prion disease

As extensively as the signatures of prion diseases have been reported, such as proteinase resistant PrP deposition, spongiform change, neuronal loss and gliosis [130], the disease mechanism in the sense of how the brain is affected by the accumulation of proteinase resistant protein particles still mostly remains an enigma. The study of metabolomics is the study of the metabolism of an organism. Metabolomics reveals the fluctuations of different metabolism processes and reactions through systematic measurements of the metabolites. Metabolomics is extremely useful for probing the underlying molecular mechanisms of the diseases and understanding the clinical phenotypes. And by studying the metabolomics of the brains along the progression of prion disease, one can track the changes in neurochemical composition which would offer useful insights into the mechanism of prion disease.

**Brief history of metabolomics**

As early as 1940s, R. J. Williams *et al.* pioneered the exploration of individual metabolic patterns and reported characteristic changes in the metabolic patterns in the urine and saliva of psychiatric patients [131]. R. J. William first brought up the term “metabolic pattern” and defined it as “the total picture of the metabolism of an individual”, which is the “total chemical processes in each and every organ and tissue and their effects on each other”; instead of “the crude summations of the total energy yielding reactions” [131]. E. C. Horning *et al* originated the term “metabolic profiling” in the early 70s; they pioneered the incorporation of gas chromatography-mass spectrometry (GC-MS) to describe the metabolic profile of urine samples. Around the same time, NMR was employed to trace and identify metabolites; using \(^{31}\text{P}, \(^{13}\text{C}\) or \(^{2}\text{H}\) isotope labeled substances NMR was used to study tissue metabolites and ethanol metabolism [132, 133].
In the later decades, with the advancement of detection techniques and chemometrics as well as the emersion of bioinformatic tools such as multivariate analysis, rapid identification, measurement and analysis of numerous metabolites became possible. All of this stimulated the expeditious evolvement of metabolic profiling. The names of metabonomics and metabolomics came up almost simultaneously around year 1999. Both terms were used to define the field of characterizing the entire site of metabolites, known as the metabolome. However despite of some strenuous effort by a few to make the two terms distinguishable, till this day they are mostly used interchangeably. J. K. Nicolson et al. defined metabonomics (metabolomics) as “the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification” [134]; which essentially, defined metabolomics as a comprehensive study of metabolic changes of the entire metabolome in integrated biological systems. Joining genomics, proteomics and transcriptomics, metabolomics quickly emerged as a new field in systems biology and has become an important tool for biomarker discovery, diagnostics, toxicology studies and drug development.

**Experimental process of metabolomics**

In general, there are two approaches in metabolomics studies based on the biological knowledge to be acquired: non-targeted or metabolic profiling and targeted. Non-targeted approach is mostly employed when *a priori* knowledge of biologically interesting metabolites is unknown and that the objective is to acquire data on a wide range of metabolites in multiple metabolic classes or pathways and the changes are suspected to be all across the metabolic network among a wide range of metabolites [135]. The subsequent analysis of the data is important for providing insights into changes in the metabolome regarding the biological inquiry under investigation and this type of studies are inductive and hypothesis-generating [135]. When the object is to focus on a specific number of metabolites that are known to be relevant in a biological process and high specificity, precision and accuracy are demanded when
quantifying the concentrations, a targeted approach is usually adopted to in this scenario.

Typical metabolomics studies usually include these following components: sample collection and preparation; separation and detection of metabolites; data acquisition; and finally the analysis of data (Fig. 1-8). The samples for metabolomics studies usually include tissue extracts, urine, serum, tears saliva, fecal matter, cerebrospinal fluid (CSF) and cell culture medium [136]. The word “metabolites”, usually refers to small molecules with molecular weight smaller than 1.5kDa are extracted and separated from the biological samples into the appropriate solvent for the downstream experimental strategy. During these steps, one crucial step is to quench the metabolic activity after the collection of the samples, where the enzymatic activity is inhibited or stopped so that the metabolic flux is eliminated in the samples [135]. This step is to preserve the biologically reflective metabolome of the samples. A number of biochemical analytical platforms have been employed in metabolomics studies, such as Fourier transform infrared, Raman spectroscopy [137] and electrochemical detection [138]; the two most commonly used instruments, however, are mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. A detailed comparison of these two techniques is covered in the next section in this chapter.

Once the raw instrumental data is collected, a correction step is performed prior to the further analysis. The raw data is usually referenced to an internal standard first, then subject to transformations including normalization and scaling. Following this step, an unsupervised algorithm is usually applied first to identify the outliers, principal component analysis (PCA), a type of multivariate analysis (MVA) is a popular choice for this purpose, after which the data is ready for further analysis. A detailed description of MVA employed in this thesis is covered in Chapter 4. Building a multivariate model can be achieved from here, a process now has a popular term “machine learning”, where the model parameters, structure and included variables are estimated according to the intended purposes [135]. This is where metabolomics studies tend to end; however, in order to have an effective model that can be used to explain the metabolic
mechanism of a biological process, model validation and multiple testing must be performed.

**NMR vs MS as metabolomics platform**

The two primary holistic analytical platforms for metabolomics studies are mass spectrometry (MS) and NMR spectroscopy. MS is an analytical technique can be used to identify molecules by measuring the mass-to-charge ($m/z$) ratio of ions. MS becomes an indispensable tool for metabolomics study especially biomarker discovery when combined with liquid phase separations such as liquid chromatography or even high performance liquid chromatography (HPLC) which greatly reduce the matrix effects and the complexity of samples [139, 140]. MS has its undeniable advantage at extremely high detecting sensitivity and its extended metabolome coverage; nanomolar levels of substance can be identified and MS can work with a big range of metabolites from small soluble molecules all the way to hydrophobic protein clusters. However, MS has its drawbacks. While MS dominates the targeted metabolomics, where only one single internal standard is needed for analysis validation; for untargeted metabolomics, either the experiment needs to be repeated multiple times using aliquots of test sample with different internal standards or a combination of separation methodologies must be utilized coherently to cover the various metabolite classes [139]. In addition, the separation of metabolites in biofluids using chromatography as well as matrix effects and ion suppression continue to pose great challenges for MS to produce robust and reproducible results [141]. Together with tricky calibration process, these are the reasons why the MS results are not always interchangeable between different labs.

NMR has indisputable advantages over MS for producing quantitative and reproducible results. The highly comparable spectra can be obtained over a short period of time and with Chenomx NMR Suite Professional (Chenomx Inc., Edmonton, Canada), the spectra can be fit with the pre-recorded peak patterns of the metabolites in the database and the quantification can be achieved by referring to the introduced internal standard at a known concentration [142]. The sample preparation is relatively
simple; all it takes is to remove the protein and lipid molecules through filtration, adjust the pH and to incorporate an internal chemical shift standard. More importantly, NMR analysis is non-invasive, meaning the samples are still intact after the experiment and can still be used for other downstream analysis. The disadvantage of NMR though, lies in its inherent sensitivity. As low as micromolar concentrations are workable with NMR but anything lower would be very challenging. In addition, for metabolites such as enzymes and other hydrophobic molecules like fat and lipid molecules, NMR is rendered incompetent, as much as it is a superior tool for untargeted metabolic profiling of water-soluble small molecules.

**Applications of metabolomics**

The sum of all metabolites in the sample composes the metabolome. The levels of metabolites are the most reflective of the metabolic reactions corresponding to the physiological, developmental, and pathological state of an organism [143]; therefore the metabolomic profile is a direct display of the phenotypes of a cell or tissue responding to different genetic or environmental influences [144]. This makes metabolomics a perfect tool for characterizing the biological states of organisms and reporting the responses to stimuli such as drug therapies and nutritional changes. In the following section, the applications of metabolomics in the aspects of biomarker discovery, drug toxicity and diagnosis will be discussed.

Biomarkers are defined as “an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” [145]. Metabolomics plays an important role in biomarker discovery as well as in acquiring greater pathophysiological understanding of the onset and progression of diseases [135]. In the field of metabolomics, a biomarker may well be a combination of several metabolites that have concerted fluctuation. One example is that in cardiovascular disease field, pseudouridine and 2-oxoglutaric acid were identified with MS as potential serum metabolic biomarkers for heart failure [146]. In addition, using liquid chromatography-mass spectrometry (LC-MS) to track the changes in the lipidome, PPARg2 was shown to have significance in controlling adipose tissue expandability as
well as preventing the accumulation of fat in peripheral tissues [147]. The more challenging task, however, is to elucidate the mechanism underlying the change in the identified biomarker. For instance, in the aforementioned cardiovascular study, as much as the correlation between the two metabolites and heart failure is established, further research work is still needed to find out whether these differences are the cause or the effect of the pathophysiology of heart failure [146].

For the past decade, metabolomics has been used for identifying and assessing toxic effects during the early stages of drug screening [148]. This is due to the fact that metabolomics is rapid and it is a non-invasive way for obtaining robust and reproducible toxicological information. Drug-induced liver injury (DILI) is one of the major causes of drug withdrawal from the market. One study conducted on 20 healthy individuals administered with 3g/day of acetaminophen has demonstrated significant differences in the metabolome in urine and plasma when compared to their controls [149]. This study demonstrated how metabolomics can be used to identify biomarkers for hepatotoxicity of acetaminophen. Another common adverse effect of drugs is drug-induced kidney injury (DIKI) or drug-induced nephrotoxicity (DIN). In one study conducted on rats, the nephrotoxicity caused by 2-bromoethylamine, gentamicin, or cyclosporine A, can be indicted by three potential biomarkers, 3-methylhistidine (3-MH), 3-indoxyl sulfate (3-IS) and guanidinoacetate (GAA); these three metabolites were selected from in total of 169 metabolites detected [150].

Clinical diagnostics, prognostication and the monitoring of diseases progresses is perhaps some of the most exciting aspects in the application of metabolomics. Promising results have been achieved in a variety of disease settings [151]. Oncology is a hot area to apply metabolomics for diagnosis. With MS of fine-needle aspirate samples, an accuracy of 93% was achieved in distinguishing the benign lesions from cancerous tissue [152]. Another study employed magic angle spinning (MAS) NMR spectroscopy to predict histological grade, immunohistochemical status as well as lymphatic spread; the sensitivity and specificity ranged from 83 to 100% [153]. Similar effective distinctions between the patients and healthy controls also happen in
pancreatic ductal cancer [154, 155], colorectal cancer [156], lung cancer [157] and genitourinary cancer [158] etc. For other disease areas, a MS-based study reported that five branched-chain and aromatic amino acids, isoleucine, leucine, valine, tyrosine and phenylalanine had highly significant associations with diabetes [159]. A NMR comparative study between 197 adults with chronic obstructive pulmonary disease and 195 healthy individuals revealed that the metabolites trigonelline, hippurate and formate are associated with baseline lung function [160].

**Metabolomics in neurodegenerative diseases**

In the neurodegenerative diseases realm, numerous metabolomics studies were carried out on diseases such as Alzheimer’s, Parkinson’s, Huntington’s disease and amyotrophic lateral sclerosis. Biomarker identification was proved to be extremely challenging in this area due to the heterogeneity in pathology and clinical presentations of neurodegenerative disorders [161]; however, with unbiased strategies such as metabolomics we can gain a better understanding of the metabolism states of these disorders which would offer helpful insight in understanding the mechanisms of these devastating diseases.

Among all these neurodegenerative diseases, combining years of metabolomics studies in this area we now have a long list of various metabolites that were demonstrated to have varied the most in each of these disorders. Unfortunately these metabolites overlap and co-occur, which makes it very challenging to find valuable metabolic biomarkers for any given neurodegenerative diseases [144]; moreover, different studies of the same disease using the same analytical platform could also yield different metabolites that appear to have changed the most during the disease progression. In Table 1-1 listed the metabolites identified to have changed the most in different studies of neurodegenerative diseases.

A few factors seem to be contributing to the relatively large variations among different studies. This variation starts from the selection of clinical cohort; patients of possible comorbidities and different clinical stages could potentially be chosen and such
heterogeneity from the origin could potentially render the downstream analysis dreadfully ineffectual. In addition, most of the metabolomics studies are simple comparison studies of sick vs healthy, whereas longitudinal studies would better map out the changes in metabolism during the disease progression as well as be more convincing in identifying metabolites that potentially can be used as signatures of the neurodegenerative disorder, if not biomarkers. Last but not least, same as metabolomics studies in other research disciplines, the lack of standardization of sample processing, data acquisition and processing heavily jeopardizes the chances of making meaningful discoveries using the metabolic profiling method in the neurodegenerative disease research area. There is still a long way ahead of us for metabolomics study to become an effective and essential tool for neurodegenerative disease research area.

**Metabolomics: Where we are at and where do we go**

While a growing body of work has contributed to disease diagnosis using metabolomics with increasing quantity of exceedingly promising results, this technique is still at its preliminary stage and still far from being included as part of routine clinical practice. And at the current stage, the biggest challenge is the lack of standardization which makes integrating all the metabolic datasets from different sources very nearly impossible. More than 200 labs world-wide carry out metabolomics research; all these labs are more or less operating with different metabolite identification and quantification standards as well as different experimental designs and data analysis tools [135]. The lack of standardized protocols in steps such as sample collection, sample storage condition and transportation condition only further exacerbates the problem. Without being able to even compare the results from different labs directly, it is hard for the whole study field to gain a momentum of moving forward as huge amount of time and resources are wasted in the process of verifying and repeating works from other labs.

A great amount of effort has already been contributed to solving these issues. Online databases such as Human Metabolome Databank (HMDB) [162], BioMagResBank
(BMRB) [163] and the Madison metabolomics Consortium Database (MMCD) [164] containing even higher dimensional NMR spectral information have been established and similar efforts were also made by scientists using MS to do metabolomic research [165]. In 2005, the Metabolomics Standards Initiative (MSI) comprised of international volunteer researchers took over the role for developing and communicating standards for the whole community. In the following years, papers regarding requirements for biological samples [166, 167], chemical analysis [168], data analysis [169] and data exchange [170] were published and they will undoubtedly benefit the entire metabolic community in the most effective way. Moreover, with the development of platforms such as Metabolights [171], Metabolomics Workbench [172], where researchers can provide the original data for all colleagues to evaluate and analyze, metabolomics research is heading towards a promising direction with the improved transparency.

In 2016, the White House announced the precision medicine initiative (PMI) in the US, stepping into a new era of individualized health care [173]. Metabolomics was put forward as an important component for the development of PMI due to its unique tight relations with phenotype outcomes [174]. This demonstrates the crucial role of metabolomics research in the future, as well as the urgency to improve this field for it to fill in its role in monitoring human health conditions and other clinical functions such as diagnosis. It will be a challenging and exciting journey.
**Fig. 1-1.** Schematic of PRNP disease-associated variants. Mutations are color coded based on different types of human prion disease. This figure was originally published by Kim et al. [175]. Permission to reuse granted.
Fig. 1-2. The proposed four-rung $\beta$-solenoid architecture of PrP$^{\text{Sc}}$. (A) three-dimensional reconstruction of a PrP$^{\text{Sc}}$ amyloid fibril with two protofilaments. (B) Cartoon representation of a four-rung $\beta$-solenoid architecture approximating the 3D reconstruction. This figure was originally published by Wille et al. on Pathog. 2018 [53]. Permitted to reuse.
Fig. 1-3. The PIRIBS model for the fold of the PrP amyloid core (residues 159-219). This figure also presents the in-register stacking motif of nearby monomers with tight interdigitation of the side chains as well as the intermolecular hydrogen bonding. The arrow indicates the long axes of the fibril. This figure was published by Cobb et al. [55]. Copyright 2007 National Academy of Sciences.
Fig. 1-4. Schematic representation of the proteolytic processing of PrP\textsuperscript{C}. The $\alpha$-cleavage site, $\beta$-cleavage site and the shedding site are indicated. This figure was published by Castle \textit{et al.} [87]. Permitted to reuse.
Fig. 1-5. Structure of mouse prion protein (121-231). The three helices are indicated in yellow, the antiparallel β-sheet is in cyan and the connecting loops are displayed in green. The disulphide bond between helix 2 and 3 is indicated in white. This figure was published by Riek et al. [68]. Permission to reuse granted.
Fig. 1-6. Diagram indicating the octarepeat region of PrP<sup>C</sup>. Normal PRNP allele contains four octapeptide repeat sequences spanning from residue 60-91. This figure was published by Glenn L. Millhauser [106]. Permission to reuse granted.
Fig. 1-7 The three distinct binding components of the OR. Low Cu $^{2+}$ occupancy favors component 3 whilst the high occupancy favors component 1. This figure was published by Glenn L. Millhauser [106]. Permission to reuse granted.
Fig. 1-8. Schematic representation of metabolomics flow. This figure was published by Botas et al. [144]. Permission to reuse granted.
<table>
<thead>
<tr>
<th>Disorder</th>
<th>Sample origin</th>
<th>Analytical platform</th>
<th>Altered metabolites</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>TgCRND8 mice and non-transgenic mice</td>
<td>$^1$H NMR 700MHz</td>
<td>Ala, Asp, Cho, Creatine, Free Fatty acids, GABA, Gln, Gly, Ile, Leu, NAA, PCho, Phosphocreatine, Succinate, Tau, Val</td>
<td>[176]</td>
</tr>
<tr>
<td>AD</td>
<td>Human brain tissue</td>
<td>$^1$H NMR 400MHz</td>
<td>Ala, Acetate, Asp, Carnitine, Cho, Creatine, Ethanoamine, Glu, Gln, Gly, Hypoxanthine, Inosine, Ileu, Lac, Leu, myo-inositol, NAA, Nicotinate, Phe, Ser, Succinate, Tau, Tyr, Val</td>
<td>[177]</td>
</tr>
<tr>
<td>AD</td>
<td>AβPPswe Tg2576 Mouse Model and controls mice</td>
<td>$^1$H NMR 400MHz</td>
<td>Cr, GABA, GPC, NAA, Tau</td>
<td>[178]</td>
</tr>
<tr>
<td>AD</td>
<td>Transgenic APP/PS1 mice brain and WT control mice</td>
<td>ESI(+/−)-QTOF-MS</td>
<td>Acetate, Acylcarnitines, Amino acids, Cho, Dopamine, Fatty acids, Glycerol-3-phosphate, GPC, GPI, GSH, Homocarnosine, Lysophospholipids, N1-acetylspermidine, Nucleotide metabolism metabolites, Phospholipids, Propionate, Pyruvate, Steroids, Tyr, Urea</td>
<td>[179]</td>
</tr>
<tr>
<td>AD</td>
<td>Human brain tissue</td>
<td>LC-ESI(+/−)-MS HPLC-UV</td>
<td>Agmatine, Arg, Citrulline, GABA, Gln, Glu, Ornithine, Putrescine, Spermidine, Spermine</td>
<td>[180]</td>
</tr>
<tr>
<td>AD</td>
<td>Human brain tissue</td>
<td>LC-ESI(+/−)-QTOF MS/MS</td>
<td>Deoxyguanosine, dGMP, Gly, Guanine, IDP, Xanthosine</td>
<td>[181]</td>
</tr>
<tr>
<td>AD</td>
<td>APP/PS1d E9 mice and WT C57BL/6J mice</td>
<td>UHPLC-QqQ MS</td>
<td>Acylcarnitine, Arg, Citrulline, Creatinine, Glu, Gln, Gly, Lysop-Heho, Phe, Pro, Putrescine, Ser, Spermidine, Spermine, Sphingolipids, t4-hydroxyproline, Thr, Tryptophan, Tyr</td>
<td>[182]</td>
</tr>
<tr>
<td>AD</td>
<td>Human brain tissue</td>
<td>UHPLC-ESI(+/−)-MS/MS</td>
<td>Lac, Phe</td>
<td>[183]</td>
</tr>
<tr>
<td>AD</td>
<td>APP/PS1 mice grain and WT mice</td>
<td>UHPLC-ESI(+/−)-QTOF MS GC-EI-ITQ MS</td>
<td>2-Hydroxyglutarate, Adenine, Adenosine, AMP, Asp, cAMP, cGMP, Cholesterol, Cho, Citrate, Creatinine, Dopamine, Ethanolamine, Glucose-6-P, Glu, Glycerol, Glyceraldehyde-3-phosphate, GPC, Gly, Guanosine, Histidine, Hypoxanthine, Inosine, Lac, Lactose, Malate, myo-inositol,</td>
<td>[184]</td>
</tr>
<tr>
<td>Disease</td>
<td>Model</td>
<td>Instrumentation</td>
<td>Detection</td>
<td>Metabolites detected</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-----------------</td>
<td>-----------</td>
<td>----------------------</td>
</tr>
<tr>
<td>AD</td>
<td>Human brain tissue</td>
<td>$^1$H NMR 500MHz</td>
<td>AD: Ala, Acetate, Gln, Glu ALS: Cho, Creatine, Lac, myo-inositol, NAA, PC</td>
<td>[185]</td>
</tr>
<tr>
<td>ALS</td>
<td>Sprague Dawley rats</td>
<td>HR-MAS $^1$H NMR 600 MHz</td>
<td>Acetate, Asp, Cho, Creatine, GABA, Glycrol, Lac, Methyl lipids, Methylene lipids, myo-inositol, NAA, NAAG, PCho, PC, Succinate, Tau</td>
<td>[186]</td>
</tr>
<tr>
<td>HD</td>
<td>R6/2 transgenic mice and WT mice</td>
<td>HR-MAS $^1$H NMR 600 MHz</td>
<td>Acetate, Ala, Asp, Cho, Creatine, GABA, Glu, GPA, Lac, myo-inositol, NAA, NAAG, PCho, scyllo-inositol, Tau</td>
<td>[187]</td>
</tr>
<tr>
<td>HD</td>
<td>R6/2 transgenic mice and WT mice</td>
<td>MRS</td>
<td>Ala, Ascorbate, Asp, Creatine, Phosphorcreatin, GABA, Glucose, Gln, Glu, GSH, myo-inositol, Lac, NAA, NAAG, PE, Tau, GPC + PC</td>
<td>[188]</td>
</tr>
<tr>
<td>HD</td>
<td>Human brain tissue</td>
<td>GC-TOF-MS</td>
<td>Urea</td>
<td>[189]</td>
</tr>
<tr>
<td>HD</td>
<td>Sprague Dawley rats</td>
<td>GC-TOF-MS</td>
<td>Arachidonic acid, Ascorbate, Asp, Citrate, Docosahexaenoic acid, Glycerol, Gly, Gln, Met, Methylmalonate, Ser, Succinate, Thr, Oleamide, Palmitate, urea</td>
<td>[190]</td>
</tr>
<tr>
<td>PD</td>
<td>$\alpha$-synuclein knockout A53T-tg positive mice and transgene-negative littermates</td>
<td>$^1$H NMR 400 MHz LC-Q/Orbitrap-MS</td>
<td>$5'$-methylthioadenosine, ADP, AMP, Cho, Creatine, Creatinine, Fumarate, Glu, Gln, GDP, IMP, Lac, myo-inositol, NAA, Oxidized GSH, Oxidized nicotinamide, Phosphorylcholine, Pyroglutamat, Succinate, Tau</td>
<td>[191]</td>
</tr>
<tr>
<td>PD</td>
<td>Parkin knockout mice, toxin-induce mice and WT mice</td>
<td>UHPLC-ESK(+) QqQ-MS</td>
<td>$\alpha$-ketoregulate, ADP, AMP, ATP, Fuctose-6-Phosphotate, Fumarate, Glucose, Glucose-6-P, Glu, Gln, Lac, Malate, NADPH, Phosphoenolpiruvate, Pyruvate, Ribose-5-Phosphate, Succinate</td>
<td>[192]</td>
</tr>
</tbody>
</table>

REFERENCES


Chapter 1- Introduction


Chapter 1 - Introduction


[70] D. B. O'Sullivan et al., "Dynamics of a truncated prion protein, PrP(113-231), from (15)N NMR relaxation: order parameters calculated and slow


Chapter 1- Introduction


Chapter 1- Introduction


[126] P. Piccardo et al., "Prion proteins with different conformations accumulate in Gerstmann-Sträussler-Scheinker disease caused by A117V and F198S


[152] C. E. Mountford *et al.*, "Diagnosis and prognosis of breast cancer by magnetic resonance spectroscopy of fine-needle aspirates analysed using a statistical


CHAPTER 2 – Canonical structure and structural stability of HRdup

A novel Gerstmann-Sträussler-Scheinker disease mutation defines a precursor for amyloidogenic 8 kDa PrP fragments and reveals N-terminal structural changes shared by other GSS alleles

Robert C.C. Mercer1, 2, Nathalie Daude1, Lyudmyla Dorosh3, 4, Ze-Lin Fu1, 5, Charles E. Mays1, Hristina Gapeshina1, Serene L. Wohlgemuth1, Claudia Acevedo-Morantes1, Jing Yang1, Neil R. Cashman6, Michael B. Coulthart7, Dawn M. Pearson8, Jeffrey T. Joseph9, Holger Wille1, 5, Jiri G. Safar10, Gerard H. Jansen7, 11, Maria Stepanova3, 4, Brian D. Sykes1, 5, and David Westaway1, 2, 5*

1Centre for Prions and Protein Folding Diseases, University of Alberta, Edmonton, Alberta, Canada
2Department of Medicine (Neurology), University of Alberta, Edmonton, Alberta, Canada
3National Research Council of Canada, Edmonton, Alberta, Canada
4Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada
5Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada
6Brain Research Centre, University of British Columbia, Vancouver, British Columbia, Canada
7Canadian Creutzfeldt-Jakob Disease Surveillance System, Centre for Foodborne, Environmental and Zoonotic Infectious Diseases, Public Health Agency of Canada, Ottawa, Ontario, Canada
8Department of Clinical Neurosciences, University of Calgary, Calgary, Alberta, Canada
9Hotchkiss Brain Institute and Calgary Laboratory Services, University of Calgary, Calgary, Alberta, Canada
10Departments of Pathology and Neurology, School of Medicine Case Western Reserve University, Cleveland, Ohio, USA
11Division of Anatomical Pathology, University of Ottawa, Ottawa, Ontario, Canada
Chapter 2 – Canonical structure and structural stability of HRdup


Contributions: ZLF synthesized the recombinant protein material used in the NMR experiments in this study. ZLF and BDS designed and performed the NMR experiments as well as analyzing the data; ZLF made Fig. 2-4 A), B), D) and wrote the “Stability and β-sheet structure of HRdup PrP” portion of the result section and NMR portion in material and method section with BDS and DW.

OVERVIEW
To explore pathogenesis in a young Gerstmann-Sträussler-Scheinker Disease (GSS) patient, the corresponding mutation, an eight-residue duplication in the hydrophobic region (HR), was inserted into the wild type mouse PrP gene. Transgenic (Tg) mouse lines expressing this mutation (Tg.HRdup) developed spontaneous neurologic syndromes and brain extracts hastened disease in low-expressor Tg.HRdup mice, suggesting de novo formation of prions. While Tg.HRdup mice exhibited spongiform change, PrP aggregates and the anticipated GSS hallmark of a proteinase K (PK)-resistant 8 kDa fragment deriving from the center of PrP, the LGGLGGYV insertion also imparted alterations in PrP’s unstructured N-terminus, resulting in a 16 kDa fragment following thermolysin exposure. This fragment comprises a plausible precursor to the 8 kDa PK-resistant fragment and its detection in adolescent Tg.HRdup mice suggests that an early start to accumulation could account for early disease of the index case. A 16 kDa thermolysin-resistant signature was also found in GSS patients with P102L, A117V, H187R and F198S alleles and has coordinates similar to GSS stop codon mutations. Our data suggest a novel shared pathway of GSS pathogenesis that is fundamentally distinct from that producing structural alterations in the C-terminus of PrP, as observed in other prion diseases such as Creutzfeldt-Jakob Disease and scrapie.

INTRODUCTION
Gerstmann-Sträussler-Scheinker Disease (GSS) is an autosomal dominant multisystemic neurological syndrome that may evolve to frank dementia and often exhibits
a protracted clinical course[1]. This inherited amyloidosis is caused by a number of mutations in the human gene, PRNP, on chromosome 20 that encodes the cellular prion protein (PrP<sub>C</sub>) [2]. PrP<sub>C</sub> is a GPI-linked protein displayed on the cell surface; it’s N-terminus is natively unstructured and contains tandem metal binding octarepeats and a hydrophobic region (HR). PrP<sub>C</sub>’s C-terminal region folds into a globular three-helix bundle with a small two-stranded β sheet [3]. In infectious prion diseases, PrP<sub>C</sub> refolds to a pathogenesis-associated form, PrP<sub>Sc</sub>, whose structure is dominated by β sheet [4-6].

GSS is of interest as there is a propensity to form amyloid, sometimes accompanied by formation of neurofibrillary tangles (NFTs) [1, 7] and because there is only a limited capacity for the generation of infectious titre and/or ability to transmit to recipient species [8] - these properties thus offering a partial parallel to Alzheimer's Disease (AD), where Aβ amyloid and NFTs are pathologic hallmarks in idiopathic and genetic disease [9] and where transmissibility of clinical disease to animal recipients is not a hallmark [10]. With poor GSS transmission first shown for experiments using non-human primates [11], transmissions into non-Tg rodents are similarly inefficient and stand in contrast to Creutzfeldt-Jakob Disease (CJD) [12]. This inefficiency may be partly overcome by using mice engineered to match the donor PrP allele [13, 14], or by using bank voles [15], which are promiscuous hosts for a number of prion diseases [16-20]. Clinical presentation of GSS can be variable and partially dependent on the causative mutation in the PRNP gene coding region, the most common clinical phenotypes are cerebellar ataxia and pyramidal signs with eventual cognitive decline before death [21]. Multicentric plaques composed of truncated PrP fragments are often found in the brains of these patients and spongiosis may or may not be present [22].

A notable PRNP mutation was recently discovered in a young GSS patient. The index case presented with status epilepticus at age 34, prefaced by night terrors at age 26. While the parasomnias subsided after a 6-year period, memory problems and behavioural changes emerged at this age [23]. PRNP sequencing revealed homozygosity for valine at the polymorphic residue 129 and heterozygosity for a
partial internal duplication, resulting in a protein with 8 extra residues. This LGGLGGGYV insertion lies at the junction between the HR - the most conserved area of PrP and one reported to adopt alternative membrane topologies that are associated with neurodegeneration [24-27] - and the globular domain, where it is predicted to duplicate some of the residues found in the first β-strand of PrP (residues 128-131 [28]). Prompted by findings that early structural rearrangements in the PrPC to PrPSc transition involve the HR [29] and the possibility of discovering a process underlying an aggressive form of GSS disease, we pursued animal modeling. Our studies define a novel misfolded form of mutant PrPc that prefigures the 8 kDa PrP fragment pathognomonic for end-stage GSS with multicientric amyloid plaques [30] and may be shared by other allelic forms of GSS; they also suggest an explanation for the early disease onset of the index case.

MATERIAL AND METHODS

GSS patients and material classification, brain samples, and PRNP gene sequencing

DNA was extracted from frozen brain tissues in all cases, and genotypic analysis of the PRNP coding region was performed as described [31-33]. On the basis of diagnostic pathology, immunohistochemistry, and western blot examination of 3 brain regions (including frontal, occipital and cerebellum cortices) with mAb 3F4 and 1E4, the pathogenic PrPSc was classified as described previously [30, 34-37]. Coronal sections of human brain tissues were obtained at autopsy and stored at -80ºC. Three 200-350 mg cuts of frontal (superior and more posterior middle gyri) cortex were taken from each brain and used for molecular analyses. The other symmetric cerebral hemisphere was fixed in formalin and used for neuropathological classification of prion disease using histological and immunohistochemical analysis of samples from 16 anatomical areas and NPDSC’s standard protocols [35, 38, 39]. We based the classification on the molecular characteristics of PrPSc on western blots developed with a panel of antibodies as described previously [35, 40, 41] to exclude GSS cases with a 21 kDa PK-resistant fragment. This criterion, as well as DNA sequencing, allowed the classification of the included cases with pathognomonic 8 kDa fragments as follows:
A117V-129V (age 35, disease duration 60 months), P102L-129M (age 37, disease duration 106 months), H187R-129V (age 42, disease duration 108 months) and F198S-129V (age 59, disease duration 120 months). For the index case, we used frozen postmortem tissue from different brain areas, as described in the Figure legends.

**Animal husbandry and inoculations**

Transgenic mice were generated using a modified half-genomic construct and standard procedures [42, 43]. Animals were housed in groups of up to five under a 12 h light/dark cycle with food and water *ad libitum*. Tg.CRND8, Tg.TauP301L and tga20 mice have been described previously [42, 44, 45]. Inoculations were performed by intracerebral injection of 30 μl of 1% (wt/vol) brain homogenate.

**Western blotting and limited proteolysis**

Brain hemispheres were homogenized in cold PBS by passage through successively larger gauge needles. Whole brain extract was subjected to 10% Tricine-SDS-PAGE and transferred to PVDF membranes (Millipore) using Tris-gly transfer buffer with 20% methanol in the Mini Trans-Blot Electrophoretic Transfer Cell (Biorad) or the XCell Blot Module (Invitrogen). Primary antibodies used were: Sha31 (α-PrP; Spi-bio), 12B2 (α-PrP; from Dr. J. Langeveld) and PrP248 and VRQ61 (α-PrP; from Dr. H. Rezaei). Secondary antibodies used were horseradish peroxidase conjugated goat α-mouse (Bio-Rad). For enzymatic digestion, 250 μg (PK) or 50 μg (thermolysin) of protein was incubated for 1 h at 37 °C (PK) or 70 °C (thermolysin) with 10 μg/ml PK (Roche) or 50 μg/ml thermolysin (Sigma) in 250 μl. For sequential digests, samples were methanol precipitated after thermolysin treatment, re-suspended in PBS and then digested with PK. 5 mM PMSF (PK) or 10 mM EDTA pH 8 (thermolysin) was used to stop the reaction and samples were centrifuged at 20,800 x g for 1 h at 4 °C. Pellets were resuspended in sample buffer containing 50 mM DTT. For removal of carbohydrates, 20 μg of protein was incubated overnight at 37 °C with 100 U PNGase F (New England Biolabs) in a volume of 20 μl according to manufacturer's instructions. Molecular weight markers are SeeBlue Plus 2 pre-stained standards (Invitrogen).
Histopathological analysis

Mice: Sagittal sections were fixed in 10% phosphate buffered formalin and embedded in paraffin. Hematoxylin and Eosin staining was done as previously described [46]. For immunodetection, slices were heated to 121 °C in 10 mM citrate buffer and allowed to cool to room temperature. Staining for PrPSc was then accomplished by treatment with formic acid and 4 M guanidine thiocyanate before an overnight incubation with biotinylated SAF83 (Cayman Chemicals) or PrP248. GFAP immunodetection was accomplished by subsequent incubation in 3% peroxide and overnight incubation with a biotinylated primary antibody cocktail (BD Biosciences; 556330). These slices are counterstained with Mayer’s hematoxylin. Index case: Brain tissue was obtained at autopsy 21 hours after death. The tissue was partly frozen at -85 °C and partly fixed in buffered formalin. The tissue was processed to blocks and embedded in paraffin 14 days after autopsy. Five micrometer sections were stained with hematoxylin and eosin, Periodic Acid-Schiff, and immunohistochemistry was performed for prion protein (12F10 1:3000 Cayman), amyloid beta (Anti B-Amyloid 17-24 (4G8) 1:20000 BioLegend), tau (AT8, 1:2500 Leica), and P62 (Anti-SQSTM1 1:200 Abcam).

PTA precipitation

An equal volume of 4% Sarkosyl was added to 1 mg of a 20% brain homogenate (PBS) and homogenized by passage through a 25g needle. A stock solution of PTA was added to have a final concentration of 2% PTA and the sample was allowed to incubate at 37 °C for 16 h with 1200 rpm shaking [47]. The sample was centrifuged at 16,000 x g for 30 minutes and the supernatant was removed. The pellet was resuspended in 50 μl 0.2% Sarkosyl and then brought to 1000 μl using 2% Sarkosyl. PTA was again added to a final concentration of 2% and the sample was incubated at 37 °C for 1 h. Following a second 16,000 x g spin, the pellet was resuspended in 0.2% Sarkosyl for downstream analysis.

Sucrose Gradient Ultracentrifugation

Linear 10-45% sucrose gradients were prepared by layering 375 μl of increasing concentrations (5% steps) of sucrose (in PBS, pH 7.4 and 1% Sarkosyl) in OptiSeal™
polypropylene tubes (Beckman Coulter). Gradients were linearized by incubation overnight at 4 °C. 250 μg of brain homogenate was brought to 300 μl in PBS (pH 7.4) containing 2% Sarkosyl and layered on top. Samples were centrifuged at 268,000 x g for 73 minutes at 4 °C using a swinging bucket rotor and eight fractions were collected from the bottom of the tube. Equivalent volumes of each fraction were then interrogated for the presence of PrP by western blot.

**NMR spectroscopy and analysis**

The genes moPrP_{118-231}, moPrP_{118-231} M128V, and moPrP_{118-231} HRdup, were synthesized by DNA2.0 with codon optimization. The N terminus of all three proteins had a 6xhis tag with a TEV cleavage site for removal. Protein sequences are listed below:

**Mouse PrP(118-231) M128V HRdup, 143aa**

HHHHHHHGASTGQMQGENLYFQGAVVGGGLGGYVLGGLGGYVLGGLGSAMSRPMIHFGNDWEDRYYRENMYRPNQVYYRPVDQYSNQNNFVHRCCNVITIQKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYKESQAYYDGRRSSG

**Mouse PrP(118-231) M128V, 135aa**

HHHHHHHGASTGQMQGENLYFQGAVVGGGLGGYVLGGLGSAMSRPMIHFGNDWEDRYYRENMYRPNQVYYRPVDQYSNQNNFVHRCCNVITIQKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYKESQAYYDGRRSSG

**Mouse PrP(118-231) WT, 135aa**

HHHHHHHGASTGQMQGENLYFQGAVVGGGLGGYMLGSGSAMSRSMPMIHFGNDWEDRYYRENMYRPNQVYYRPVDQYSNQNNFVHRCCNVITIQKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYKESQAYYDGRRSSG

Vectors were transformed into BL21(DE3) cells. The proteins were expressed and purified based on the methodology previously described for the expression of human PrP [48]
For urea denaturing experiments, the samples were prepared the same as described previously [2]. Whilst for two-dimensional (2D) $^1$H-$^1$H NOESY experiment, 2 mg protein was dissolved in 500 μl solution that was made up with D$_2$O (99.9%, Cambridge Isotope Laboratories), 10 mM sodium acetate, and 0.3 mM DSS-d$_6$; pD was adjusted to 5.3 using 2 M DCl. The sample was flash frozen, lyophilized and re-suspended using 500 μl D$_2$O (99.996%, Cambridge Isotope Laboratories) to minimize the H$_2$O signal within the spectra thereby providing a clear view of β-sheet-correlating signals.

NMR experiments were performed on an 800 MHz Varian INOVA NMR spectrometer at 25 °C. One-dimensional (1D) proton NMR spectra were acquired for urea denaturing experiments with 256 transients, with a spectral width of 15.0 ppm and a time delay of 2.5 s; while two-dimensional (2D) $^1$H-$^1$H NOESY spectra were acquired with 64 transients, 768 increments, spectral widths of 12.0 ppm for both proton dimensions. The time delay was 1.5 s and the mixing time was 100 ms. All NMR spectra were processed with VnmrJ software v4.0 (Varian Inc.) and line broadening of 0.5 Hz was applied. Further data processing and analysis for both urea denaturing experiments and 2D experiments are the same as previously described [49].

**Modeling of the insert structure**

The insert model HRdup was built using the SWISS-MODEL homology modelling server [50]. The moPrP 23-239 sequences with the M128V polymorphism were modified by insertion of a LGGLGGYV sequence between V128 and L129 and uploaded to SWISS-MODEL. The server template search and alignment with BLAST and HHBlits software [51, 52] were performed. Direct insertion of a LGGLGGYV between V128 and L129, or GLGGYVLG between G130 and S131 with target-template alignments were also attempted. Those constructs tended to exhibit similar conformations as in the model we adopted for analysis (S3 Table). Both the global and per-residue model quality were assessed using the QMEAN scoring function [53]. Two scores were evaluated: the global model quality estimation (GMOE; scores closest to one indicate the highest quality) and score composite scoring function to estimate the
global and local model quality (QMEAN4; the highest negative scores indicate a higher local per-residue reliability of the model). The homology search for mouse PrP sequence with insert produced 285 templates, from which 10 were chosen to build models. These included PDB ID codes 4MA8, 4MA7, 2L39, and 2L1H (MoPrP) and 4KML, 1QLZ (HuPrP), see S3 Table. The best matching models chosen based on the range, sequence identity, and coverage, included constructs 2L39 with identity 98.25%, GMQE score 0.41 and QMEAN4 score -4.03; 2L1H with identity 99.11%, GMQE score 0.45, and QMEAN4 score -1.94; and 4MA7 with identity 99.12%, GMQE score 0.46 and QMEAN4 score -1.56. The coordinates of the eight amino acid insert model based on the 2L39.pdb construct were uploaded and prepared for the simulations. The best-matching 3D model of HRdup constructed by the SWISS-MODEL tool initially represented a L124-D226 sequence. The N-terminal fragment G89-G123 was added to this initial model using the Accelrys VS software (Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, Release 2017, San Diego: Dassault Systèmes, 2016). The choice of pdb structure and the length of the sequence for the control M128V structure without the insert were based upon the best-matching model 2L39.pdb. C-terminal fragment 227-232 was removed to match the HRdup model, and fragment 89-117 and polymorphic mutation M128V were introduced using the Accelrys VS. In all simulation runs, the C-terminal and N-terminal extremities of the main chains were kept charged (−COO- and −NH3+), whereas all other titratable amino acids were assigned their canonical state at pH 4.5 with the PropKa server software [54].

**Molecular dynamics simulations**

The HRdup and M128V constructs were subjected to minimizations, equilibrations and production molecular dynamics (MD) simulations in Gromacs v 4.5.3 package with OPLS forcefields [55, 56]. Starting models were minimized in vacuo for 10000 steps of steepest descent minimization. Then the models were solvated in single point charge extended (SPC/E) rectangular periodic water box, after which Cl− or Na+ ions were added to neutralize the systems. Subsequent solvent minimizations with decreasing position restraints ($K_{posre} = 1 \times 10^5, 1 \times 10^4, 1000, 100, 10$ and $0$ kJ mol$^{-1}$ nm$^{-2}$) on non-
hydrogen protein atoms have been made to relax solvent and protein. Subsequent heating with the Berendsen thermostats from 0 K to 310 K and NPT equilibration with adjustment of solvent density to 1 g/cm$^3$ followed the minimizations. The last equilibration step and the production simulations were conducted at 310 K temperature and at a pressure of 1 atm with isotropic pressure coupling (NPT ensemble). The bond lengths were restrained with the LINCS algorithm with a fourth order of expansion. The short-range electrostatic and van der Waals interactions cut-off radii were equal to 14 Å each. Long-range electrostatic interactions were treated with the particle-mesh Ewald (PME) summation with grid spacing of 0.135 nm for the fast Fourier transform and cubic interpolation. The simulations were performed for 20 ns for each system; 1 fs time steps were employed, and snapshots saved every 20 fs in order to analyze the essential collective dynamics. For each of the HRdup and M128V constructs, the MD simulations were duplicated from the same starting coordinates, using different starting velocities of atoms. The corresponding MD trajectories are denoted as “I” and “II” in the discussion.

Analysis tools
To analyze the PrP constructs from MD trajectories, their secondary structure content, numbers of hydrogen bonds and salt bridges, contact maps, and solvent accessible areas (SASA) have been calculated using scripts implemented in Gromacs [56, 57] and VMD [58] packages. Final SASA analysis was made according to solvent exposure level defined in [59]. Snapshots from trajectories and graphical representation of models was done with VMD or Accelrys VS [59]. Protein docking and predictions of residues involved in oligomer contacts for representative snapshots from M128V and HRdup trajectories were made through the InterEVDock server integrated in the RPBS Mobyle portal [60].

Essential collective dynamics
To analyze and compare dynamics of PrP alleles in greater depth we employed the novel essential collective dynamics (ECD) method [61-66]. The method stems from the statistical-mechanical analysis of the generalized Langevin dynamics of proteins [61, 64], according to which persistent correlations between atoms’ motion in the
protein can be determined from principal eigenvectors of the covariance matrix of a protein’s MD trajectory. A suite of dynamics descriptors has been derived within this framework, including in particular the main-chain flexibility profiles and pair correlation maps [62, 65]. Previously the method has undergone an extensive validation against NMR-derived [61, 64] and X-ray based structural data [62, 65, 66], and was demonstrated to predict accurately the main-chain flexibility, pair correlations, and other dynamics trends from short fragments of MD trajectories. In this work, the ECD main chain flexibilities and pair correlation maps are obtained for the HRdup and M128V constructs using the techniques described in detail elsewhere [63, 65, 66].

**Protein docking**

Representative conformations from the four production MD trajectories were used as templates for protein docking on the InterEVDock server integrated in the RPBS Mobyle portal [67]. 10,000 decoys were created, scored and clustered resulting in 10 models for each of three scoring methods (InterEVScore, SOAP_PP atom-based statistical potential, and FRODOCK). Conservation of residues was assessed with the rate4site.

**RESULTS**

**Spontaneous and transmissible prion disease in transgenic HRdup PrP mice**

In analyses of RK13 cells lacking endogenous PrP<sup>C</sup> we concluded that glycosylation, endoproteolysis, biotinylation of cell-surface PrP and immunocytochemistry with and without cell permeabilization was similar between WT and HRdup and M128V allelic variants of mouse PrP (Fig. A-1.). We generated three lines of Tg mice expressing mouse PrP with the 8-residue insertion in the HR between amino acids V128/L129 (mouse numbering scheme, Fig. 2-1A, Table 2-1), Tg.Prnp.HRdup.M128V-32, Tg.Prnp.HRdup.M128V-26 and Tg.Prnp.HRdup.M128V-10 (for brevity, Tg.HRdup-32, Tg.HRdup-26 and Tg.HRdup-10). We also created two control lines expressing a mouse Prnp allele modified to incorporate the equivalent of the human valine 129 polymorphism, Tg.Prnp.M128V-39 and Tg.Prnp.M128V-25 (for brevity, Tg.M128V-39 and Tg.M128V-25). Since GSS alleles are dominant in their natural setting, transgenes were expressed on an FVB/N Prnp<sup>+/+</sup> genetic background. The Tg lines
differed in their net steady-state levels of PrP^C, with expression levels ranging from 3.4-1.6 \times \text{endogenous for the mutant lines and 1.9-2.0} \times \text{for the control M128V lines (Table 2-1). With aging, Tg.HRdup mice, but not Tg.M128V mice, developed a spontaneous neurologic syndrome with accompanying neuropathological changes wherein animals with the highest transgene expression levels succumbed to the syndrome faster than their lower expressing counterparts (Tg.HRdup-32 > Tg.HRdup-26 > Tg.HRdup-10; \textbf{Table 2-1, Fig. 2-1, Fig. A-2 and Fig. A-3}); while the life expectancy of Tg.HRdup-10 mice was not significantly shorter than that of non-Tg FVB/N littermates in our colony, neurological presentation and a discrete pathological signature (below) distinguished it from age-related death.

The most notable clinical feature of Tg.HRdup mice was slowly progressing ataxia. As the disease advanced, weight loss was apparent and the animals were euthanized when showing kyphosis and hypokinesia. Tg.M128V mice with comparable PrP expression levels did not present with any neurologic abnormalities and were used throughout the study as negative controls (\textbf{Table 2-1}). Since myopathy has been described in GSS mice expressing the P102L allele (P101L in mouse PrP [68]), we sought these pathological changes in Tg.HRdup-26 animals with clinical disease; these studies failed to define necrotizing myopathy or neuropathy (\textbf{Fig. A-4}), suggesting that these types of lesions do not contribute to the clinical presentation of Tg.HRdup mice.

As attempts to transmit GSS isolates to primate and rodent models have yielded varied successes [15, 21, 69] we explored the issue of transmissibility using Tg.HRdup-10 animals as recipients for brain homogenate from sick Tg.HRdup-26 animals and observed acceleration of disease course from 662 ± 76 to 270 ±38 days post-inoculation (\textbf{Table 2-1}). Tg.HRdup-10 animals inoculated with non-transgenic, healthy brain homogenate culled at 399 days post inoculation had no signs of disease (\textbf{Table 2-1, Fig. A-5}) while \textit{tga20} mouse (which express WT PrP^C at ~6-7x endogenous levels and thus higher than the other Tg lines in our study [42]) remained healthy and displayed no pathology when inoculated with the same brain homogenates from Tg.HRdup-26 animals in the clinical phase of disease. These data demonstrate that pathogenic
processes in the brains of the Tg.HRdup mice extend to the generation of infectivity and offer a parallel to prior studies of transmission/host-range effects when a GSS P102L mutation is inserted into a mouse Prnp gene [14].

**Histopathology in the index case and induced by HRdup PrP transgenes**

Previously reported neuropathological data on the index case were restricted to a right frontal lobe biopsy. The formalin-fixed tissue obtained at autopsy showed only mild to focal moderate spongiosis, with mild gliosis in the neocortex and cerebellum molecular layer (Fig. 2-1.C; Fig. 2-1D I and IV); other areas showed no significant spongiosis. Vacuolation and gliosis of the hippocampus, cortex and cerebellum were present in all three Tg.HRdup lines but these features were most prominent in Tg.HRdup-26 animals (Fig. 2-1 E).

The index case showed most intense multicentric plaque (MCP) burden, as visualized with α-PrP antibodies and Periodic Acid-Schiff (PAS) stain, in the neocortex and slightly less in the cerebellum molecular layer and hippocampus. Moreover, in the neocortex there was a layer oriented distribution of MPCs. Layer 5 and 6 showed most immunostaining with aggregates of MCPs (Fig. 2-1 D; II, III, V and VI), followed by layer 1, and the least MCPs in layer 2 and 3. The cerebellum showed less MCPs at the arachnoidal side of the molecular layer and most at the Purkinje cell side. There was no association between plaques and vacuoles (Fig. 2-1 D; I and III). Somewhat less intense PrP deposits were found in other grey matter structures including striatum and thalamus with the most sparse and smallest deposits present in the brainstem with granular synaptic deposits in inferior olivary nucleus and dentate nucleus.

Tg.HRdup-26 animals showed a similar distribution of PrP plaques, with other affected areas being the anterior olfactory nucleus, corpus callosum, thalamus, anterior commissure and medulla. Some focal PrP deposits in mice stained with PAS (Fig. 2-1 E; II, inset). PrP deposits stained with monoclonal antibodies for C-terminal and N-terminal residues are shown in Fig. 2-1 F. In the case of Tg.HRdup-10 animals, the most intense PrP deposition was found in the cerebellum (Fig. A-2) whereas,
interestingly, analogous deposits were scarce in Tg.HRdup-32 animals which succumbed to disease ~400 days earlier. These data are summarized in Table A-1. Aged-matched Tg.M128V-39 animals were negative for all these histopathological hallmarks (Fig. A-3). In the case of disease produced by inoculation of Tg.HRdup-10 mice with Tg.HRdup-26 brain extracts, there was an accentuation of spongiform change and focal PrP deposition in the cerebellum (Fig. A-5) versus changes seen in aged un-inoculated Tg.HRdup-10 mice (Fig. A-2).

In the index case, Aβ, phosphorylated tau and proteasome-targeted proteins or inclusions were not detected. In agreement with this, attempts with AT8 antibody (phospho Ser202, phospho Thr205) to detect hyper-phosphorylated tau in Tg mice with spontaneous disease were unsuccessful.

**Proteinase-K resistant signature fragments in HRdup PrP transgenic mice**

Proteinase K (PK)-resistant fragments from the center of PrP and of molecular mass estimated at between 6-8 kDa distinguish GSS from the C-terminal fragments that accumulate in Creutzfeldt-Jakob Disease [70-73]. For simplicity, we will refer to this fragment in the context of the index case and Tg mice as being "8 kDa" (but also noting that its molecular mass is predicted as 0.735 kDa greater than a PK-resistant fragment generated from the corresponding WT sequence). We investigated this situation using frozen brain tissue obtained at autopsy from the index case (Fig. 2-2 A). The deep white matter had a paucity of neuropathological change and comprised an internal control for these analyses; notably, the abundance of the 8 kDa fragment normalized for protein loading was correlated with neuropathology visualized by light microscopy. This 8 kDa signature was abundantly present in aged Tg.HRdup-26 mice with spontaneous disease (Fig. 2-2 B, C) and present in brain extracts from Tg.HRdup-10 mice with spontaneous disease (Fig. 2-2 D), but not in Tg.HRdup-32 mice (Table 2-1). It was absent from aged Tg.M128V-39 control mice (Fig. 2-2 B, D). Antibody mapping confirmed the 8 kDa species originates from the center of the PrP molecule, with 12B2 (88-92) and Sha31 (145-155) epitopes present, but with PrP248 (octarepeat region; 55-96) and VRQ61 (β2-α2 loop, 165-175) epitopes absent (Fig. 2-2 C); thus the fragment
minimally consists of residues 88-155, in broad agreement with studies of a GSS-associated 7 kDa amyloid extracted fragments mapped to residues ~90-153 [73]. In terms of spontaneous pathogenic processes, the 8 kDa signature fragment was absent in young (healthy) Tg.HRdup-26 mice but was observed in animals with neurological symptoms of disease (Fig. 2-2 B). With respect to pathogenesis produced by inoculation, a paucity of this fragment in young, un-inoculated Tg.HRdup-10 mice was overcome in age-matched animals administered brain samples from sick Tg.HRdup-26 mice, whereas the same inoculum administered to over-expresser tga20 mice did not yield the signature fragment (Fig. 2-2 E). Regarding the ability to propagate the 8 kDa fragment, these data are in agreement with GSS inoculations performed using prion-susceptible bank voles and Tg mice [15].

**Assembly of the HRdup PrP into high molecular weight complexes**

In samples from aged human brain, the action of endogenous proteases upon misfolded PrP (but not WT PrP) can generate a natural protease-resistant domain. As phosphotungstate (PTA) precipitation [47] has been used previously to enrich for abnormal PrP in brain homogenates from P101L mice [74], we applied this procedure to process brain samples from our Tg mice. Subsequent western blot analysis of PrP revealed an 8 kDa signature in sick Tg.HRdup-26 mice, but not in control Tg.M128V-39 mice (Fig. 2-3 A). Immunoreactive species migrating between the positions of the 16 and 17 kDa molecular weight size marker were also noted. These data suggest a process wherein misfolded PrP assembles into higher molecular weight aggregates such that a core structure is not digested to completion by endogenous proteolytic processes (Fig. 2-3 A). To assess the quaternary state of PrP, we performed ultracentrifugation assays utilizing a linear gradient of 10-45% sucrose in the presence of 1% Sarkosyl; we analyzed brain homogenates from three biological replicates for each allelic type. These velocity centrifugation studies (Fig. 2-3 B-G) revealed that HRdup PrP populates higher molecular weight gradient fractions (highest fraction numbers) in comparison to M128V PrP. Comparing sick Tg animals to their healthy genotypic counterparts (Fig. 2-3 C versus D, E versus F) revealed more signals in high molecular weight fractions such as fractions 7 and 8 whereas signal was not detected.
at all in fraction 4-8 of Tg.M128V animals (Fig. 2-3 G). Sick Tg.HRdup-32 mice showed a profile similar to healthy Tg.HRdup-26 mice, which is in line with the absence of PK-resistant PrP in these animals (Fig. 2-3 B versus Fig. 2-3 C). These analyses suggest that, in the context of the HRdup mutation, the development of PK resistant PrP detected by immunoblotting and focal PrP aggregates revealed by immunostaining is associated with detergent insoluble aggregates that are present in the bottom fractions of the gradient (summarized in Fig. 2-3 H).

**Stability and β-sheet structure of HRdup PrP**

We next turned to secondary and tertiary structure. Recombinant PrPs corresponding to WT, M128V (residues 118-231) and HRdup (residues 118-231 with an extra eight amino acids) were expressed by standard procedures. 1H NMR spectroscopy revealed all proteins were highly enriched and folded into a primarily α-helical structure (Fig. A-6). To evaluate the stability of HRdup PrP, we performed a urea denaturation series (Fig. 2-4 A and B). Using previous chemical shift assignments of WT PrP, five resonances were identified in protons of the following residues: isoleucine181, Hγ; phenylalanine197, Hα; tyrosine161, Hα; tyrosine217, Hδ; and tyrosine162, Hε. Stacked 1D spectra of samples under different urea concentrations are shown in Fig. A-7. As previously performed to determine the stability of hamster, mouse, rabbit and bovine PrP [75], peak area values as a function of added urea were normalized to the largest value during denaturation process and then plotted against urea concentration. Two thermodynamic parameters, [D]1/2, the urea concentration at half point of unfolding and m, the slope of the denaturation curve which reflects the sensitivity of each resonance towards urea were extracted. Denaturing curves and [D]1/2 or m half values (S2 Table) for all five chosen resonances reveal that HRdup PrP had a similar sensitivity towards urea denaturation as WT and M128V PrP. Next, since the insertion in the HRdup allele includes extra residues as found within β-strand 1 in WT PrP, 2D NMR experiments were used to assess the short β-sheet structure in HRdup PrP. Here two characteristics indicate the presence of β-sheet structure: the downfield shift of Hα resonances [76] and presence of nuclear Overhauser effects (NOEs) within the α-H region (4 ppm to 5 ppm), due to the short distance between α-Hs from opposing
residues in β-sheets. Within the α-H region of both HRdup and M128V we observed the NOEs from the two spectra as having the same chemical shifts, which indicate identical β-sheet structures formed by the same residues in the two sequences (Fig. 2-4 C-E, Fig. A-8). These NOEs are also consistent with those reported for the same region of WT mouse PrP [28]; this spectral information allowed assignment of the NOEs in our spectra leading to the conclusion that the eight extra amino acids do not disrupt the short β-sheet structure present in all reported mammalian PrP NMR structures. However, these analyses are not able to determine if β-strand 1 of HRdup is composed of residues 127-130 (Tyr, Val, Leu and Gly) or residues 7-8 of the insertion and residues 129-130, or a mixture of the two.

**Molecular dynamics simulations of M128V and HRdup PrP**

Further analyses used molecular dynamics (MD) to probe PrP structure. Two MD trajectories for molecules with GYVLGGLG inserted between G125 and G126, the preferred model identified by homology modeling, ("HRdup-I" and "HRdup-II") and two for molecules without the insert ("M128V-I" and "M128V-II") were simulated at a temperature 310 K and pH 4.5 for 20 ns each, with secondary structure (SS) elements shown in Fig. 2-5A and Fig A-9 (see also Table A-3). Both M128V structures retained all major secondary structure elements including β-strands S1 and S2 and three α-helices H1, H2, and H3, with the exception of a small C-terminal part of H2 (Fig. 2-5 A and Fig. A-9 A and B). In system HRdup-I, the N-terminus of helix H2 and almost the entire helix H3 unfolded (Fig. 2-5 A and Fig. A-9 C). Residues insG6 and insL7 replace G123, and L124, forming a new beta strand. In system HRdup-II (Fig. A-9 D) helix H2 and the middle part of helix H3 unfolded. Transient β-content was occasionally observed in N-terminal area of M128V-I, and to a greater extent in the areas of the insert and loop S2-H2 in HRdup-I and HRdup-II (Table A-4). To quantify structural differences between the alleles we averaged the per-residue solvent-accessible areas (SASA) [77] over the last 2 ns from the four MD trajectories (Fig. 2-5 B and Table A-5 and A-6). Two trajectories for HRdup and those for M128V exhibit close total SASAs for various groups of residues, as well as for the entire protein (Table A-5), and average per-residue SASAs were also close in HRdup and in M128V.
However, the insert caused local effects (Fig. 2-5 B). Changes in SASA were evident in extended regions such as P104-K109, V111-V120, V121-G125, A132-D146, P164-Q167, N173-V179, V188-M204. (Fig. 2-5 B and Table A-5). Only a few residues with charged side-chains exhibited large differences in SASA suggesting a dominating effect on the hydrophobic residues.

Next, PrP dynamics were analyzed by an essential collective dynamics tool (ECD). ECD results supported the aforementioned trends, with main-chain flexibility profiles [63, 66, 78, 79] of the PrP constructs shown in Fig. 2-5 C. High levels of the flexibility descriptor represent loops, whereas minima indicate rigid areas such as α-helices or β-strands. In both constructs the areas of helices H2 and H3 are characterized by broad minima indicating a relative rigidity whereas helix H1 is less stable, which is not unusual for PrP [63, 78]. The insertion caused extended regions of increased rigidity N-terminal to the insert and in the area of L108, where a systematic decrease in SASA was also observed (Fig. 2-5 B). The insert also seemed to destabilize the N-terminal part of helices H2 and H3 (Fig. 2-5 C) while increasing SASA in the same area (Fig. 2-5 B). Overall, data in Fig. 2-5 B and Fig. 2-5 C indicates that an increase in per-residue SASA is often associated with greater main-chain flexibility, and vice versa.

ECD pair correlation maps (Fig. A-10) provide complementary information on dynamic correlations of atomic motion within the same statistical-mechanical framework [66, 79], with S7 Table listing non-consecutive residues that show the strongest dynamic correlations in this type of analysis; Fig. 2-5 D and E graph these data and compares them with the 2D NOESY experiments. In summary, for M128V the strongest correlations were found between β-strands S1 and S2, between β-strand S2 and helix H2, between helices H2 and H3, and between β-strand S2 and helix H3. For HRdup, the correlated elements are similar but the number of strongly correlated residues in H2 and H3 was notably decreased, whereas an additional correlation was observed between residues insG8 and Y127 with N-terminal residue A116.
Lastly, given the role of protein assembly effects in prion biology, protein/protein docking was assessed using representative trajectories (Fig. A-11 and Table A-8). Beyond intermolecular contacts found in top dimer model based on M128V-I and M128V-II, new contacts appeared or became more frequent in the N-terminus (W98, K100, A114, A115) and helix H3 (Y225, D226) of the HRdup-I and HRdup-II constructs. N-terminal contacts involving W98, K100, V111, and A116 were also frequently observed in heterodimers (Table A-8).

**A thermolysin-resistant PrP signature fragment in GSS**

The protease thermolysin was used as an enzymatic probe of conformation to interrogate PrPC\(^{-}\)'s natively disordered region [80, 81]. Thermolysin will completely degrade PrPC\(^{-}\) (see below) but produces a signature that includes protease-resistant full-length PrP\(^{Sc}\) from diseases such as mouse-adapted scrapie, hence providing information about the accessibility of residues in the natively disordered N-terminal region. For samples taken at autopsy, digestion of normalized protein samples from the index case from neuroanatomical areas with notable pathology (cerebellum (Cb); frontal cortex (FC); parietal cortex, PC)) yielded an intense 16 kDa thermolysin resistant fragment (Fig. 2-6 A), while the 16 kDa signal from an area with less pathological staining (deep white matter (WM)) was less notable. A 16 kDa species was also present in PTA precipitations from the same four tissue samples processed without any *in vitro* protease digestion (Fig. 2-6 B, Sha 31 antibody, lanes1-4). Here the relative abundance Cb ~FC ~PC > WM was again noted. While an 8 kDa species was not apparent in these analyses of the PTA precipitates of human material (Fig. 2-6 B, lanes 1-4), the same samples did yield an 8 kDa PrP fragment after PK digestion (Fig. 2-6 B, lanes 5-8) with a similar profile of signal intensity as noted above with thermolysin digestion, namely three robust signals (Cb ~FC, ~PC) versus a smaller signal (WM). A 16 kDa species was also seen with an octarepeat antibody in undigested PTA-precipitates (Fig. 2-6 C). Next, to address disease specificity, we sought similar signatures in sporadic CJD material or in normal brain. Two types of sCJD case containing a V129 polymorphism did not yield strong 16 kDa species but instead multiple species closer to the mobility of undigested PrP (Fig. 2-6 D). A normal control did not produce any TL-resistant
species (Fig. 2-6 D) while cerebellar material from the proband yielded a doublet running slightly slower than the 16 kDa marker. To assess generality and the possibility that the 16 kDa signature can co-exist with or prefigure the generally accepted appearance of 7-8 kDaPK resistant species present in different types of GSS [72, 82, 83], we analyzed other GSS cases; these were pre-selected from samples within the US CJD surveillance system as harboring 7-8 kDa PK-resistant fragments. As shown in Fig. 2-6 E, a 16 kDa signature was shared by brain material obtained from four other GSS alleles (Fig. 2-6).

**Origins and properties of the thermolysin-resistant PrP fragment**

As anticipated, a thermolysin-resistant 16 kDa fragment was also present in the brains of TgHRdup-10 and TgHRdup-32 mice in clinical phase of disease, but not in the control Tg.M128V-39 line (Fig. 2-7 A); electrophoretic mobility of the 16 kDa fragment was unaffected by the use of PNGase F, suggesting that the C-terminus must lie N-terminal to the glycosylation sites (Asn180 and Asn196); this assumption was validated by four antibodies with epitopes N-terminal to these two positions (Fig. 2-7 B). Our data exclude that the 16 kDa fragment corresponds to an un-glycosylated, thermolysin-resistant form of the C1 fragment produced by physiological endoproteolysis of PrP<sup>C</sup> - this is because two antibody epitopes (248 and 12B2) lie N-terminal to mouse PrP C1 N-termini at residues 109, 100 yet can detect the 16 kDa fragment and also because VRQ61 antibody that detects a C-terminal epitope present within C1 PrP nonetheless fails to detect the 16 kDa species. As the 16 kDa species appears as a doublet with the PrP248 and 12B2 antibodies but less clearly so with Sha31, there may be raggedness at the C-terminus in the vicinity of the Sha 31 epitope (Fig. 2-7 B). Heterogeneous cleavage of PrP has also been noted in vivo during analysis of A117V, F198S and Q217R cases and following PK digestion of GSS cases harbouring the P102L mutation. Levels of this fragment increased with age (Fig. 2-7 C). Sick Tg.HRdup mice could contain some immunoreactivity after thermolysin treatment indicative of full-length glycosylated PrP, but this was not the major species. Most notably, the 16 kDa fragment was detected at ages as young as 7 days in the brains of Tg.HRdup-26 brain mice (Fig. A-12), underscoring occurrence preceding the 8 kDa
species and hence a potential role as a precursor. With regards to accumulation, level of this fragment changed by 5-fold and net thermolysin-resistant PrP signal increased by 13.7-fold between pre-symptomatic (200d) and end-stage (442d) line 16 animals (Fig. 2-7 C). We next used an in vitro double-digest experiment (Fig. 2-7 D) to explore the ability of 16 kDa thermolysin resistant PrP species to engender an 8kDa PK resistant fragment. TL digests of 50 µg of protein from two aged Tg.HRdup-26 animals revealed a 16kDa fragment as the major species, along with two sub-molar species, confirming and extending the inference of ragged fragment termini. As anticipated, an additional PK treatment after a TL digest yielded an 8kDa species as the predominant product; in both cases this was accompanied by two sub-molar species, albeit with slower (not faster) electrophoretic mobility than the predominant species (Fig. 2-7 D).

We extended these analyses to three other aged Tg.HRdup-26 animals and quantitated the % autoradiographic in the vicinity of the 16 or 8kDa species versus the complete integrated signal in a Mr range from 34-4kDa. 16 kDa or 8 kDa species comprised > 80% signal in the TL or TL plus PK digests, respectively, and these percentages were not significantly different (Fig. 2-7 E). Thus, in an in vitro situation, the major product of a TL digest is efficiently converted after an additional PK digestion to a major product that has similar mobility to the major product of PK digest; in short, the novel 16 kDa TL-resistant species can engender 8 kDa PK-resistant hallmark species.

Concerning templated seeding and 16 kDa species, the inoculum from Tg.HRdup-26 mice that produced early clinical disease in Tg.HRdup-10 increased the level of 16 kDa fragment above genotypic controls treated with inoculum from healthy WT mice. No such fragment was detected in tga20 mice (i.e. expressing WT mouse PrP) seeded with Tg.HRdup-26 inoculum (Fig. 2-7 E), suggesting an allelic barrier to seeding.

DISCUSSION
Genetic versus transmissible aspects of GSS disease
Prion diseases can be sporadic, infectious or genetic; GSS, in particular, while caused by germline mutations, can also be infectious in certain lab settings. Mice expressing either WT murine PrP with a stop codon before the GPI anchor signal sequence or a
natural I109 allele of WT bank vole PrP are reported to develop a spontaneous GSS-like syndrome [20, 84] but, for the most part, the goal of attaining GSS-like neuropathology and plaque deposition has been met by Tg mice that introduce mutations within the framework of mouse Prnp [85-87]. Conversely, P101L knock-in mice with 1x endogenous expression levels remain asymptomatic [88, 89], as do mice expressing human PrP with a P102L or A117V mutation [13, 90] and neither line are described as having spontaneous accumulation of pathognomonic 7-8 kDa PK-resistant fragment. This work presents a third transgenic model of GSS with spontaneous disease appearance where the human GSS mutation is used in the context of WT mouse PrP (the Prnp^a allele) but without synthetic epitope tags [91], thus joining Tg.P101L and Tg.A116V models [85, 86]. While PK-resistant PrP has not been observed in the brains of Tg.P101L mice or P101L knock-in mice [14, 68, 85, 88, 92], here we were able to detect robust levels of a pathognomonic 8 kDa PK-resistant fragment in two Tg lines with spontaneous disease. A similar low molecular weight PK-resistant PrP fragment present in A117V patients has been detected in the brains of sick Tg.A116V mice, though seemingly less prominent than observed in the HRdup expressing animals analyzed here [86] (Fig. 2-2 B-E). An 8 kDa fragment is seen in transmissions from GSS tissue into bank voles [15], and in spontaneously sick Tg mice expressing bank vole PrP or GPI-anchorless mouse PrP [20, 84]. Association between the 8 kDa fragment and transmissibility is also present in our studies based on the HRdup GSS allele, where disease presentation in the lowest-expressor line, Tg.HRdup-10, was accelerated and levels of the 8 kDa PK-resistant fragment were enhanced over age-matched (asymptomatic) mice of the same transgenotype (Fig. 2-2 F, Table 2-1). This finding on species-barrier effects aligns with studies using other GSS models and a β-enriched recombinant peptide approximating to the 8 kDa fragment [14, 68, 92, 93], but noting that transmissions into mice expressing P102L human PRNP can reveal different host-range properties than for mouse Prnp [94].

**Misfolding of HRdup and a 16 kDa thermolysin-resistant signature**

What might be the *cis* effects of the 8 amino acid insertion upon folding of HRdup PrP? We failed to see significant distinctions between alleles in urea denaturation studies by
monitoring five residues in the C-terminal globular domain of WT, M128V and HRdup PrP (Fig. 2-4 A, B; Fig. A-7), suggesting the HRdup insertion does not affect the allele's PrP\textsuperscript{C}-like global fold. Sequence inspection of HRdup reveals that the first β-sheet is split by the insertion, creating the potential for a second β strand separated by a short linker region (Fig. 2-4 C). As conversion from PrP\textsuperscript{C} to PrP\textsuperscript{Sc} involves an increase in β-content, one might speculate that HRdup has intrinsically more β-structure than a WT counterpart that could drive its conversion to a PK resistant form [5, 95]; while no additional β-sheets were noted in the ensemble measurements of recombinant protein preparations by NMR (Fig. 2-4 D), transient occupation of additional β-structure versus WT controls was noted in MD modeling following the trajectory of individual molecules (Fig A-9, Table A-4). In the HR of WT PrP there are four conserved glycines within three GxxxG motifs, while HRdup contains 6 equally spaced glycines within 5 of these motifs. GxxxG motifs were first identified as mediators of transmembrane helix-helix association [96]; there are indications from prion infections that these extracellular sequences may feature in uptake of prions in the replicative cycle [97]. \textit{In vitro} work has shown that interruption of these GxxxG repeats decreases the PK resistance of recombinant PrP folded into a β-rich form. This finding has some parallels in allelic alterations in tandem repeats in Shadoo, a PrP family member with a HR lacking an N-terminal palindrome but instead composed solely of GxxxG repeats [98-100]. Further appraisal of the properties of GxxxG sequence motifs in PrP\textsuperscript{C}'s HR may thus be fruitful. Beyond this, as the HR is only one part of PrP\textsuperscript{C}'s natively disordered region and since perturbations were noted in sequences N-terminal to the insertion in MD simulations of monomers and in a dimer interface in \textit{in silico} docking studies (Fig. A-11), the question arises as to whether a larger region is impacted by the HRdup mutation. As a chemical probe, we employed thermolysin, an enzyme used previously to assess A116V PrP expressed in cell culture [101] and where the authors reported a greater resistance of this GSS associated allele to thermolysin relative to PK. However, to the best of our knowledge, the discrete N-terminal 16 kDa thermolysin-resistant fragment present at robust levels (and using the enzyme at 70 °C [102]) is a previously unreported feature, being in line with the distinct
molecular nature of GSS in that it differs from digestions of CJD and scrapie brain material.

Multimerization and acquisition of protease-resistance often go hand-in-hand for prion diseases and here we note the HRdup PrP holoprotein has different assembly properties for than M128V PrP (Fig. 2-3) and that a self-aggregation determinant [103] lies within the boundaries of the 16 kDa fragment defined by epitope mapping. Moreover, unlike the natural C1 endoproteolytic fragment of PrP, the boundaries of 16 kDa thermolysin-resistant fragment do not preclude a precursor relationship to the 8 kDa form (Fig. 2-8 A). Indeed, there is an interrelationship in the levels of 16 and 8 kDa species in i) different brain areas from the index case (Fig. 2-6 D), ii) in time-course analyses of normalized brain protein extracts of aging mice where the 16 kDa fragment occurs earlier than the 8 kDa fragment (Fig. 2-2 E, 2-7 C) and iii) in sequential digest experiments (Fig. 2-7 D). Our data support a series of events wherein a fraction of full-length HRdup undergoes misfolding in the N-terminal domain such that it starts to assemble to multimers, becomes precipitable with PTA and yields a 16 kDa signature upon thermolysin digestion (Fig. 2-2, 2-3, 2-6, 2-7 and 2-8). Enhanced accumulation of the 16 kDa fragment in inoculated Tg.HRdup-10 mice (Fig. 2-7 F) is also compatible with some templating and self-assembly capacity that may warrant further investigation. Over the course of time this conformationally altered form of PrP may yield 8 kDa PrP through the action of endogenous degradative processes (Fig. 2-3 A and 2-8) [73]. Conversely, the use of distinct proteases and antibody mapping exclude an alternative explanation that the 16 kDa PrP species detected by immunoblot reflect SDS-resistant dimers of 8 kDa PrP species.

16 kDa thermolysin resistant PrP and disease pathogenesis

The presence of the 16 kDa signature in four other GSS alleles in cis to Val129 and increased fragment levels in more pathologically affected areas of the brain from the index case (Fig. 2-6 A, B and F) both argue for intimate involvement in the disease process. In addition, the C-terminal boundary of the thermolysin-resistant fragment adjacent to the end of helix 1 offers an unexpected parallel to pathogenic stop codon
mutations such as PrP145X, where structural biology and transmission properties have been studied in depth [104-106]. Finally, Tg.HRdup-32 mice succumb to spontaneous disease earliest of all the Tg lines described here and accumulate 16 kDa thermolysin resistant PrP but not the 8 kDa PK resistant fragment (and, hence, cannot accelerate disease in recipient animals), so a straightforward inference from these net data is that the misfolded form of PrP revealed by thermolysin digestion has neurotoxic activity. Hypothetical toxic forms of PrP (PrP\textsuperscript{L}) have been inferred for prion infections, with clinical disease emerging when levels of PrP\textsuperscript{L} transcend a threshold; in this scheme, autocatalytic propagation of PrP\textsuperscript{Sc} ("Phase 1") precedes toxicity ("Phase 2") [107]. We suggest that a permutation of these concepts may come close to representing the GSS pathogenic process; specifically, the potentially toxic species could start to accumulate early on and be diagnosed in tissue by thermolysin treatment as the 16 kDa fragment. The later disease process may be characterized by two events. First, noting that the level of the 16 kDa thermolysin-resistant fragment rises with chronological age in Tg mice, levels of toxic misfolded forms of PrP may transcend a threshold needed for clinical manifestation. Second, the 8 kDa PK-resistant species that is adept at autocatalytic propagation events - "infectivity", as measured in experimental transmissions from human source material in bank voles [15] or from Tg mouse material (Fig. 2-2 F) - begins to rise. Delayed occurrence of the 8 kDa species may reflect its derivation from the form of misfolded PrP represented by thermolysin resistance. This scheme of pathogenesis is summarized in Fig. 2-8 B. While full-length thermolysin-resistant forms of PrP were present at end-stage in Tg mice (Fig. 2-7 A, C) and might also have toxic potential, levels of full-length thermolysin-resistant forms of PrP were unremarkable in human GSS tissue (Fig. 2-6 A, E) and hence were not considered an obligatory item in pathogenesis. Future studies to inventory the levels of abnormally folded full-length PrP by use of PTA (Fig. 2-6 B, C) or by detergent insolubility, may be useful probe and refine this new view of pathogenesis.

In terms of accumulation of 16 kDa thermolysin-resistant PrP and toxicity thresholds, for the Tg.HRdup-26 line, the 16 kDa signature is apparent at almost the very time PrP transcript levels increase in rodents above the levels present in embryos, this transition
correlating with neuronal precursor cells ceasing proliferation and beginning to differentiate [108-111]. Clinical disease is scored 400 days later in Tg.HRdup-26 mice and ~170 days later for Tg.HRdup-32 mice. These data beg the question of whether an early start to the accumulation of abnormal PrP species resulted in a corresponding early transit across a threshold for manifestation of clinical symptoms. Of interest, the index case had psychiatric events in his twenties and is reported to have a son affected with an autism spectrum disorder condition (Asperger's) [23]. Both WT PrP\textsuperscript{C} and HRdup PrP undergo physical interactions with DPP6 [112, 113], a type II membrane protein [114] that controls dendritic morphogenesis and one wherein gene disruptions link to different neurodevelopmental disorders [115-117]. Thus, a number of fascinating possibilities may emerge from further investigation of this kindred. More broadly, the Tg.HRdup mice comprise a new tool for structural biological investigations to investigate changes in PrP's N-terminal region [6] and links to neurotoxicity. Given that other GSS mutants may share a chemical signature, there may be an opportunity to understand common disease mechanisms in GSS and, pending the generation of new molecular probes for this form of PrP, perhaps more common CNS disorders as well.
**Fig. 2-1.** HRdup PrP causes GSS in transgenic mice. A) Line diagram of PrP demonstrating the site of the 8 amino acid insertion (yellow box) in the center of PrP. The hydrophobic region, "HR", is shown in orange, β-strands (strand 1 and strand 2, "S1" and "S2") are in red and helix A in dark blue. The presence of valine at res. 128 is indicated. B) Kaplan-Meier survival plot of the transgenic lines expressing HRdup PrP. Tg.HRdup-32, 166 ± 39 days (black, SD; n=9); Tg.HRdup-26, 410 ± 40 days (red, SD; n=23) and Tg.HRdup-10, 662 ± 76 days (blue, SD; n=16). Panels C & D. Pathological features of the index GSS patient. C) Focal neocortical moderate spongiosis (middle frontal gyrus), not preferentially associated with MCP. H&E staining, scale bar = 50 μm. D) Low power views of the cortex in H&E, PAS and 12F10 antibody stain (I-III) and a higher power view (IV-VI) with size bars of 500 μm and 25 μm, respectively. E) Pathological features of Tg animals with spontaneous disease. Photomicrographs of sagittal brain sections of Tg.HRdup-26 mice at terminal stage of disease (panels I-XII). Immunostaining for PrP was performed after treatment of the slices with formic acid while slices for examination with H&E (third column) or GFAP (fourth column) were left untreated. The CA1 region of the hippocampus is shown. Inset in panel VI shows a PAS-positive plaque in the cortex. Scale bar = 25 μm. F) High power view of cortical (I and II) and cerebellar (III and IV) PrP deposits in Tg.HRdup mice stained with two antibodies, as indicated. Scale bar = 50 μm.
Fig. 2-2. PK-resistant PrP in the index case, spontaneously sick and inoculated Tg mice. A) Presence of proteinase K (PK)-resistant PrP in different brain regions of the index case. Cb, cerebellum; FC, frontal cortex; PC, parietal cortex; WM, white matter. Black arrow indicates the position of the 8 kDa PrP fragment. B) Age-dependent accumulation of the 8 kDa fragment in Tg.HRdup-26 mice but not in Tg.M128V-39 control animals. C) Immunoreactivity of the 8 kDa fragment versus four α-PrP antibodies. PrP248 recognizes the octarepeat region (residues 55-96) of PrP and fails to detect the 8 kDa fragment. 12B2 (residues 88-92) is reactive with the 8 kDa fragment. VRQ61 recognizes the β2-α2 loop (residues 165-175) of PrP is not immunoreactive with the 8 kDa fragment. Using these epitopes as minimal termini, the fragment maps to residues 88-155 using the numbering scheme of WT mouse PrP, which corresponds in turn to a molecular weight of ~7.2 kDa. D) PNGase F treatment. The 8 kDa fragment is not affected by treatment with PNGase F. Presence of protease-resistant PrP in Tg.HRdup-10 mice but not in aged Tg.M128V-39 mice. "H", healthy"; "S", sick. At the terminal stage of disease, HRdup prions are resistant to digestion by 10 μg/ml PK. E) Presence of the 8 kDa fragment in low-expresser Tg.HRdup-10 mice is hastened by inoculum with brain homogenate from a sick Tg.HRdup-26 animal. Low expressing Tg.HRdup-10 animals and tga20 (overexpressing WT PrP ~6x) animals received an intracranial inoculation with brain homogenate from a sick Tg.HRdup-26 animal. Tg.HRdup-10 animals inoculated with healthy brain homogenate were euthanized after 400 days. Tg.HRdup-10 animals inoculated with brain homogenate from a sick Tg.HRdup-26 animal succumbed to disease at 276±38 days. Brain homogenates from inoculated animals either undigested or exposed to PK is presented. Note that only Tg.HRdup-10 animals inoculated with Tg.HRdup-26 brain homogenate display the 8 kDa PK resistant fragment (black arrow) versus genotype-matched animals treated with control inoculum (right four lanes). Analyses in A, B, D and E were performed with Sha31 antibody.
Fig. 2-3. *In vivo* and *in vitro* properties of HRdup prion. A) The 8 kDa PK-resistant fragment is formed *in vivo*. Following PTA precipitation, the 8 kDa fragment can be observed by western blot without the use of PK (indicated by black arrow), while it is absent from control mice (Tg.M128V-39). B-G) Sedimentation properties. PrP in both sick and healthy Tg.HRdup animals is found in lower fractions of linear 10-45% sucrose gradients in 1% Sarkosyl than in aged Tg.M128V animals. PrP immunoreactivity emerges in the lowest fraction at end-stage of disease, when PrP becomes PK-resistant while, in asymptomatic mice, PrP remains in the upper fractions of these gradients. B) A sick Tg.HRdup-32 animal. C) A healthy Tg.HRdup-26 animal. D) A sick Tg.HRdup-26 animal. E) A healthy Tg.HRdup-10 animal. F) A sick Tg.HRdup-10 animal. G) A healthy aged Tg.M128V-39 animal. The Sha31 antibody was used for all blots. H) Quantification of PrP in fractions 4-8 from B, D, F and G. n=3; statistics done using a one-sided students t-test; n.s., not significant; *, p<0.05; **, p<.005; ***, p<.0005.
**Fig. 2-4.** NMR spectra and beta-sheet signatures of recombinant proteins. A) Representative urea titration of one (Y162) of five assayed residues positioned in PrP's globular domain, as assessed for three PrP constructs (as indicated). B) Half maximal urea concentrations of each of the five residues with allelic origin color code as per panel A and presented with residues in numerical order. No pair-wise comparisons between the same residue measured in the three alleles reached significance. C) Position of beta strand S1 in WT PrP (top line) and a potential, additional beta strand, both indicated by S, that might arise as a consequence of the tandem duplication encompassing the residues YV1G. Lower panels; Part of the 2D $^1$H-$^1$H NOESY spectra of M128V (D) and HRdup (E) where NOEs due to the presence of β-sheet are displayed.
Fig. 2-5. Molecular dynamics assessment of PrP HRdup.

A) Representative conformations from 20ns molecular dynamics runs for M128V and HRdup PrP. α-helices are colored with magenta, β-strands with yellow, turns with cyan, random coils are white and the insert (for HRdup) is red. For the alignment, regions in HRdup that show differences with respect to M128V are depicted in green, and the insert is indicated in red. B) Differences in weighted average per-residue SASA for HRdup versus M128V models were averaged over two trajectories per allele (positive or negative differences indicate, respectively, greater or lesser solvent exposure in HRdup in comparison with M128V, as listed in the last column of Table S5). Positions of helices and beta strands are shaded in grey (see also panel C). C) Main chain flexibility profiles in the PrP systems. For M128V, the flexibility profile is shown with red and black lines, and for systems with insert the profiles are shown with yellow and blue lines. The insert area is marked with vertical dashed lines. The main secondary structure elements in PrP are shaded in grey are indicated at the top of the plots with purple lines for α-helices and yellow lines for β-strands. Panels D) and E): The strongest correlations of non-consecutive residues from ECD pair correlation maps of the four PrP systems, as listed in Table S5. D) Data for M128V-I (red dots) and M128V-II (black dots); E) Data for HRdup-I (blue crosses) and HRdup-II (yellow crosses). The regions of strongest correlations S1-S2, S2-H2, S2-H3, and H2-H3 are indicated. Correlations S1-S2 determined from 2D NOESY experiments are shown with magenta crosses in D and E.
**Fig. 2-6.** A 16 kDa thermolysin-resistant PrP signature in human brain material. A) Exposure of homogenates from different brain regions of the HRdup GSS patient to thermolysin. Cb, cerebellum; FC, frontal cortex; PC, parietal cortex; WM, white matter. Open arrow indicates the position of the 16 kDa PrP fragment. B) PTA precipitates of lysates identifies a prominent 16 kDa fragment (open arrow) and an 8 kDa species after PK digestion (black arrow), Sha31 antibody. C) PTA precipitates analyzed with PrP248 antibody (octarepeat epitope) reveals a 16 kDa species in Cb, FC, and PC. D) Thermolysin digests of MV1 and VV2 forms of sporadic CJD samples alongside a normal brain and an HRdup control. Relative protein loading in micrograms before digestion are indicated under the gel lanes. E) Thermolysin digests of brain material from other forms of GSS; with the exception of P102L, the presented mutations are in cis to a valine codon at residue 129. Gel loadings are described as per D. Heterogeneity in the 16 kDa species is indicated by arrows with solid or dashed perimeters; see also Fig. 7B. Analyses in A, B, D and E were performed with Sha31 antibody.
A 16 kDa thermolysin-resistant PrP signature in mouse brain material. 

A) 16 kDa thermolysin-resistant PrP signature (open arrow) found in Tg.HRdup-10 and -32 mice is not altered in mobility by PNGaseF treatment. Animal ages are line 16, 445d; line 10, 630d and line 32, 176d. A healthy TgM129V-39 control (684d) is also shown. B) The immunoreactivity of the 16 kDa fragment to various α-PrP antibodies was determined. PrP248 (residues 55-95), 12B2 (residues 88-92) and Sha31 (residues 145-152) reacted whereas VRQ 61 (res 165-175) did not. A second lower Mr species, suggesting TL-digestion generates ragged termini, by is indicated by an open arrow with a dotted perimeter. These antibody data suggest the larger TL-resistant fragment maps to residues 23-155 (mouse numbering scheme for WT PrP), which corresponds in turn to a molecular weight of ~14.7 kDa. C) Time-course of appearance of thermolysin-resistant HRdup PrP (indicated by open arrow); this signal increases as the animals age. Some residual full-length PrP of ~34 kDa can be noted at the fourth time-point. D) Double-digests to define a precursor-product relationship between 16 kDa TL-resistant and 8 kDa PK-resistant PrP species. Single and double digests of a 398d and 340d Tg.HRdup-26 brains. In each blot, the first lane represents 50 µg protein digested with thermolysin, second lane represents 100 µg protein digested with thermolysin and then digested with PK, with the major product migrating at 8kDa (open and black arrows, respectively). The third lane represents the sample digested only with PK but overloaded (100 µg load) to emphasize the 8kDa product. Heterogeneity, indicative of ragged termini, is apparent for these gel loadings, as per panel B. E) Quantitation of signal in TL and TL plus PK digests. Data represent digests of brain material from five Tg.HRdup-26 mice (range 340-398 d.) Left hand column represents %
autoradiographic signal in TL digest in the vicinity of 16 kDa versus net integrated signal in gel lane between 34 and 4kDa. Right hand column represents % autoradiographic signal in TL plus PK digest from 12 to 4 kDa species versus net integrated signal in gel lane between 34 and 4kDa. Percentage values between single and double digests are not significantly different (n.s.). F) Low expressing Tg.HRdup-10 and overexpressing tga20 (~6x) animals received an intracranial inoculation with brain homogenate from a sick Tg.HRdup-26 animal. Tg.HRdup-10 animals inoculated with healthy brain homogenate were euthanized after 400 days. Tg.HRdup-10 animals inoculated with clinical Tg.HRdup-26 brain homogenate succumbed to disease at 276±38 (SD) days. Thermolysin exposure of brain homogenates from inoculated animals is presented. Only Tg.HRdup-10 animals inoculated with Tg.HRdup-26 brain homogenate display the ~16 kDa thermolysin-resistant fragment (open arrow).
**Fig. 2-8.** Ontogeny and properties of abnormal PrP species in Tg.HRDup mice. A) In vitro proteolysis products of HRDup and WT PrP derived in young animals. The hydrophobic region (HR) is presented in orange, helices in dark blue, beta strands in red and all other residues indicated by mid-blue shading. The approximate location of the epitopes of the anti-PrP antibodies used to map the boundaries of the protease resistant domains are indicated. In the presence of thermolysin (TL) and proteinase K (PK) WT PrP¢ is completely degraded (indicated by dotted outline). In contrast the disposition of HRdup PrP is different, with conformational effects of the 8-residue insert (yellow) insert spreading both in an N-terminal direction (horizontal black arrow) and in a C-terminal direction (horizontal grey arrow); these changes in conformation are indicated by the mid-blue shading of PrP¢ being replaced by wavy lines. While being completely degraded by PK (like PrP¢) HRdup can adopt a conformation whereby the N-terminal region is TL-resistant and hence only the C-terminal portion of the molecule is degraded by this protease. From gel analyses the cleavage sites may be heterogeneous ("ragged termini"); these termini have not been mapped in detail but are shown in a provisional manner by small vertical arrows). The action of C1 protease (large vertical arrow) immediately N-terminal to the HR would preclude the formation of 16 or 8 kDa protease-resistant species. B) Proposed evolution of protease resistant PrP species from PrPHRdup. For simplicity, ragged termini have been omitted from this schematic. The aging process is shown on the vertical axis. The signature 16 kDa thermolysin fragment of the HRdup PrP is already present in young mice, exhibits a slow accumulation with aging but remains PK sensitive. A similar species may accumulate spontaneously from endogenous protease action without the need for in vitro thermolysin digestion (asterisk) At a later stage in the disease course the 16kDa species are hypothesized to undergo a different conformational change (indicated by dark blue wavy fill) such that its C-terminal region (i.e. the central region of PrP) acquires resistance to endogenous proteases and accumulates; it also has the property of being resistant to PK digestion performed in vitro and yields the GSS signature 8 kDa PK-resistant fragment. The 8 kDa species can be amplified by templated refolding as shown in transmission experiments. Some full-length protease-resistant PrP may also be present at disease end-stage (e.g., Fig 7C, lane 6) but this is neither the predominant species nor consistently present and hence is not represented here. The presence of 16 kDa TL-resistant species in a number of GSS cases containing 7-8 kDa PK-resistant species (Fig. 6E) suggest that this scheme may be generally applicable.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Tg.HRDup-32</th>
<th>Tg.HRDup-26</th>
<th>Tg.HRDup-10</th>
<th>Tg.128V-39</th>
<th>Tg.128V-25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net expression level relative to WT (1x)</td>
<td>3.4x</td>
<td>2.6x</td>
<td>1.6x</td>
<td>2.0x</td>
<td>1.9x</td>
</tr>
<tr>
<td>Age euthanized (days)</td>
<td>166 ± 39</td>
<td>410 ± 40</td>
<td>662 ± 76</td>
<td>spontaneous neurologic disease not detected</td>
<td>spontaneous neurologic disease not detected</td>
</tr>
<tr>
<td>8 kDa PK-resistant fragment in diseased micea</td>
<td>not detectable</td>
<td>yes</td>
<td>intermittent</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>16 kDa thermolysin- resistant fragment in adolescent micea</td>
<td>ND</td>
<td>yes (7 days)</td>
<td>yes (15 days)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>16 kDa thermolysin- resistant fragment in aged micea</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>not detectable</td>
<td>ND</td>
</tr>
<tr>
<td>Age euthanized post inoculationb</td>
<td>N/A</td>
<td>N/A</td>
<td>270 ± 38</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Table 2-1.** Expression level, disease onset and signature PrP fragments in transgenic mice.

ND, not done; N/A, not applicable.

a analysis of brain homogenate without PTA precipitation.

b inoculated with brain material from clinical phase Tg.HRDup-26 mice. Tga20 mice inoculated with the same material were held for an observation period of 400 days with no signs of clinical disease.
REFERENCES


Chapter 2 – Canonical structure and structural stability of HRdup

297/5580/353 [pii].


Chapter 2 – Canonical structure and structural stability of HRdup


http://www.pnas.org/cgi/content/full/97/1/145
http://www.pnas.org/cgi/content/abstract/97/1/145.


[54] C. R. Sondergaard, M. H. Olsson, M. Rostkowski, and J. H. Jensen, "Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and


Chapter 2 – Canonical structure and structural stability of HRdup


[88] J. C. Manson et al., "A single amino acid alteration (101L) introduced into murine PrP dramatically alters incubation time of transmissible spongiform


[95] J. R. Requena and H. Wille, "The structure of the infectious prion protein: experimental data and molecular models," *Prion*, vol. 8, no. 1, pp. 60-6, Jan-
Chapter 2 – Canonical structure and structural stability of HRdup


CHAPTER 3 – Nascent β structure at the N-terminus of HRdup

Nascent β structure in the Elongated Hydrophobic Region of a Gerstmann-Sträussler-Scheinker PrP Allele

Ze-Lin Fu\textsuperscript{1,2}, Peter C. Holmes\textsuperscript{2}, David Westaway\textsuperscript{1,2}, Brian D. Sykes\textsuperscript{2}\textsuperscript{*}

\textsuperscript{1}Centre for Prions and Protein Folding Diseases, University of Alberta, Edmonton, Alberta, Canada, T6G 2R3;
\textsuperscript{2}Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada, T6G 2H7

Publication: Fu et al. 2019. Journal of Molecular Biology, 431, 2599-2611

Contributions: ZLF, DW and BDS designed the experiments. ZLF prepared all the recombinant proteins. ZLF and BDS performed NMR experiments. ZLF and PCH analyzed the NMR data. ZLF wrote the manuscript with BDS, DW and PCH.

OVERVIEW

Prion diseases are neurodegenerative disorders caused by the misfolding of the cellular prion protein (PrPC). Gerstmann–Sträussler–Scheinker syndrome is an inherited prion disease with one early-onset allele (HRdup) containing an eight-amino-acid insertion; this LGGLGGYV insert is positioned after valine 129 (human PrP\textsuperscript{C} sequence) in a hydrophobic tract in the natively disordered region. Here we have characterized the structure and explored the molecular motions and dynamics of HRdup PrP and a control allele. High-resolution NMR data suggest that the core of HRdup has a canonical PrP\textsuperscript{C} structure, yet a nascent β-structure is observed in the flexible elongated hydrophobic region of HRdup. In addition, using mouse PrP\textsuperscript{C} sequence, we observed that a methionine/valine polymorphism at codon 128 (equivalent of methionine/valine
Chapter 3 – Nascent β-structure at the N-terminus of HRdup

129 in human sequence) and oligomerization caused by high protein concentration affects conformational exchange dynamics at residue G130. We hypothesize that with the β-structure at the N-terminus, the hydrophobic region of HRdup can adopt a fully extended configuration and foldback to form an extended β-sheet with the existing β-sheet. We propose that these structures are early chemical events in disease pathogenesis.

INTRODUCTION

Prion infections are caused by the misfolding of cellular form of the prion protein, PrP\textsuperscript{C} into an aggregation-prone pathogenic form commonly referred to as PrP\textsuperscript{Sc} [1]. Encoded by the \textit{PRNP} on chromosome 20 in humans [2], PrP\textsuperscript{C} is a glycosylphosphatidylinositol (GPI)-anchored protein highly expressed in neurons. It consists of a structured, C-globular domain (residues 127-231, mouse numbering) and a flexible, unstructured N-terminus (residues 23-126). The structured C-terminus contains three α-helices and an antiparallel β-sheet [3], while the intrinsically disordered N-terminus features an octarepeat region, which is associated with metal ion binding [4] and also a conserved hydrophobic region (HR). While there have been speculations on the functions of PrP\textsuperscript{C} that include associations with copper homeostasis, given its ability to coordinate copper ions [5], this remains an active area of research.

Production of misfolded PrP may be initiated under a number of conditions, these including the inheritance of germline mutations. Thus Gerstmann-Sträussler-Scheinker (GSS) disease is the prototype form of genetic prion disease; it is an autosomal dominant multi-systemic neurological syndrome that can develop into frank dementia and with the patients commonly displaying a protracted clinical course [6]. GSS can arise from a range of mutations within the open reading frame of \textit{PRNP} gene. These mutations are associated with different clinical presentation, yet the most common clinical phenotypes include cerebellar ataxia and pyramidal signs admixed with cognitive decline [7].
One notable PRNP mutation discovered recently contains a 24-nucleic-acid insertion which encodes an 8-amino-acid internal duplication of the residues LGGLGGYV in the PrP hydrophobic region (residue 110-130). This mutant allele is referred to as HRdup [8] and occurs in cis to valine at codon 129 in PRNP, with a common M129V polymorphism being present in many human populations [9]. Previous studies have revealed that thermolysin digestion of brain homogenates transgenic (Tg) mice expressing the HRdup allele yields a 16kDa thermolysin-resistant signature fragment. This fragment derives from the N-terminal and central hydrophobic region of the HRdup form of PrP [10], and in turn comprises a potential precursor to the observed 8kDa PK-resistant fragment that is a common feature of most types of GSS disease [11].

Prior analyses have demonstrated that the insertion in HRdup does not disrupt the canonical antiparallel β-sheet present in wild type (WT) PrP or affect the structural stability of the globular domain of HRdup [10], leaving the structural perturbations that can lead to the spontaneous misfolding and assembling of HRdup unidentified. In this study, we have used high resolution multinuclear NMR spectroscopy to compare the structures and explore the molecular motions and dynamics of WT mouse PrPC (118–231, M128; i.e. methionine at residue 128 equivalent to human residue M129), mouse PrPC (118–231) V128, and mouse PrPC (118–231) V128 HRdup (henceforward referred to as WT, V128 and HRdup, respectively). Our data suggest that while the core of HRdup has a canonical PrP\textsuperscript{C} structure, a nascent β-structure was observed in the elongated hydrophobic region of HRdup, a region which is natively unstructured in the M128 and V128 forms of mouse PrP. We also observed that the exchange dynamics of glycine 130 are correlated with methionine/valine polymorphisms at codon 128 and with the propensity to form PrP oligomers by high protein concentration. We hypothesize that the newly-discovered, nascent β-structure initiates long-range alterations that convert PrP's N-terminal unstructured region into β-enriched forms that resist digestion by thermolysin and proteinase K.
**MATERIAL AND METHODS**

**Protein purification**

Expression constructs for moPrP (118-231), moPrP (118-231) V128, and moPrP (118-231) HRdup (referred herein to as WT, V128, and HRdup, respectively) and based on the PD444 vector were synthesized by ATUM (California, USA), with codon optimization in the sequences previously described [10]. The N-terminus of all three proteins had a histidine tag (his-tag) composed of six histidine residues plus a TEV cleavage site (ENLYFQ/G). Extra residues before the cleavage site were added for visualization purposes on SDS-PAGE when removing the his-tag. The natural-abundance-isotope material was expressed and purified in standard growth media, as used previously to prepare human PrP [12]. $^{13}$C and $^{15}$N isotope enriched proteins were prepared with the 1L LB substituted with one liter culture medium containing 1g of $^{15}$NH$_4$SO$_4$ (Aldrich, USA), 3g $^{13}$C–glucose (Isotech Inc., USA), 1mM MgSO$_4$ (Caledon, Canada), 0.1mM CaCl$_2$ (Sigma-Aldrich, USA), 1mg of biotin (Sigma-Aldrich, USA) and 100mg of thiamine (Sigma-Aldrich, USA).

For the removal of the his-tag, unlabeled V128 lyophilized protein material was dissolved at 0.5mg/ml with 20mM sodium acetate (pH 5.5) and TEV (also containing a his-tag) was used at one part in one hundred by mass for 3.5 hours at 34°C in the presence of 0.5mM EDTA. After digestion was verified by SDS-PAGE, the solution was adjusted to pH7 then loaded onto a nickel affinity column equilibrated with solution (pH7) containing 10mM imidazole (Sigma-Aldrich, USA), 300mM NaCl (Fisher Scientific, USA) and 20mM Tris-HCl (Fisher Scientific, USA); the same solution was also applied to wash out the his-tag cleaved protein. The flow-through containing the V128 protein minus his-tag was dialyzed against water and then lyophilized.

**NMR samples**

2mg of lyophilized [$^{15}$N, $^{13}$C]-labeled or non-labeled protein material were dissolved in 500µl of NMR buffer comprising 95% double distilled water and 5% D$_2$O. The NMR buffer also contains 10mM of sodium acetate (pH5) and 0.25mM of deuterated 2.2-
dimethyl-2-silapentane-5-sulfonic acid (DSS-d$_6$) which acts as a chemical shift standard and is also used for estimating the protein concentration of the samples. The pH of the final solution was further titrated to pH5 with microliters of 2M hydrochloric acid, using the $^1$H NMR spectrum of acetate as a pH indicator.

**NMR spectroscopy**

The NMR experiments were all performed on a 600MHz Varian INOVA NMR spectrometer equipped with a triple resonance probe with Z-pulsed field gradients and a computer-controlled variable temperature (VT) module for temperature regulation. All the pulse sequences used are from Biopack (VnmrJ 3.2B, Varian Inc).

All the experiments were carried out at 30°C. The 1D spectrum were acquired with 14ppm spectral width, 1.5s relaxation delay and 2s acquisition delay and 1024 transients. The 2D $^1$H, $^{15}$N-HSQC NMR spectra were acquired with the water and gNhsqc pulse sequences of Biopack (Varian) with spectral widths of 14 ppm ($\omega_2$) and 40 ppm ($\omega_1$), 511 and 128 complex points, respectively; 32 transients were taken with relaxation delay of 1.5s. The 2D $^1$H, $^{13}$C-HSQC NMR spectra were taken with spectral widths of 14 ppm ($\omega_2$) and 80 ppm ($\omega_1$), 511 and 320 complex points, respectively; 64 transients were taken with relaxation delay of 1.5s. The 3D CBCA(CO)NNH and HNCACB were acquired with spectral widths of 14 ppm ($\omega_3$), 30 ppm ($\omega_2$) and 80 ppm ($\omega_1$) with 511, 42, and 64 complex points, respectively; 32 scans were taken and relaxation delay was set to 1.3s. The 3D $^{15}$N-NOESYHSQC spectra were taken acquired with spectral widths of 14 ppm ($\omega_3$), 10 ppm ($\omega_2$) and 30 ppm ($\omega_1$) with 511, 125 and 42 complex points, respectively; mixing time was set to 125ms and relaxation delay was set to 1.3s.

While the 1D $^1$H spectra were processed, and plotted with VnmrJ, all the 2D and 3D spectra were processed with NMRPipe and analyzed with NMRViewJ (v9.2.0-b4, One Moon Scientific) and CcpNmr (Version 2.4).
**Backbone amide relaxation data**

A series of 2D $^1$H-$^{15}$N HSQC $^{15}$N T$_2$ spectra were acquired at 60 MHz $^{15}$N frequency at 30°C using the gNhsqc pulse sequence (VnmrJ Biopack) to obtain the backbone amide relaxation data. These relaxation experiments were acquired with $^1$H and $^{15}$N spectral widths of 14 ppm ($\omega_2$) and 40 ppm ($\omega_1$), respectively, and with 2048 ($t_1$) * 192 ($t_2$) complex points. The $^{15}$N-T$_2$ measurements were taken with relaxation delays set to 10, 30, 50, 70, 90, and 110ms. The delay between repetitions of the pulse sequence was 4s to insure no sample heating. In the absence of extra relaxation measurements (such as chemical exchange) the R$_2$ (=1/T$_2$) values are given by:

$$R_2 = \text{Constant} \times S^2 \times t_M$$

where the rotational correlation time of the protein, $t_M$, is the macromolecular rotational correlation time which is proportional to the molecular weight and the order parameter $S^2$ for fast internal motions ranges from 0 to 1 (flexible to ordered). In the presence of fast to intermediate chemical exchange between conformations, R$_2$ would be enhanced [13, 14]:

$$R_2 = \text{Constant} \times S^2 \times t_M + R_{2\text{ exchange}}$$

R$_2$ values were obtained by fitting peak intensity from NMRViewJ versus time using in-house software xcrvfit with the following equations:

$$M_{xy}(t) = M_{xy}(0) \times e^{-t/T_2}$$

$$R_2 = \frac{1}{T_2}$$

**RESULTS**

**NMR spectra of V128 and HRdup versions of mouse PrP**

In general, the 2D $^1$H,$^{15}$N-HSQC NMR spectrum of a protein is one of the most sensitive indicators of the structure and dynamics of the protein. The 2D $^1$H,$^{15}$N-HSQC NMR spectra of WT, V128, and HRdup under these same experimental conditions are presented in **Fig. B-1 to B-3**. The M128 (WT) and V128 proteins have very similar 2D $^1$H,$^{15}$N-HSQC NMR spectra, and the spectrum of WT mouse PrP was comparable to previous studies [15].
The solubility of HRdup PrP was found to be highly sensitive to pH and to reach high enough concentration (~0.1mM) for NMR analysis, pH 4.0–5.0 was used for all the following studies. Despite of the high environmental sensitivity of the amide NH resonances, where even minor differences in sample condition can cause obvious chemical shift changes on spectra, only a few resonances from V128 were shifted versus the 2D $^1$H,$^{15}$N-HSQC NMR spectrum of HRdup (Fig. 3-1); these were mainly attributed to glycines in the flexible hydrophobic region that includes the LGGLGGYV insert. These glycines (including the extra 4 glycine resonances due to the insert in HRdup) are shown in detail in Fig. 3-1. In the 2D $^1$H,$^{13}$C-HSQC NMR spectra of HRdup, a total of six extra methyl groups were observed versus V128 as expected from the methyl groups of two leucines and one valine present in the insert (Fig. 3-2).

In summary, we observe that the resonances on the spectra of V128 remain at very similar, if not identical chemical shifts to HRdup, while the additional resonances of HRdup appear as “extra” peaks on the spectra. This suggests that even with an insert in its sequence, the core of the HRdup protein adopts the canonical PrP globular domain fold, as does V128 mPrP (118-231).

**Chemical shift assignment and novel β-structure detected in HRdup (118-231)**

NMR resonances corresponding to amino acids in the N-terminal region of the HRdup construct proved challenging to assign because of the highly repetitive sequence of the hydrophobic region of HRdup (mPrP 121-128 plus the insert) and its intrinsically disordered nature. The backbone chemical shift assignments (HN, N, C$_\alpha$ and C$_\beta$ nuclei) were completed to 80% with 3D CACBCONNH and NHCACB NMR spectra. The assignment for HRdup from V124 through the sixth resonance of the insert is shown in Fig. B-4. For clarity when dealing with duplicated sequences, for the following text we used a nomenclature scheme (Fig. 3-3 a) where the amino acids attributed to the additional resonances (i.e., those absent from analyses of the control V128 allele) are denoted as "X". (“X” stands for “extra”). Note that this scheme is different from the assignment of extra amino codons as assigned from inspection of the nucleotide sequence of the HRdup PRNP allele [8].)
Chemical shifts of NH, N, Cα and Cβ nuclei were loaded into TALOS+ program to calculate the secondary structure of HRdup [16] (Fig. 3-3), with only the assigned residues with a high propensity (>0.5) to be structured being presented. Residues with high α-helix propensity make up three α-helices, consistent with the helices in WT mPrP. Due to broadening of the resonances from residues in the anti-parallel β-sheet (HRdup 127-130) caused by chemical exchange (see below), the canonical β structure of PrP cannot be completely characterized in this analysis. It has been previously reported that the residues forming the first β-strand do not possess strong β character [17], but the presence of this canonical β-sheet in HRdup was already demonstrated strongly by dαα(L129, Y161) and dαα(Y127, R163) nOe’s observed in our previously published data [10].

Using the SEQSEE program [18], the residues VX2, LX3, GX4 and GX5 are predicted to have high β-propensity while the proceeding and following sequences are predicted to be random coil (data not shown). The dataset of BetaTPred 3.0 consists of 20142 protein chains (BT20142); the sequence VLGG is present in total of 220 counts, and in the 71 out of 220 counts a β-turn is formed by VLGG [19]. Consistent with these β-propensity prediction approaches, the most interesting feature of the TALOS+ analysis is a novel high β-propensity structure at valine X2 (VX2) and leucine X3 (LX3) (Fig. 3-3), indicating that a nascent β-structure is at the highly flexible N-terminus (see following section) of HRdup.

We did not observe the nOes between the residues with high β-propensity making contacts with the existing β-sheet. Rather, we propose that this high β-propensity structure is a nascent β-turn formed by valine at X2 (VX2), leucine at X3 (LX3) and glycine at X4 (GX4). In the 3D NHSQC-NOESY spectrum we observe the dNN(i, i+1) nOes that are consistent with a type 2 β-turn, where the NH of GX4 points at the carbonyl group of GX5 [20] (Fig. B-5).
In our previous molecular dynamics (MD) simulation analysis using HRdup fragment from residue 118 to 231 [10], this nascent β-turn existed in all three trajectories of HRdup and the structure(s) observed is consistent with the proposed type 2 β-turn. The MD simulation also shows the flexible and dynamic feature of this novel β-structure. 

**Fig. 3-4** presents a snapshot of VX2, LX3, GX4 and GX5 forming a type 2 β-turn at 7.28ns in trajectory 3 of HRdup (50ns in total).

**N-terminus 15N relaxation measurements**

The 15N transverse (R2) relaxation rates were measured to gain insight into the backbone motions of the WT, V128 and HRdup alleles. Given that the dynamics of the N-terminus is of the most interest, R2 values of the residues 118 to 155 of the three alleles are presented (**Fig. 3-5**).

The region from residue 119 to 125 is flexible region in all three alleles, as indicated by R2 values <10 s\(^{-1}\). These values are typical for intrinsically disordered parts of proteins [8]. The extra residues of HRdup at the N-terminus are also flexible given their small R2 values; the aforementioned nascent β-structure in this region does not appear to hinder the flexibility of the backbone in this region. Starting from residue 127 the three alleles become structured, supported by the increased R2 values resulting from the slow rotational tumbling of the structured core of the protein. The typical value of R2 for the structured core of a globular protein is approximated ~12 s\(^{-1}\) [14] which is close to our measured R2 at a protein concentration of 0.1mM (**Fig. 3-5**). Our data is consistent with previously reported R2 values of similar length PrP constructs such as mPrP (118-231) and bovine PrP (121-231) previously reported [15, 21].

When the concentration of the WT protein is ~0.2mM, the R2 values increased for all residues (albeit with the specific exception of residue G130, to be discussed below). Given that R2 values are directly correlated with molecular weight, these data suggest that oligomerization process takes place as the concentration of the solution increases. The increased R2 values are consistent with the R2 values of longer forms of PrP such as those starting at residue 90. For instance, average R2 values of ~15, 19, and 20 s\(^{-1}\)
for residue 127-140 have been published for PrP 90-231 at concentrations of 0.5mM, 0.8mM, and 1.0mM, respectively [22-24]. It has been inferred that presence of an N-terminal extension is correlated with increasing R₂ values; combining previously reported data with our own, it is likely that the central flexible region (90-120) of PrPC can enhance the PrPC oligomerization process [15].

Residue G130 has a higher R₂ value than the other residues in this region and Sullivan et al. [15] have suggested that it is due to chemical exchange causing broadening. An interesting observation is that the R₂ broadening is decreased at the higher concentrations, implying that motions at G130 near the first β-strand are altered due to the oligomerization process [25].

**Polymorphism at codon 128 affects the chemical exchange at G130**

G130 is located at the end of the first β-strand. In the 2D ¹H, ¹⁵N-HSQC NMR spectra of WT allele, its resonance locates at ~115ppm, ¹⁵N dimension and ~9.3ppm, ¹H dimension. However, under our experimental conditions, we noticed that in both the 2D ¹H, ¹⁵N-HSQC NMR spectra for V128 and HRdup, the G130 resonance is unobservable compared with WT where it is weak but observable. In the high signal-to-noise 1D ¹H NMR spectrum of V128 (using non-¹³C, ¹⁵N labeled protein), we observe two resonances, one at each side of G130 in WT spectrum (Fig. 3-6). This is highly consistent with what was reported by Liu et al. that there are two conformations exist for G130; the exchange between these two conformations can result in broadening of the resonance of the amide of G130 in 2D ¹H, ¹⁵N-HSQC NMR spectra [17], and that the separate resonances observed in the 1D ¹H NMR spectrum can correspond to different conformational states.

We attempted to simulate the lineshape of G130 at different exchange rate constant (kex) between the two conformations. Assuming the two observed peaks are from G130 in slow exchange, we found that for the lineshape of G130 to vary from a broadened singlet to two separate resonances, the kex decreased from approximately > 500 s⁻¹ to approximately < 50 s⁻¹. Therefore, under our experimental conditions, the Met/Val
polymorphism at codon 128 affects the chemical exchange at G130 and it is highly feasible that the slowed exchange is observed when a valine is at codon 128 instead of a methionine.

DISCUSSION

In the study of the HRdup protein, additional β structure was observed involving residues VX2 and LX3 of the insert. This raises the question that how the β1–β2 sheet is formed—do residues X1–X4 or 127–130 pair with β2 (see Fig. B-6), especially in light of the higher β-propensity of the former calculated by TALOS based upon the NMR chemical shifts? As described in detail in the Results section, we conclude that in the predominant structure of the HRdup protein, the structured globular C-terminal domain of HRdup has a canonical PrP_C structure, with a nascent β-structure in the flexible elongated HR, which we have suggested is important in the oligomerization of the protein. The reasons are summarized below: (1) all of the resonances from the insert appear without any perturbation of the resonances from the V128 protein such as residues 131–135 contiguous to β1 (Fig. 3-1); (2) the pattern of the R2 values reflecting mobility would be very different especially for VX2 and LX3, which would be immobilized; (3) the chemical shifts of the β1 strand resonances in previous structures do not show a high TALOS propensity; (4) the proposed structure of a flexible β-turn segment separated from the core is consistent with the MD simulations; and (5) secondary structure predictions show β-turn propensity for the residues in the insert. Nonetheless, it is impossible to rule out exchange between the two conformations. Importantly, as this β-structure maps within protease-resistant fragments observed after thermolysin treatment of brain material from young Tg HRdup mice and proteinase K treatment of material from aged Tg HRdup mice [10], a causal relationship to pathogenesis becomes plausible.

The genotypic composition of a high frequency Met/Val polymorphism at position 129 in the human PRNP gene (i.e., M/M versus M/V versus V/V) [9] has a profound impact upon the progression and manifestation of prion diseases. This effect applies irrespective of whether disease is caused by infection (from prion-contaminated food or medical products), or by germline mutation or if the disease arises in idiopathic form,
as is the case for sporadic CJD. Generally, heterozygosity at this polymorphism site can lead to substantial resistance towards prion disease; for the individuals from families with inherited prion disease if with Met/Val heterozygosity at this site, the age of the disease onset is significantly later [26, 27]. A spectacular example of how this polymorphism site can determine disease phenotype is the D178N germline mutation can result in genetic CJD or fatal familial insomnia depending upon whether V129 or M129 polymorphisms, respectively, are present in the chromosomal allele carrying the D to N mutation in codon 178 [28].

It was previously reported by Collinge and co-workers [24] that hPrP M129 and V129 possess the same canonical PrP\(^\text{C}\) structure and that similar molecular dynamics and stability were observed for both alleles. Our previously published results also demonstrated that mPrP M128 and V128 have the same regional stability when subjected to progressive denaturation induced with urea [10], and the results presented herein support the same canonical core structure (see above); however, the \(^{15}\text{N}\) relaxation data for mPrP M128 (WT) and V128 displayed different exchange dynamics at G130. We presented that under our experimental condition, when a valine is present at codon 128 the conformational exchange is slower as compared with methionine. Combined with our results where we demonstrated the relationship of the altered exchange dynamics of the first \(\beta\)-strand and the oligomerization of PrP\(^\text{C}\), we conclude that the Met/Val polymorphism at codon 128 can affect the initial oligomerization process of PrP\(^\text{C}\) by altering the exchange dynamics at G130, nearby the first \(\beta\)-strand of PrP\(^\text{C}\). This can be the structural implications which lead to different biological effects of the polymorphisms.

Incorporating all the aforementioned information, we hypothesize that in HRdup, with nascent \(\beta\)-turn structure present at the unstructured N-terminus, the elongated hydrophobic region can spontaneously adopt a fully extended configuration and fold back to contact the existing \(\beta\)-sheet and altogether develop into an extended antiparallel \(\beta\)-structure that consists of three \(\beta\)-strands (Fig. 3-7). Such process would be the very beginning of the misfolding and self-assembling of HRdup into a conformation high in
β-structure contents. Possibly, the stacking of this proposed β-sheet structure among molecules could subsequently lead to the formation of bigger molecular assemblies and eventually, amyloids. Thus, we hypothesize that the elongated hydrophobic region could be the birthplace of the misfolding of HRdup.

According to our hypothesis, the hydrophobic region could be highly involved in the initialization process of PrP\textsuperscript{C} transforming into its infamous form PrP\textsuperscript{Sc}, which is consistent with the determinant effects of a few point mutations located in this otherwise highly-conserved region; the drastically different downstream outcomes of these point mutations include different phenotypes of neurological diseases [29, 30] and complete resistance towards prion infection [31]. Meanwhile, as the canonical anti-parallel β-sheet of PrP\textsuperscript{C} is incorporated into this proposed misfolding-triggering structure, this also helps explain the correlation of the dynamics of the first β-strand and the misfolding initialization process.

The proposed β-structure to be formed by hydrophobic region and the existing β-sheet has already been observed by Legname group using WT hPrP (structure shown in Fig. 3-7). Their elegant work demonstrated that when crystalized with nano-body at the N-terminal end, the hydrophobic region of full length hPrP adopts a fully extended conformation and forms a three-strand β-sheet with the canonical PrP\textsuperscript{C} anti-parallel β-sheet; most importantly, a β-turn structure was confirmed in the hydrophobic region [32]. Compared with this β-structure observed in full length human prion protein crystal, our proposed three-strand antiparallel β-sheet model for HRdup is longer thus potentially more stable, more likely to happen under biological conditions, which can explain the higher propensity for HRdup to self-transform into a β-sheet rich structure. Bearing in mind that hydrophobic region and the anti-parallel β-sheet (Y145stop) can form amyloids that can cause clinical prion disease with 100% attack rate [33], one can speculate that our proposed β-structure can also be an important component of a prion disease causing particle.
Fig. 3-1. Superimposition of $^1$H-$^{15}$N 2D HSQC NMR spectra of V128 and HRdup. The asterisk symbol indicates an overlap of two resonances in the $^1$H-$^{15}$N 2D HSQC NMR spectrum of HRdup; the blue arrows show resonances that exist on both spectra, while the red and black arrows indicate resonances that are unique on the spectrum of V128 and HRdup, respectively. The residues with asterisk in the numbering are from the residues artificially added in between the N-terminal His-tag and TEV cleavage site (discussed in Method section).
Fig. 3-2. Methyl region of 2D $^1$H, $^{13}$C-HSQC NMR spectra of V128 and HRdup. Upper spectrum is V129 and lower spectrum is HRdup. The extra resonances on the HRdup spectrum (indicated by blue square boxes) correspond to the methyl groups from the extra two leucine and one valine residues from the insert of HRdup.
**Fig. 3-3.** Secondary structure of HRdup calculated by TALOS+. Top panel shows the sequence numbering of the hydrophobic region of HRdup; the residues corresponding to the extra resonances on the NMR spectra due to the insertion are numbered X1 to X8. Lower panel shows the secondary structure prediction of HRdup from TALOS+. The canonical secondary structure is shown at the top of the plot (BMRB ID: 1AG2) [34]. All the sequences are aligned by the C-terminus. The residues with high propensity to be α-helices are indicated with positive values, while the residues with high propensity to be β-structures are indicated with negative values.
Fig. 3-4. Snapshot of MD simulation trajectory of HRdup. This diagram shows the novel \( \beta \)-structure that is consistent with NMR data and the proposed type 2 \( \beta \)-turn structure. This structure was taken from state No. 728 (out of 5001 states) in the 50ns trajectory 3 of HRdup [10].
Fig. 3-5. $^{15}$N R$_2$ relaxation data of HRdup and controls. In this diagram is data for 0.1 mM HRdup (blue), 0.1 mM V128 (green), 0.1 mM WT (yellow) and 0.2 mM WT (orange). The gaps are due to overlapping of resonances or missing resonances due to chemical exchange broadening.
Fig. 3-6. NMR spectra of G130. 2D $^1$H, $^{15}$N-HSQC NMR spectrum for M128 (top panel), 2D $^1$H, $^{15}$N-HSQC NMR spectrum for V128 (middle panel), and 1D $^1$H NMR spectrum of unlabeled V128 (bottom) with N-terminal his-tag removed. The arrows indicate the two resonances corresponding to the two conformations for G130; the asterisk in the 1D $^1$H NMR spectrum indicates the resonance caused by unknown impurity.
Fig. 3-7. Proposed misfolding mechanism of HRdup. At the top is the proposed misfolding-initializing structure of HRdup. The elongated hydrophobic region turns, folds back and contacts the canonical β-sheet. At the bottom is the three-strand antiparallel β-sheet structure reported by the Legname group [32].
REFERENCES


CHAPTER 4 – Brain metabolomics study of mouse scrapie

Metabolomic study of disease progression in scrapie prion infected mice; validation of a novel method for brain metabolite extraction

Ze-Lin Fu\textsuperscript{1,2}, Pascal Mercier\textsuperscript{3}, Ghazaleh Eskandari-Sedighi\textsuperscript{1,2}, Jing Yang\textsuperscript{2}, David Westaway\textsuperscript{1,2}, Brian D. Sykes\textsuperscript{1}

\textsuperscript{1} Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada
\textsuperscript{2} Centre for Prions and Protein Folding Diseases, University of Alberta, Edmonton, Alberta, Canada
\textsuperscript{3} National High Field Nuclear Magnetic Resonance Centre (NANUC), Edmonton, Alberta, Canada

Publication: This chapter has been re-submitted to journal of Metabolomics.

Contributions: DW, BDS and ZLF conceived and designed the research. JY and GES collected the samples. GES performed western blot analysis. ZLF conducted the metabolite extraction, NMR experiments and analyzed the data with PM. ZLF wrote the manuscript with DW and BDS.

Data availability: The metabolomics data reported in this paper are available at Metabolomics Workbench, study ID1871.

OVERVIEW

Introduction Prion disease is a form of neurodegenerative diseases caused by the misfolding and aggregation of cellular prion protein (PrPC). The neurotoxicity of the misfolded form of prion protein, PrPSc still remains understudied. Here we try to investigate this issue using a metabolomics approach.
Chapter 4 – Brain metabolomics study of mouse scrapie

**Objectives** The intention was to identify and quantify the small-in-size and water-soluble metabolites extracted from mice brains infected with the Rocky Mountain Laboratory isolate of mouse-adapted scrapie prions (RML) and track changes in these metabolites during disease evolution.

**Methods** A total of 73 mice were inoculated with RML prions or normal brain homogenate (NBH) control; brains were harvested at 30, 60, 90, 120 and 150 days post-inoculation (dpi). We devised a high-efficiency metabolite extraction method and used nuclear magnetic resonance (NMR) spectroscopy to identify and quantify 50 metabolites in the brain extracts. Data were analyzed using multivariate analysis.

**Results** Brain metabolome profiles of RML infected animals displayed continuous changes throughout the course of disease. Among the analyzed metabolites, the most significant changes included increases in myo-inositol and glutamine as well as decreases in 4-aminobutyrate (GABA), acetate, aspartate and taurine.

**Conclusion** We report a novel metabolite extraction method for lipid-rich tissue. As all the significant metabolites are identifiable and quantifiable by magnetic resonance spectroscopy (MRS), this study suggests that tracking of neurochemical profiles could be effective in monitoring the progression of neurodegenerative diseases and useful for assessing the efficacy of candidate therapeutics.

**INTRODUCTION**

Scrapie is the prototype of a class of neurodegenerative diseases caused by the misfolding and aggregation of normal cellular prion protein (PrP\(^{c}\)) into the beta-sheet-rich isoform PrP\(^{Sc}\), which composes the infectious prion agent [1]. Prion diseases are characterized by spongiform change, neuronal loss and astrocytic activation [2]; currently all are incurable and fatal. Metabolomics is a type of “omics” study that characterises the global biochemical compositions of biological samples [3]. By monitoring the metabolic profiles of biological matrices such as tissues, cells or the entire organisms during disease progression or other means of stimulations, metabolomics is capable of revealing the biological mechanism of the condition under study [4]. Nuclear Magnetic Resonance (NMR) is a non-invasive, untargeted and high-throughput approach for metabolites identification and quantification and has its
unique advantage in high reproducibility [5]. Using NMR metabolic profiling technique, the brains of animal models such as Alzheimer’s disease and Huntington disease were distinguished from their healthy controls [6-8], and by profiling serum and blood NMR also identified the metabolic and energetic differences between Huntington disease animals and their healthy controls in presymptomatic phase [9].

Measurements of brain metabolites in a patient of Creutzfeldt-Jakob disease (CJD; a human form of prion disease) using magnetic resonance spectroscopy (MRS) have been reported previously [10], as well as the measurements of the metabolites extracted from CJD-infected hamster brains [11]. Studies by Bourgognon J-M. and co-workers also analyzed the biochemical changes in mice 10 weeks after being inoculated with Rocky Mountain Laboratories scrapie prion (RML) using ultrahigh-performance liquid chromatography-tandem mass spectroscopy [12]. Nonetheless, we still lack a detailed profiling of the aqueous neurochemicals in prion infected brains, especially the fluctuations happening during silent (pre-clinical) phase of disease evolution.

In this study, we invented a novel extraction method which significantly improved metabolite yields from lipid-rich tissue to improve the further enable the detailed profiling of aqueous metabolites in the brains; and using this extraction method we investigated the metabolic perturbations in the brains of RML inoculated mice along the progression of prion disease at 30, 60, 90, 120 and 150 days post-inoculation (dpi). The results indicate that the brain metabolic profile goes through a continuous change as prion disease progresses, and that the changes of a few metabolites can be distinguishable before the presence of neurological symptoms.

**MATERIAL AND METHODS**

**General data analysis for metabolomics**

When analyzing the data of metabolomics studies, the key is to identify the correlations of the changes in the metabolite concentrations. Finding the association among the metabolites is essential for interpreting the mechanism of the biological process that contributes to the change in the metabolic profile. When doing metabolomics studies,
instead of working with one or few variables from a few samples, we operate with an immense amount of data generated from a vast number of samples. To identify the changes and especially the correlations among the numerous variables, the conventional univariate methods such as ANOVA and Student’s t-test are rendered ineffective and inadequate.

Multivariate analysis (MVA) is a series of statistical techniques that are capable of analysing data sets containing numerous variables and has valuable advantage of working with correlated variables for information extraction and classification [13]. In this thesis, two of the most commonly used multivariate analysis methods were employed and are discussed in the following sections.

**Principal Component Analysis (PCA)**

PCA is the basis for multivariate analysis; it is also one of the most popular statistical techniques and applied in many diverse scientific fields [14]. PCA has a long history, the origin of this statistical method can be traced back as far as 1901 by Pearson et al. [15] whilst the term principal component (PC) was first introduced by Hotelling at 1933 [16]. When given a matrix of data, PCA can help achieve several goals such as simplification, data reduction, outlier detection, classification and prediction [17].

The best way to interpret PCA is through a geometric representation. In the case of a data matrix X with N objects and K variables, we can represent the matrix X as an ensemble of N points in a K-dimensional space [17]. In PCA, by projecting the K-dimensional space into a lower-dimensional space while maintaining the maximum amount of the spatial distribution, we simplify the representation of the data matrix and that the correlations among the variables can be revealed. This is achieved by fitting the data with the first dimension (first principal component) through the direction that represents the most amount of variance in the K-dimensional space; mathematically this means in the least squares sense (minimizing the sum of the squares of the differences between the observed values and the estimated values). The subsequent
dimensions are fitted orthogonal to the previous dimension while representing the most amount of variance of the ensemble distribution [18] (Fig. 4-1).

The mathematical way of interpreting PCA is that PCA generates two smaller matrices, T matrix and P matrix to approximate the original X matrix using the following linear relationship [17]:

\[ X = TP^T + E \]

The T matrix is the score matrix, and the P matrix is named loading matrix. E accounts for the matrix of unexplained variance left by the model. While the X matrix is composed of N rows and K columns, the T matrix has N rows and D columns whilst the P matrix consists of D rows and K columns (Fig. 4-2). The number D is the number of PCs one decides to fit the model with, and to each PC there is a score vector \( t \) of all the observations that is determined by the K variables and the loading vector \( p \). The relationship between \( t \), \( p \) and PC is as following:

\[ t_i = p_i \cdot x_i \]

\[ PC: t = p_1 x_1 + p_2 x_2 + p_3 x_3 + \cdots + p_n x_n \]

On the score plot of PCA, the scores from different PC of each observation are plotted while the loading factors of the variables are plotted in the loading plot. The clustering of the observations on the score plot indicates the similar scores of these observations along the PCs presented in the graph. The loading plot demonstrates how the variables are ‘weighted’ towards the scores of the observations, which means how important the variables are for the patterns to show on the score plot and for the model building [17].

In the data generated in metabolomics studies, the object (observation) number N stands for sample numbers whilst the variable number K stands for the levels of all the metabolites measured in the experiments. Given the fact that PCA is an unsupervised approach, the natural correlations among the variables (metabolites) are best revealed by PCA. On the score plot, the clustering of the samples indicate their similar metabolic profiles and in the loading plot the important metabolites contributing to the distribution of samples can be identified by their larger loading factor values along different PCs.
With the correlations of the metabolites, we are offered clues of the possibly correlated biological events in the organisms and this is fundamental for deciphering the mechanisms of the biological event under study.

**Partial Least Squares Discriminant Analysis (PLS-DA)**

While PCA describes the variance of one data set, Linear Discriminant Analysis (LDA) is a different type of multivariate analysis that is optimized for minimizing the variance within each data set while maximizing the covariance between two data sets [19]. In the field of metabolomics, the most commonly used LDA algorithm is partial least squares regression (PLS-DA) [18]. In this thesis, PLS-DA was used to classify the samples into two different classes, that is to come up with an algorithm to generate a set of latent variables to maximize the correlation between the data set obtained from experiments and a regressor matrix containing two classes. This is the process that is generally referred to as ‘model building’ in metabolomics studies, and the end result should be a clear separation between two clusters on the PLS-DA score plot. A detailed mathematical deconstruction of PLS-DA is beyond the scope of this thesis.

The validity of the models built with PCA and PLS-DA are usually assessed by two parameters, $R^2$ and $Q^2$. $R^2$ stands for goodness of fit, which is an indicator of the scope the explained variance; $Q^2$ represents goodness of prediction, which is used to specify the predicting ability of the model. It can be arbitrary to estimate the effectiveness of the models solely using the values of $R^2$ and $Q^2$ as it is hard to set a uniform standard for the models built for different types of studies, but generally for biological application a model is considered acceptable when the $Q^2$ value exceeds 0.5.

Another rule of thumb for evaluating a model is that the difference between $R^2$ and $Q^2$ should not exceed 0.2. As the number of latent variable increases in the analysis, the model becomes closer and closer to the original data matrix thus the $R^2$ value keeps increasing; however, as more and more the biological meaningful variance is included into the model, more and more noise variance is also accounted which leads to the
worsened predicting ability of the model and a decreased $Q^2$ value (Fig.4-3). For this reason, as a general rule, an cut-off value 0.2 is set.

**Data combing: scaling and orthogonal projections to latent structures (O-PLS)**

It is unlikely that the data gathered from metabolomics studies are ready for the downstream analysis. The concentration of metabolites measured from experiments can differ by as much as 5000 times [20], and a variation like this can easily mask the biological significance of the metabolites at low levels as well as magnify the importance of the abundant metabolites. Herein, to avoid the extreme distortion of the biological relevance of the metabolites, a scaling step is mandatory for metabolomics experimental data. The scaling method employed in this thesis is called *pareto scaling* [21]; the data is first mean centered and then divided by the square root of the standard deviation. Pareto scaling is a commonly applied scaling method in the metabolomics field [6, 22, 23]; and the advantage of this scaling method is that not only does it effectively reduce the relative importance of large values, but it also keeps the structure of the original measurement partially intact [20].

During the experimental process of this thesis, another pre-treatment method was considered for the data combing process. Orthogonal projections to latent structures (O-PLS) is a generic preprocessing method for multivariate data analysis [24]. O-PLS enables to remove the variation of one data set that is not correlated to another set of variables; or in another way, the variations in one data set that is orthogonal to the covariance between two data sets can be removed using O-PLS. In this way, non-correlated noises can be filtered out safely instead of risking to lose meaningful information during the ‘de-noise’ steps while attempting to eliminate systematic noises [24].

In the metabolomics study covered in this thesis, O-PLS was considered to be used to remove some systematic error introduced during the experiment process due to different sample volume received. An alternative way to analyse the experimental data other than the PCA method employed in the end would be to remove the latent variables
in the O-PLS analysis that separates the samples by their volumes (instead of different time points); then with transformed data whose volume effect had been removed, a PLS analysis would be able to demonstrate the clear separations of the clusters of different time points. Although this method was not employed in the final version of this thesis, the potential of O-PLS removing the orthogonal variances is very much worth noticing and mentioning.

Method for metabolomics experimental study

Animal Procedures: RML inoculation and brain collection

Wild type (WT) FVB/Crl mice were used for this study. These animals were bred from the FVB mice purchased from Charles River laboratory. All the mice were housed in groups of up to five animals, maintained in ventilated racks (Tecniplast, Green Line). The animals were fed irradiated chow (LabDiets, 5053) and were under a 12h light/dark cycle with food and tap water *ad libitum*. The cages were cleaned and maintained on daily basis. The mice were subjected to intracerebral injection to be inoculated with 30µl of 1% (w/v) RML brain homogenate or normal brain homogenate (NBH) at 3-6 weeks of age. At each timepoint, a sub-set of RML and NBH inoculated animals were euthanized (numbers are listed in Table C-1). The brains were harvested shortly after cerebral dislocation was performed; each brain was dissected into two hemispheres by cutting along the middle line, flash frozen with dry ice and then stored at -80°C.

Brain tissue extraction for NMR spectroscopy

We devised a new metabolite extraction method for this study. Instead of introducing any strong acid or hydrophobic chemical reagents as per conventional brain metabolites extraction methods, we employed ultracentrifuge and centrifugal filtration to remove the unwanted high molecular weight hydrophobic molecules. The extraction efficiency comparison between this new protocol and a conventional protocol is detailed in the discussion section. The right hemispheres of the brains (the left hemispheres were used for other experimental purposes) were weighed and homogenized in fresh double distilled water (ddH2O) using glass Dounce homogenizers to make 10% (w/v) homogenate. The homogenizers were washed with 5% LPH and thoroughly rinsed with
water between the samples. The homogenate was subsequently subject to ultracentrifuge at 100,000xg for 1 hour at 4°C, and the resulting supernatant was neutralized with Tris buffer (0.5M, pH 6.8) one tenth the volume of the supernatant. The pH neutralized homogenate was then passed through 3kDa molecular weight cut-off (MWCO) filters at 4,000 rpm for 45 mins at 4°C; the filters had been washed with 0.1M NaOH solution once and ddH2O for three times. The filtered supernatant was dried on a SpeedVac Vacuum Concentrator at low heating setting until dry pellets were received and ready for NMR analysis.

**NMR analysis of brain sample extracts**

The pellets from the speed-vac were dissolved in 530µl of ddH2O and 60µl of D2O containing 5mM of 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS-d6, Chenomx Inc., Canada). The samples with a 600µl final volume were subsequently transferred into 5mm NMR tubes. The DSS-d6 (0.5mM in the samples) serves as a chemical shift indicator and concentration standard. All spectra were acquired at 30°C on a 500MHz INOVA NMR spectrometer equipped with an Agilent OneProbe with VT, PFG which has two channels. The pulse sequence used was tnnoesy (Vnmr 6.1B software, non-Biopack, Varian Inc.), and the experiments were acquired with 1s pre-acquisition delay, 12 ppm sweep width, 4s of acquisition time and 0.1s of mixing time. For each spectrum 256 transients were taken.

**Metabolite identification**

After data acquisition, Fourier transforming and initial phasing of the spectra were done in VnmrJ (version 3.2B, Varian Inc.). All spectra were then loaded into Chenomx NMR Processor version 8.4 (Chenomx NMR Suite 8.4, Chenomx Inc., Canada) for more fine phasing and baseline correction. The metabolite assignment was carried out in Chenomx NMR Profiler version 8.4 (Chenomx NMR Suite 8.4, Chenomx Inc., Canada); using the concentration of DSS-d6 as standard, the metabolites in the brain extract were identified and quantified.

**Multivariate Statistical analysis**
Data sets were normalized to the level of total creatine (creatine and phosphocreatine), a commonly used normalization method in neurochemical profiling [25]. In SIMCA-P+ 12.0.1 software (Umetrics, Sweden), the concentrations of all profiled metabolites were scaled using Pareto scaling and then used as variables for unsupervised principal component analysis (PCA).

The parameter R² is generated by SIMCA-P software to evaluate the performance of the model. The R² score indicates the amount of the total variance of the data described by one of the principal components, and R²cum indicates the amount of variance described by all of the components of the model (fitting of the model). It is a value between 0 and 1; the higher the value the more variance of the datasets is captured by the model.

T-test was performed in Microsoft Excel 2016 to determine the significance of the differences between RML and NBH groups at different timepoints. The t-test performed is two-tail t-test assuming unequal variance.

RESULTS

Multivariate statistical analysis of brain metabolites

In this study, we examined the metabolic profile of brain extracts of RML infected mice and the negative control animals inoculated with normal brain homogenate. The RML inoculated animals displayed neurological symptoms between 120 and 150 dpi. A step-by-step description of our procedure is found in the methods section. Fig. 4-4 shows the aliphatic region of the overlay of thirty 1D ¹H NMR spectra obtained using this approach with the major metabolites identified (downfield region of the spectral overlay is shown in Fig.C-1). The quality of the spectral overlay is very high indicating the stability of the extraction procedure with respect to variables such as pH and overall purity relative to higher MW components. Among all 73 spectra, in total of 50 metabolites were identified and measured (a list of all the metabolites is shown in Table C-2). In Table 4-1 we list the concentrations measured in our study as well as the values reported for mouse brain metabolites extracted using perchloric acid method [7].
Notably, there is a general increase of the concentrations when compared with previously reported values [26].

**Brain metabolomic changes with the progression of prion disease**

The PCA score plot of our data is shown in Figure 4-5. The first component on the x-axis explains 41% of the variation whilst the second component on the y-axis explains 23% of the variance. In total, there are five principal components and $R^2_{cum}=0.816$ (the explained variance of all components are plotted in Fig. C-2). In this score plot, all the control samples of different timepoints, together with the two clusters from the first and second timepoint (30 and 60 dpi) of the RML infected animals form a large cluster on the lower right part of the plot. In Fig. 4-5 b a close-up view demonstrates the overlay of the first and second RML timepoints (green and blue) with their controls (empty squares in corresponding color) and emphasizes the similarity of the metabolic composition. In short, up to 60 dpi, the brain metabolism of RML-infected animals is close to that of healthy brains.

The third timepoint of RML infected animals (yellow solid boxes) separate into a band slightly to the left and above the aforementioned large cluster, while the fourth (red solid boxes) and fifth (purple solid boxes) timepoints of RML inoculated animals are displaced progressively further from the large cluster. The continuous spread along a vector starting in the bottom-right and moving to the upper-left direction indicates a progressive change in brain metabolism as disease advances; the separation from the third timepoint onwards, i.e. at 90 dpi, is notable and coincides with a transition documented in prior studies [27].

Two outliers from the fourth and fifth timepoint of RML-infected groups were noted as not fitting the overall pattern. They are animal ID# 82806 from the fourth timepoint and #82864 from the fifth timepoint. In Figure 4-5, they are located in the large cluster on the bottom-right part of the plot, next to the NBH control animals. The spectra of both samples are similar to the spectra of the other samples from the same group; a comparison of the spectra is presented in Fig. C-3. Both animal were scored as showing
clinical symptoms of disease before they were euthanized and both show clear signals of PK-resistant PrP in western blot analyses of brain homogenates (data not shown). It is unknown why these two RML-positive animals do not present with a similar neurochemical profile to the other animals in their experimental groups; their levels of the key metabolites identified in this study (see below) are plotted in Fig. 4-7 next to the average levels of their groups.

**Metabolites contributing to distribution of the samples groups**

In the loading plot of this PCA (Fig. 4-6), the first and second principal component are displayed on the x and y axis, respectively. The loading figure shows the contribution of each metabolite for the separation among the sample types in the score plot; the metabolites furthest away from the center of the plot (with large coordinate values) are the ones most responsible for the clustering of the groups. And similarly located metabolites on the loading plot and the samples on the score plot reflects the increased relative concentration of these metabolites in the corresponding samples; whilst the metabolites with decreased relative concentrations are situated on the opposite side of the loading plot across from the zero point of the plot. From Figure 4-5 and Figure 4-6, one can see that with the progression of the disease, the concentrations of myo-inositol, glutamine and glucose increased while aspartate, acetate, 4-aminobutyrate (GABA) and taurine have decreased in the sick animals. The levels of these metabolites at different timepoints are plotted in Figure 4-7 and this indicates that these metabolites change continuously as the disease progresses (glucose is discussed separately in later section). In this figure, RML sample 82806 from the fourth timepoint and RML sample 82864 from the fifth timepoint are indicated differently with triangles; one can see how the unchanged concentrations of these metabolites would not distinguish these two samples from their controls.

At third timepoint, which is 90 dpi and is at pre-clinical stage, on the PCA score plot we observe see some separation between the RML animals and NBH animals. In Figure 4-7, it appears that there is some differentiation at this timepoint for 4-Aminobutyrate, taurine, acetate and aspartate. To verify this observation, we carried
out two-tail t-test and we saw evidence for a difference in the concentrations of these four metabolites (p<0.05). The other p values of these metabolites at different timepoints are listed in Table C-3.

**PCA does not distinguish different timepoints of NBH controls**

In our study, different timepoints of NBH samples are randomly distributed on the score plot. A separate PCA using only the NBH samples also showed a similar scattered pattern, with $R^2_{cum}=0.70$ (score plot is in Fig. 4-8). This result demonstrated that in our study, age-dependent neurochemical changes are not the dominating differentiating factor. The samples at the first timepoint (30 dpi) were obtained when the animals passed 8 weeks of age, at which time the animals have finished the great majority of development and reached sexual maturity; whilst at the last timepoint, the brains were collected when the animals were about 6 months old, which is greatly in advance of the time when the aging factor should be taken into consideration [28], relating to an average life expectancy in excess of 2 years.

**DISCUSSION**

**Novel efficient brain metabolite extraction protocol**

In this study, we employed a novel method to extract the small-size water soluble metabolites from brains. The two most commonly used metabolite extraction protocols are perchloric acid (PCA) extraction and methanol-chloroform-water (M/C) extraction. PCA extraction method utilizes perchloric acid to precipitate out the fat and lipids in the brain homogenate; while in M/C extraction methanol and chloroform are introduced to create hydrophobic phase and aqueous phase, the latter of which contains small and soluble metabolites and is used for metabolomics studies. The elegant work form J. E. Le Belle *et al.* [29] compared these two extraction methods and concluded that M/C method is more efficient and consistent in extracting metabolites from both brain tissues and cells than PCA method.

By including an ultracentrifuge step and a centrifugal filtration step, our metabolite extraction protocol circumvents the introduction of strong chemical reagents, therefore
keeping the samples in a milieu that is closer to physiological condition. This greatly preserves the original net metabolic profiles. This procedure is efficient in removing the high-molecular-weight compounds (Fig. C-1); more importantly, it yields much higher extraction efficiency than the conventional extraction methods. For metabolites such as alanine, creatine, glutamate, glutamine and taurine, the concentrations are close to their in vivo measurements determined by MRS [30]. The drawback of this method, however, is that inactivation of metabolic enzymes within the tissue samples was not addressed when compared, for example, with PCA extraction. This is suggested by the increased lactate level compared with in vivo measurements; and this is also the reason why lactate (and only lactate among all 50 detected metabolites listed in Table C-2) was left out in our data analysis. This also jeopardized our ability to analyze glucose and NAA (Fig. C-4); the enzymes in the tissue kept on breaking these two metabolites hence the relatively big detecting range. Nonetheless, by incorporating enzymatic inhibitors during the homogenization stage this issue may be avoided in future iterations. Thus, this novel method promises to be an effective new protocol for hydrophilic brain metabolites extraction.

In year 2018, Bourgognon et al. carried out an elegant metabolomics study on the changes of neuronal metabolism caused by prion disease [12]. Using liquid chromatography-tandem mass spectroscopy, they detected hundreds of biochemicals and reported that there are significant changes associated with glucose metabolism, neuropeptides, fatty acids, L-arginine/nitric oxide and prostaglandins at 70dpi. Compared with this study, instead of bigger biomolecules such as lipids and enzymes which are removed in the filtration step in our extraction step, our study primarily focuses on small-size soluble metabolites. Our results agree on the decrease of the levels of NAA and GABA, but we further provide information on the changes of many more water soluble metabolites such as myo-inositol, glutamine, acetate and aspartate. More importantly, more than just the increase or decrease end results, we also report the quantitated metabolite concentrations and their accurate changes in the brain; and by employing a longitudinal experimental design, our results are more indicative of the changes of the metabolic states in the brain during the progression of the disease.
Metabolic changes identified during the disease progression

We observed continuous changes for these following metabolites: glutamine, myo-inositol, acetate, GABA, aspartate and taurine during the progression of the disease. Myo-inositol is a sugar-like molecule and is an important osmolyte in glial cells [31]. Elevated myo-inositol concentration has been observed in numerous neurodegenerative diseases including Alzheimer’s disease [32, 33], Huntington’s disease [34] and Parkinson’s disease [35] along with some other diseases such as bipolar disorder [36] and human immunodeficiency virus infection (HIV) [37]. For prion disease studies, in one human CJD patient, the level of in cerebral myo-inositol measured by MRS increased by 30% after the patient was in coma [10]; whilst in CJD-infected Syrian hamsters, after the display of clinical signs as well as astrocytic activation and vacuolization, a 67% increase in myo-inositol concentration in the brain extract was reported [11]. In our study, we observed consistent increase of myo-inositol concentration in the RML infected animals as they advanced in the course of the disease and at the last two timepoints, an increase of 17.4% and 26.2% is observed. At the third-timepoint (i.e., 90 dpi) when the RML animals can be distinguished from their controls with 91% accuracy, myo-inositol did not, however, show a significant increase (p>0.05).

The increase of myo-inositol concentration has been correlated with microglial activation [31]. In prion disease, with gliosis being one of the disease signatures [2], this elevation is speculated to be caused by the breakdown of glycosylinositol phospholipids; and the free myo-inositol spilled from the dying neurons would further accentuate the activation of astroglia due to the uptake of myo-inositol by astroglia [11].

This mechanism fits with the commonly observed simultaneous decrease of N-acetylaspartate (NAA) and upregulation of myo-inositol. NAA is an abundant amino acid in CNS and a decrease in NAA concentration is widely considered as a marker of neuronal loss or dysfunction (Jones and Waldman, 2004). In our study, lack of a systematic interventions to inhibit metabolic enzymes precludes the direct study of
NAA itself; however, the levels of acetate and aspartate, the two products of NAA breakdown, both steadily decreased during the disease progression. In fact, at 90 dpi, both acetate and aspartate have significantly lower concentrations ($p<0.05$) in RML infected animals as compared with their controls.

One puzzling finding, was that the level of taurine decreased in the RML animals when compared with their age matched NBH-inoculated controls. Both being osmolytes found in mammalian cells, taurine and myo-inositol concentrations can react synchronously when neuronal metabolism is perturbed [38, 39]. Moreover, in APP mice model where taurine is considered to play the major role of osmoregulator, an increase in taurine concentration has been reported [40]. Nevertheless, to the best of our knowledge, the change of taurine concentration in prion infected mice has never been reported and the decrease of taurine could comprise a prion disease-specific event.

Other significant results are the upregulation of glutamine and the downregulation of GABA. In a Huntington disease mouse model, similar opposite changes for glutamine and GABA concentrations were observed [8]. Glutamine is the storage and precursor form of the major excitatory neurotransmitter glutamate while itself is inert in synaptic transmission. Glutamine plays the key role in the metabolism of both glutamate and the primary inhibitory neurotransmitter GABA [41]. GABA is produced in GABAergic neurons by the action of glutamate decarboxylase (GAD) on glutamate [41]. The glutamate however, arises from glutaminase acting on glutamine substrate synthesized by astrocytes [42]. One way to explain the simultaneous increase of glutamine and decrease of GABA is in terms of dysfunction of GABAergic neurons; as the activity of GABAergic neurons is affected by prion disease, the reduced synthesis of GABA leads to the elevation of glutamine concentration as the substance. It is worth mentioning that at the third timepoint, the level of GABA is significantly decreased while the level of myo-inositol remained unchanged. If our inference is correct, this result suggests that the dysfunction of GABAergic neurons is likely to happen before the activation of astrocytes in the course of disease progression.
CONCLUSION

In this study, we applied a novel brain aqueous metabolite extraction protocol which significantly improved the removal of lipid molecules and metabolite extraction efficiency and consistently yields ubiquitous sample condition for detection without further adjustments. By employing an experimental design that tracks the neurochemical changes along the progression of the disease, this study demonstrates that in the case of prion disease, the change of brain metabolism is discernible before the symptoms start to show; and that the changes in neurochemical composition continue to advance as the disease progresses into the late stage. As all the discussed significant metabolites are identifiable and quantifiable by MRS [30], this study demonstrates that tracking of brain neurochemical profiles promises to be an effective way to monitor the progression of neurodegenerative diseases in humans. Thus, this study shines light on the potentially important role of brain metabolomics in assisting evaluate the therapeutic efficacy in clinical trials for treating these devastating illnesses.
**Fig. 4-1.** Model of Projection. The observations are indicated with yellow circles and $t_1$ and $t_2$ represent the first and second principal components. This figure was published by Susanne Wiklund [43].
Fig. 4-2. Geometric representation of PCA matrices. The dataset (X-matrix) are decomposed into the score matrix (T), loadings matrix (P) and error matrix (E). $d$ is the number of principal components; $k$ stands for the number of variables (metabolites) and $n$ stands for the number of observations. This figure was adapted from a publication by Susanne Wiklund [43].
Fig. 4-3. Number of latent variables and goodness of fit and prediction. As the number of the number of latent variables increases along the x-axis, both goodness of fit and goodness of prediction increases; however, when all the meaningful biological variance is explained, the model would start to fit on noise and decreases the predicting ability of the model. This figure was published by Susanne Wiklund [43].
Fig. 4-4. The aliphatic region of the overlay of 30 high resolution 500 MHz 1H-NMR spectra. Some major metabolites are indicated. GABA, γ-aminobutyric acid; Glu, glutamate; Gln, glutamine; NAA, N-acetyl-aspartate.
Fig. 4-5. Unsupervised PCA model plots. They are of metabolic changes identified in brain extracts from all 73 RML and NBH treated mice at all five timepoints. All metabolites but lactate and creatine (48 in total) are used to build this model. **A)** Score plot showing the progressive change in the brain metabolism as the disease advances, x-axis: first principal component (PC); y-axis: second PC; R²=0.816, Q²=0.608. RML from first to the fifth timepoints are indicated in solid boxes and colored green, blue, yellow, red and purple, respectively; whilst all five timepoints of NBH control animals are indicated in empty squares and uniformly colored grey for simplification; **B)** Close-up view of the lower right region of the previous plot, now NBH controls (empty squares) are also color coded by different timepoints the same way as RML samples (solid boxes).
Fig. 4-6. PCA loading plot of all samples. This diagram demonstrates the details of how the metabolites contribute to the distribution of the samples. The metabolites on the upper left corner have increased concentrations in RML animals in the later timepoints; the decreased metabolites in these samples are located in the bottom right corner.
**Fig. 4-7.** The concentrations of the metabolites of different timepoints. They are diagrams of NBH at 30, 60, 90, 120, 150 dpi timepoints (NBH1 to NBH5, solid circles) plotted next to their concentrations of RML at corresponding timepoints (RML1-RML5, solid boxes). Each dot represents one sample. The outlier samples 82806 from RML4 and 82864 from RML5 are specially indicated with triangles.
**Fig. 4-8.** Unsupervised PCA on all 34 NBH control samples. The dataset contains 48 metabolites (with creatine and lactate removed from all the detected metabolites). The first to fifth timepoints are colored green, blue, yellow, red and purple, respectively. \( R^2_{\text{cum}} = 0.702, A=3. \)
### Table 4-1

Comparison of different extraction methods. This table shows the comparison between the brain metabolite extraction efficiency between the novel protocol and conventional PCA extraction reported by Graham S. F. et al. [22]. The concentrations are in µmol per gram of wet brain weight. The metabolite levels of the novel extraction protocol are the average value of three NBH animals (82661, 82662, 82664) euthanized at 150dpi while the reported concentrations obtained via PCA extraction are averages from six 12 month old C57BL/6J mice.

<table>
<thead>
<tr>
<th>Metabolites</th>
<th>PCA extraction*</th>
<th>Average</th>
<th></th>
<th>SED</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>umol/g</td>
<td>SED</td>
<td>umol/g</td>
<td>SED</td>
<td></td>
</tr>
<tr>
<td>Lactate</td>
<td>2.34</td>
<td>0.329</td>
<td>7.51</td>
<td>1.465</td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>0.103</td>
<td>0.018</td>
<td>0.938</td>
<td>0.341</td>
<td></td>
</tr>
<tr>
<td>GABA</td>
<td>0.215</td>
<td>0.024</td>
<td>6.206</td>
<td>0.914</td>
<td></td>
</tr>
<tr>
<td>NAA</td>
<td>0.5</td>
<td>0.042</td>
<td>0.669</td>
<td>0.369</td>
<td></td>
</tr>
<tr>
<td>Glutamine</td>
<td>0.331</td>
<td>0.04</td>
<td>2.375</td>
<td>0.529</td>
<td></td>
</tr>
<tr>
<td>Glutamate</td>
<td>0.096</td>
<td>0.007</td>
<td>2.952</td>
<td>0.279</td>
<td></td>
</tr>
<tr>
<td>Succinate</td>
<td>0.079</td>
<td>0.008</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Creatine</td>
<td>1.148</td>
<td>0.112</td>
<td>6.229</td>
<td>0.526</td>
<td></td>
</tr>
<tr>
<td>Choline</td>
<td>0.532</td>
<td>0.134</td>
<td>0.728</td>
<td>0.188</td>
<td></td>
</tr>
<tr>
<td>Glycerophosphocholine</td>
<td>0.984</td>
<td>0.148</td>
<td>0.235</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>Taurine</td>
<td>0.372</td>
<td>0.045</td>
<td>6.93</td>
<td>0.645</td>
<td></td>
</tr>
<tr>
<td>Myo-inositol</td>
<td>0.102</td>
<td>0.012</td>
<td>5.136</td>
<td>0.641</td>
<td></td>
</tr>
<tr>
<td>Ascorbate</td>
<td>0.032</td>
<td>0.003</td>
<td>1.044</td>
<td>0.114</td>
<td></td>
</tr>
<tr>
<td>Adenosine</td>
<td>0.007</td>
<td>0.001</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>
REFERENCES


Chapter 4 – Brain metabolomics study of mouse scrapie


CHAPTER 5 – Conclusions

For the longest time I did not realize that the primary key word for my PhD thesis would be “prion”. With all the time and energy I have spent on doing protein cultures, preparing samples, running NMR experiments, assigning and profiling NMR spectra, it was hardly clear to me that my work was to contribute to a better understanding of the mechanism of prion diseases.

SECTION I: Structural study of HRdup

The first part of my research work was focused on characterizing the molecular signatures of a novel PrP mutant with an 8-amino-acid insert in the central hydrophobic core region of PrPC; this mutant is termed HRdup. Unlike other point mutations of PrPC, HRdup is a specially intriguing research subject as these extra residues duplicate the sequence for the first $\beta$-strand of PrPC, which means potentially we could discover a different $\beta$-structure in the molecule as opposed to the canonical antiparallel $\beta$-sheet of PrPC. Meanwhile, as the detachment of the $\beta$-sheet from the helices was proved to be the initial structural change when PrPC is subject to denaturation [1], this insert was suspected to affect the molecular stability of the mutant as well as altering the canonical PrPC structure. If proven to be true, the decreased stability of the molecule could be responsible for the spontaneous misfolding and aggregating that leads to GSS disease.

Employing the methodology established by Julien et al. [1], I added increasing amount of deuterated urea into the protein solution while taking 1D NMR spectrum to monitor the misfolding process. With the resonances from five residues which act as probes for detecting the structural changes in HRdup molecule as well as the controls, we were able to evaluate the structural stability of HRdup which is similar to WT PrPC molecule. And 2D nuclear Overhauser effect spectroscopy (NOESY) experiment revealed that HRdup maintains a canonical antiparallel $\beta$-sheet. At this point, I was just starting to realize the power of NMR technique and truly enjoyed my experience working with
the 800 MHz spectrometer at NANUC (and those were some long working days); however these solid and meaningful scientific observations at the time definitely made me feel rather frustrated because these “negative” results certainly busted my hopes and dreams of winning any major scientific awards for this work, or even at the very least, one or two publications in Nature or Science. I was so disappointed.

After having decided that it was probably just me being naïve, I proceeded with my experiments and made $^{13}$C, $^{15}$N-double-labeled material of HRdup and the controls. HRdup, with all the duplicated repetitive sequence of a string of glycine residues in the intrinsically disordered region, is surely not the best choice of material to learn how to assign NMR spectra. Months of work went in and I was making poor progress, especially in the key region where the insert is located. It was some fortune of mine that Dr. Peter Holmes had finished his PhD and returned to work as a post-doc in Sykes lab; him sorting through the assignment of this very challenging part as well as pioneering the TALOS analysis of HRdup were crucial for this project to come through.

With this portion of the data, we learned that the extra residues sit in the central hydrophobic region of HRdup before the first $\beta$-strand; however, a few residues in the duplicated sequence displayed significantly increased $\beta$-propensity that is indicative of the presence of a $\beta$-turn in this otherwise highly disordered region. Because HRdup has a valine at codon 129 (human numbering), we included both V128 (mouse numbering) and M128 as our controls. Interestingly, our data revealed that the exchange dynamics between two conformations at G130 (mouse numbering) is affected by both the M/V polymorphism at 128, as well as the oligomerization induced by high protein concentration. These are all very interesting experimental evidence for understanding the initial misfolding process of PrP$^C$. The paper took a few back-and-forth maneuvers to be published, but it felt fantastic because I could finally label myself as someone who had contributed a small part to the frontier of human knowledge.
The following steps

In year 2013, the year when I started my graduate school, Abskharon et al. reported that when the N-terminus of PrP\textsuperscript{C} is tethered to an antibody, a three-strand antiparallel extended $\beta$-sheet containing a $\beta$-turn was observed in the central hydrophobic region using X-ray [2]. With this piece of evidence, it is sensible to hypothesize that in HRdup the extra residues bring in a pre-existing $\beta$-turn and that this structure would lead to a spontaneous extension of the canonical $\beta$-sheet. Subsequently, these $\beta$-sheet could seed the stacking and aggregation of the PrP\textsuperscript{C} molecules therefore causing the disease. It is worth mentioning that the building block of the pK resistant material in GSS disease is the central region of PrP\textsuperscript{C} (70-150) [3], encompassing the entire region of the proposed $\beta$-sheet. The discovery of an additional moiety with high $\beta$-propensity in HRdup is potentially revealing a crucial step during the initial PrP\textsuperscript{C} misfolding process. Some following experiments should be performed to verify this promising hypothesis.

One of the experiments to bring this research topic to the next level is to monitor the structural change of the monomers during an induced oligomerization process. Glaves et al. explored the methodology that using two-dimensional solution NMR spectra, the residue-specific structural change of PrP\textsuperscript{C} can be monitored during the oligomerization [4]. Using this method, by observing the resonances of the central region of HRdup, we should be able to monitor the structural change of this flexible region and to verify the formation of the proposed extended $\beta$-sheet. This highly attainable experiment could potentially yield some very significant results regarding the initial misfolding process of HRdup.

The protein material to be used, however, should be the C-terminal fragment of the mutant starting from roughly residue 90. For previous studies I went with a protein design starting from residue 118 which is bit short for observing the behavior of the central region of HRdup. A fragment starting from 90 would correspond to the C2 fragment coming from $\beta$-cleavage of PrP\textsuperscript{C} and it might be a more biologically interesting fragment to study.
The next frontier of structural biology?

After spending the past few years researching the structural dynamics of HRdup, I am still rather amazed at how subtle the differences are between HRdup and WT PrP<sup>C</sup> molecules in their structural dynamics, even with eight extra residues. In fact, the additional β-turn in HRdup is already a big structural alteration when compared with some other disease-causing mutants of PrP<sup>C</sup>, where only minor changes such as a loop region exposing the hydrophobic side more to the solvent, different orientation of helices and surface electrostatic potential [5-7]. Meanwhile, some variations of PrP<sup>C</sup> molecules that have strong protective roles towards prion diseases also display very minor structural changes. The famous G127V protective mutation was only demonstrated to have two of its β-strands slightly further apart and the first β-strand is more flexible, which is not surprising given the bigger side chain of valine, as well as some different loop and helices orientations [8]. Moreover, the initial study of the polymorphisms at codon 129 barely even observed any differences between M129 and V129 PrP<sup>C</sup> [9]. How would these miniscule structural differences lead to such drastically different disease outcomes?

From my own experience of working with proteins, I learned that while some proteins are very well behaved and can tolerate relatively big changes of experimental conditions, some can be very tricky and can display some very different properties at only slight different conditions. For instance, the concentration of HRdup solution can reach up to about 0.1mM when the pH is at 5 or slightly lower, whilst when pH is higher than 5 (once I tried pH 5.3) only very miniscule of protein would go into solution. For WT PrP<sup>C</sup> 90-231 fragment, when the protein concentration is at 0.5, 0.8 and 1.0 mM, the average R<sub>2</sub> values for residue 127-140 are roughly 15, 19 and 20 s<sup>-1</sup> [10]. It is likely that the increase of R<sub>2</sub> values is caused by the oligomerization induced by increased protein concentration. Evidence like this not only makes me cautious when reading about the experiment results from other labs, but also makes me wonder, if we could get closer to the physiological condition, would we be seeing different structural behavior and dynamics of PrP<sup>C</sup>? Would we observe much larger differences among the mutants of PrP<sup>C</sup>?
The prevalent usage of NMR spectroscopy is partially due to the first PrP\textsuperscript{C} structures being deciphered using NMR [11, 12]; on the other hand, X-ray crystallography probably would not have been as informative as NMR in reporting the structural dynamics of PrP\textsuperscript{C} molecules, as the molecules would be in crystalized form instead of in solution. Having a clear idea of the structure of the proteins is the fundamental step of studying these biomolecules, however, to learn more about their true properties and functions we need more than just high definition. In my opinion, we need to study the proteins at body temperature, at pH 7, in the midst of molecules that would “dirty” up the systems.

When I first started my PhD program, as I was trying to figure out who my supervisors are and what they have done, I came upon Brian’s article “When biochemistry became chemistry” [13]; I remember having a hard time going through the paper with my at the time extremely limited structural biology knowledge (still rather limited now). I now seem to understand how strong of a pull Brian as well as other scientists at the time must had felt, seeing biological molecules become chemicals when written in their structural formulas and that their intrinsic properties can be characterized using physical approaches. After decades of brilliant and diligent work contributed to solving protein structure better and better; maybe it is time to switch the focus to studying the properties of proteins in conditions closer and closer to physiological conditions. Maybe the time has come “when chemistry needs to become biology”.

SECTION II: Brain metabolomics study

By the time I joined Brian’s lab, there were fascinating ongoing metabolomics studies on mice urine. One of the research topics was to distinguish the animals infected with mice prion RML from their healthy controls; and there were already some promising results. I remember learning from Dr. John Paul Glaves to use the tube-in-tube method [14] where a small capillary full of mice urine is inserted into a 5mm NMR tube and held to the centre by a little piece of rubber. I still remember thinking it was kind of funny to watch a big hockey guy maneuvering those tiny little parts (I have a weird sense of humor).
I do not seem to remember who made the call of switching the subject to mice brains; however, one of the first few things I did in my graduate program was to work out a protocol to extract the metabolites out from the brains and I do remember I was quite scared the whole time during that experience. From choosing a way to homogenize the brains and a buffer to work with to finding out whether I should lyophilize or speed dry the homogenate as well as how to remove the big particles in the sample for NMR experiments, I had too much to worry about each step of the way. In the end, after spending much more time than what was necessary, I was able to establish a protocol that seemed to have worked fine with the setup we have in the prion centre containment area.

Nonetheless, I could not have foreseen that the biggest mistake I would make for this project has nothing to do with the extraction protocol. For the first half of the mice brains, upon receiving the samples I homogenized all of them then stored the brain homogenate in -80°C for months before proceeding with the following steps in the extraction; yet for the second half of the samples I processed samples right after they were homogenized. In the later analysis, there was a clear separation between the two groups of samples processed with different timing and we would not be able to proceed with this project for the following time. Luckily, there came more brains inoculated the same way as the previous batch of brains I worked with for other experimental purposes, and I was able to get some homogenate to repeat this experiment. This time I made sure everything was processed the exact same way with the exact same timing and later analysis revealed noteworthy results, which are covered in Chapter 4 of the thesis. Dr. Pascal Mercier helped profiling all the spectra using a semi-automatic approach, which not only assured the quality of assigning but also shortened my PhD program by a few good months.

The results from Chapter 4 indicate that the brain metabolomic profile displayed continuous changes along the progression of mouse prion disease. While the concentrations of myo-inositol, glutamine have increased, the concentrations of taurine, 4-aminobutyrate, acetate and aspartate have decreased during the process. Importantly,
these findings are backed up by the evidence that systematic age-dependent changes in neurochemical profiles are not readily discernible; therefore, the differentiation between the infected and non-infected samples are not caused by the aging factor. And to my greatest relief, the extraction protocol I came up with turned out to have a better extraction efficiency when compared with the conventional methods.

The best experiment to verify these results would be to carry out longitudinal studies using MRS on live animals. Since the identified compounds are detectable using MRS, by running MRS on prion inoculated animals as well as animals of other neurodegenerative disease models, it would be exciting to see how these metabolites change these diseases progress.

The painful detour I walked in this project made me fully realize how important it is to have a standardized system for sample collection, storage as well as processing. Not only the steps should be carried out carefully, but the waiting time in between the steps must also be kept as close as possible. This experience reflects that perfect standardization must be the premise of generating any meaningful results in metabolomics studies, also one can have a good glimpse of how challenging it can be to integrate metabolomics into health care system for clinical usage.
REFERENCES


Chapter 5 – Results and discussion


BIBLIOGRAPHY


[92] D. G. Donne *et al.*, "Structure of the recombinant full-length hamster prion protein PrP(29-231): the N terminus is highly flexible," (in eng), *Proc Natl*
Bibliography


[118] J. Glaves, M. Li, P. Mercier, R. Fahlman, and B. Sykes, "High-throughput, multi-platform metabolomics on very small volumes: 1 H NMR metabolite


PPATHOGENS-D-12-00720 [pii].


[291] M. Stevens, S. Lattimer, M. Kamiyo, C. Van Huysen, A. Sima, and D. Greene, "Osmotically-induced nerve taurine depletion and the compatible osmolyte


[326] E. D. Walter, M. Chattopadhyay, and G. L. Millhauser, "The affinity of copper binding to the prion protein octarepeat domain: evidence for negative


**Fig. A-1.** The cellular location of HRdup is comparable to that of M128V PrP

A) RK13 cells were transiently transfected with pcDNA3 vectors encoding WT, M128V and HRdup PrP. Lysates without PNGase F treatment run in Tris-glycine gels showed a similar degree of glycosylation. Lysates after PNGase F treatment showed similar amounts of full-length (FL) PrP, as well as similar amounts of C1 endoproteolytic cleavage products. B) RK13 cells were transiently transfected with WT, M128V and HRdup PrP using a pcDNA3 vector. The biotinylation reagent is impermeable to the cell membrane. Total PrP (T) and biotinylated PrP (S, cell-surface) are loaded side-by-side and electrophoresed using a Tris-glycine buffer. HRdup is comparably accessible to the biotinylation reagent as control PrP. C) Transfected cells as per panel A were analyzed for cell-surface PrP without permeabilization (top row) or after permeabilized with Triton X-100 ("+TX"; bottom row) to analyze intracellular PrP. No qualitative difference in the localization of the three proteins was observed (n=6).
All images represent a merge of Hoechst (nuclei, blue) and Alexa fluor 594 (red) signals. Conjugated secondary antibody was used with SAF83 primary antibody (both 1:2000). 40x magnification.

Tg.HRdup-10

PrP
H&E
GFAP

Hippocampus

Thalamus

Cortex

Cerebellum

Tg.HRdup-32

Cortex

Cerebellum

Hippocampus

H&E

GFAP
**Fig. A-2.** Histopathological examination of Tg.HRdup-10 and Tg.HRdup-32 animals.

Upper set of panels. Photomicrographs of sagittal sections of a sick Tg.HRdup-10 mouse (585d). Immunostaining for PrP was performed after treatment of the slices with formic acid, slices destined for examination with H&E or GFAP were left untreated. The hippocampus (A-C), thalamus (D-F), cortex (G-I) and cerebellum (J-L) are shown. PrPimm immunostaining is absent the hippocampus, thalamus and cortex but is present in the cerebellum. Vacuolation is found in the hippocampus, thalamus and cerebellum. GFAP immunostaining is restricted to the cerebellum. Scale bar = 100 μm. Lower set of panels, photomicrographs of sagittal sections of a sick Tg.HRdup-32 mouse (162d) showing spongiform change. M, P cortex, N and Q cerebellum and O and R, hippocampus. H & E (M-O) and GFAP immunostains (P-R) are shown. Scale bar = 100 μm for cortex and cerebellum, 500μM for hippocampus. Insets (squares, O, R) show expanded views of spongy change but mild astrocytic activation.
Fig. A-3. Histopathological examination of Tg.M128V-39 animals. Photomicrographs of sagittal sections of healthy, aged Tg.M128V-39 mouse (576d). Immunostaining for PrP was performed after treatment of the slices with formic acid, slices destined for examination with H&E or GFAP were left untreated. The hippocampus (A-C), thalamus (D-F), cortex (G-I) and cerebellum (J-L) are shown. Vacuolation, PrPres and GFAP immunostaining are absent in all regions in the brains of these mice. Scale bar = 100 μm.

Fig. A-4. Lack of myopathy and neuropathy in Tg.HRdup-26 mice. H&E stain of cross-section of quadriceps from aged WT (211d, A) and a Tg.HRdup-26 mouse (437 d, B). Scale bars 50 μm. Toluidine blue staining of a cross section of sciatic nerve from an aged Prnp^{0/0} (446d, C) showing irregular fibre diameters. A sick Tg.HRdup-26 mouse (437d, D). Scale bar = 25 μm.
Fig. A-5. Brain homogenate from sick Tg.HRdup animals induces early pathological changes in mice expressing homologous PrP.

Low expressing Tg.HRdup-10 animals and tga20 (overexpressing WT PrP ~6-7x) animals received an intracranial inoculation with brain homogenate from a sick Tg.HRdup-26 animal or a healthy WT animal. The onset of clinical disease and pathological change was accelerated in the animals receiving the HRdup brain homogenate. A) Kaplan-Meier death curve for animals in the study. n=4 for all groups. tga20 animals, with the exception of one that had to be euthanized due to intercurrent illness, were euthanized after 460 days. Tg.HRdup-10 animals inoculated with healthy brain homogenate were euthanized after 400 days. Tg.HRdup-10 animals inoculated with clinical Tg.HRdup-26 brain homogenate succumbed to disease at 276±38 days. B) Immunohistochemistry of the cerebellum of a Tg.HRdup-10 (top row) and tga20 animal (middle row) inoculated with the brain homogenate from a
Appendix A – Supporting Information for Chapter 2

sick Tg.HRdup-26 animal or Tg.HRdup-10 animal inoculated material from a healthy WT animal (bottom row). i, iv, vii, PrP\textsuperscript{Sc} immunoreactivity with SAF83; ii, v, viii, H&E stain; iii, vi, ix: GFAP immunoreactivity. Scale bar = 75 \( \mu \text{m} \). Note that only animals that succumbed to illness (and harboring protease resistant PrP) display PrP immunoreactivity, vacuolation and astrocytic gliosis associated with prion disease.

Fig. A-6. 1D NMR spectra for three PrP alleles.

1D 1H NMR full spectra are presented for HRdup, M128V and WT PrP (top to bottom).
Fig. A-7. 1D NMR spectra of urea titrations for three PrP alleles.

1D $^1$H NMR experiment of WT (A), M128V (B) and HRdup PrP (C) titrated with urea from 0 to 10M. Upper panels show stacked plots of the aromatic region of spectra, ranging from 5.2 to 6.7 ppm (resonances of Y162H$_x$, Y161H$_x$, F197H$_x$ and Y217H$_b$ are indicated) while lower panels show stacked plots of the methyl region of spectra, ranging from -0.14 to 0.86 ppm (resonance of H181H$_y$$^2$ is indicated in the graph).
**Fig. A-8.** Complete 2D NMR spectrum.

Full spectra of 2D $^1$H-$^1$H NOESY of HRdup (A) and M128V PrP (B) are presented.
Fig. A-9. Secondary structure evolution in MD simulations.
Shown are data from the M128V-I system (A), M128V-II system (B), HRdup-I system (C), and HRdup-II system (D). The total simulation time was 20 ns. Along the vertical axes, N-terminals are on the top and C-terminals are on the bottom. The legends underneath indicate the structure-color associations: T – turn (cyan), E – extended β conformation (yellow), B – isolated β bridge (dark yellow), H – α-helix (purple), G – 3-10-helix (blue), I – π-helix (red), and C – coil (white). The figures were obtained with the VMD package.
Fig. A-10. ECD correlation maps.

These are calculated for side-chain atoms (above the diagonal) and main-chain atoms (below the diagonal) in M128V-I system (A), HRdup-I system (B), M128V-II system (C), and HRdup-II system (D). High levels of correlations are indicated with red and yellow, medium levels are indicated with green and light-blue; and low levels with dark blue. Some of the dark blue lines in the side chain correlation maps correspond to glycine (no side chains) and alanine (short side chains) residues, which were excluded from the side-chain analysis.
Fig. A-11. Docking analysis.
Close-up of predicted homodimers M128V-I (A) and HRdup-I (B), and heterodimer M128V-I HRdup-I (C). The residues on the binding interface are shown with lines and translucent surfaces colored according to the electrostatic charge (translucent pink – negative charge; translucent blue – positive charge; white – neutral). In the display of the first chain in the dimers, α-helices are indicated with opaque red color; β-strands with light blue; random coils with grey; and turns or helix disruptions with green.
Fig. A-12. A 16 kDa thermolysin-resistant PrP fragment is present in neonatal Tg.HRdup-26 mice. Upper panel, western blot analysis of brains of eight neonatal animals (A-H) from the Tg.HRdup-26 line. Antibody is Sha31. The four left-hand samples have the thermolysin-resistant fragment. Lower panel, corresponding PCR analyses of genomic DNA from eight neonatal mice with reaction products displayed on SYBR safe stained agarose gel. The diagnostic fragment for the PrP transgene migrates at 550 bp.
<table>
<thead>
<tr>
<th>Brain region</th>
<th>Spongiosis</th>
<th>PrP&lt;sup&gt;B&lt;/sup&gt;</th>
<th>Tau</th>
<th>Amyloid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortex</td>
<td>Line 10: 3 Mi, 1 Mo, 3 Neg; Line 26: 10 Mi, 5 Mo, 3 St, 2 Neg; Line 32: 6 Mi, 4 Mo)</td>
<td>Focal deposit (0/7 line 10; 0/1 line 13; 8/20 line 26; 0/7 line 32)</td>
<td>None (0/1 line 13; 0/4 line 26; 0/1 line 32)</td>
<td>None (0/1 line 13; 0/3 line 26; 0/3 line 32)</td>
</tr>
<tr>
<td>White matter</td>
<td>Line 10: 4 Mi, 1 Mo, 2 Neg; Line 26: 6 Mi, 6 Mo, 2 St, 6 Neg; Line 32: 3 Mi, 4 Mo)</td>
<td>Focal deposit (1/7 line 10; 0/1 line 13; 9/20 line 26; 0/7 line 32)</td>
<td>None (0/1 line 13; 0/4 line 26; 0/1 line 32)</td>
<td>None (0/1 line 13; 0/3 line 26; 0/3 line 32)</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>Line 10: 1 Mi, 3 Mo, 1 St, 2 Neg; Line 26: 4 Mi, 6 Mo, 6 St, 4 Neg; Line 32: 6 Mi, 1 Mo)</td>
<td>Focal deposit (3/7 line 10; 0/1 line 13; 9/20 line 26; 0/7 line 32)</td>
<td>None (0/1 line 13; 0/4 line 26; 0/1 line 32)</td>
<td>None (0/1 line 13; 0/3 line 26; 0/3 line 32)</td>
</tr>
<tr>
<td>Striatum</td>
<td>Line 10: 1 Mi, 5 Mo, 1 Neg; Line 26: 6 Mi, 4 Mo, 4 St, 6 Neg; Line 32: 4 Mi, 2 Mo, 1 Neg)</td>
<td>Focal deposit (0/7 line 10; 0/1 line 13; 8/20 line 26; 0/7 line 32)</td>
<td>None (0/1 line 13; 0/4 line 26; 0/1 line 32)</td>
<td>None (0/1 line 13; 0/3 line 26; 0/3 line 32)</td>
</tr>
<tr>
<td>Thalamus</td>
<td>Line 10: 1 Mi, 5 Mo, 1 Neg; Line 26: 6 Mi, 9 Mo, 4 St, 1 Neg; Line 32: 2 Mi, 5 Mo)</td>
<td>Focal deposit (0/7 line 10; 0/1 line 13; 8/20 line 26; 0/7 line 32)</td>
<td>None (0/1 line 13; 0/4 line 26; 0/1 line 32)</td>
<td>None (0/1 line 13; 0/3 line 26; 0/3 line 32)</td>
</tr>
<tr>
<td>Brainstem</td>
<td>Line 10: 3 Mi, 1 Mo, 3 Neg; Line 26: 8 Mi, 6 Mo, 0 St, 6 Neg; Line 32: 1 Mi, 6 Mo)</td>
<td>Focal deposit (3/7 line 10; 0/1 line 13; 7/20 line 26; 0/7 line 32)</td>
<td>None (0/1 line 13; 0/4 line 26; 0/1 line 32)</td>
<td>None (0/1 line 13; 0/3 line 26; 0/3 line 32)</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>Line 10: 1 Mi, 1 Mo, 3 St, 2 Neg; Line 26: 10 Mi, 5 Mo, 4 St, 1 Neg; Line 32: 1 Mi, 4 Mo, 2 St</td>
<td>Focal deposit (3/7 line 10; 0/1 line 13; 8/20 line 26; 0/7 line 32)</td>
<td>None (0/1 line 13; 0/4 line 26; 0/1 line 32)</td>
<td>None (0/1 line 13; 0/3 line 26; 0/3 line 32)</td>
</tr>
</tbody>
</table>

**Table A-1.** Pathology in mouse Tg.HRdup lines

Line 32 age range examined is 149-237d, Line 26 age range 111-482d, line 10 age range 308-636d

Mi = Mild, Mo = moderate, St = strong and neg = negative
<table>
<thead>
<tr>
<th>m-Value (kJ mol(^{-1}) M(^{-1}))</th>
<th>WT</th>
<th>M128V</th>
<th>HRdup</th>
</tr>
</thead>
<tbody>
<tr>
<td>I181H(\gamma)2</td>
<td>3.2±0.5</td>
<td>3.8±0.6</td>
<td>6.8±2.1</td>
</tr>
<tr>
<td>Y162H(\varepsilon)</td>
<td>2.9±0.5</td>
<td>3.9±0.4</td>
<td>5.1±0.9</td>
</tr>
<tr>
<td>Y161H(\alpha)</td>
<td>4.1±0.9</td>
<td>6.5±2.6</td>
<td>5.5±0.6</td>
</tr>
<tr>
<td>F197H(\alpha)</td>
<td>7.0±1.3</td>
<td>5.0±0.5</td>
<td>4.1±1.0</td>
</tr>
<tr>
<td>Y217H(\alpha)</td>
<td>2.9±0.4</td>
<td>3.6±0.3</td>
<td>4.6±1.6</td>
</tr>
</tbody>
</table>

**Table A-2.** \(m\) values Table for 5 PrP residues in a urea denaturation experiment
### Table A-3.
The results of SWISS-MODEL template library search with BLAST and HHBlits

The search was for evolutionary related structures matching the target HRdup construct with the M128V substitution and LGGLGGYV insert sequence. For GMQE and QMEAN4 scores description, see Methods above.

<table>
<thead>
<tr>
<th>Template (PDB ID. model#. chain ID)</th>
<th>Sequence Identity, %</th>
<th>Oligo-state</th>
<th>Found by</th>
<th>Method</th>
<th>Resolution</th>
<th>Sequence Similarity</th>
<th>Range (WT)</th>
<th>Coverage</th>
<th>Description</th>
<th>GMQE</th>
<th>QMEAN4</th>
</tr>
</thead>
<tbody>
<tr>
<td>4kml.1.A</td>
<td>90.34</td>
<td>hetero-oligomer</td>
<td>HHblits</td>
<td>X-ray</td>
<td>1.50Å</td>
<td>0.62</td>
<td>116 - 224</td>
<td>0.95</td>
<td>Major prion protein</td>
<td>0.50</td>
<td>-0.97</td>
</tr>
<tr>
<td>1qlz.1.A</td>
<td>90.34</td>
<td>monomer</td>
<td>HHblits</td>
<td>NMR</td>
<td>NA</td>
<td>0.62</td>
<td>114 - 227</td>
<td>0.95</td>
<td>PRION PROTEIN</td>
<td>0.49</td>
<td>-1.32</td>
</tr>
<tr>
<td>4ma7.1.A</td>
<td>99.12</td>
<td>hetero-oligomer</td>
<td>HHblits</td>
<td>X-ray</td>
<td>1.97Å</td>
<td>0.63</td>
<td>118 - 227</td>
<td>0.53</td>
<td>Major prion protein</td>
<td>0.46</td>
<td>-1.56</td>
</tr>
<tr>
<td>4ma8.1.A</td>
<td>99.12</td>
<td>hetero-oligomer</td>
<td>HHblits</td>
<td>X-ray</td>
<td>2.20Å</td>
<td>0.63</td>
<td>118 - 225</td>
<td>0.53</td>
<td>Major prion protein</td>
<td>0.45</td>
<td>-1.49</td>
</tr>
<tr>
<td>2l1h.1.A</td>
<td>99.11</td>
<td>monomer</td>
<td>HHblits</td>
<td>NMR</td>
<td>NA</td>
<td>0.63</td>
<td>123 - 226</td>
<td>0.52</td>
<td>Major prion protein</td>
<td>0.45</td>
<td>-1.94</td>
</tr>
<tr>
<td>1b10.1.A</td>
<td>93.62</td>
<td>monomer</td>
<td>HHblits</td>
<td>NMR</td>
<td>NA</td>
<td>0.62</td>
<td>124 - 227</td>
<td>0.65</td>
<td>PRION PROTEIN</td>
<td>0.43</td>
<td>-3.25</td>
</tr>
<tr>
<td>2ku5.1.A</td>
<td>98.21</td>
<td>monomer</td>
<td>HHblits</td>
<td>NMR</td>
<td>NA</td>
<td>0.62</td>
<td>120 - 231</td>
<td>0.52</td>
<td>Major prion protein</td>
<td>0.43</td>
<td>-3.13</td>
</tr>
<tr>
<td>2139.1.A</td>
<td>98.25</td>
<td>monomer</td>
<td>HHblits</td>
<td>NMR</td>
<td>NA</td>
<td>0.63</td>
<td>118 - 231</td>
<td>0.53</td>
<td>Major prion protein</td>
<td>0.41</td>
<td>-4.03</td>
</tr>
<tr>
<td>Residue</td>
<td>M128V-I</td>
<td>M128V-II</td>
<td>HRdup-I</td>
<td>HRdup-II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<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>Q90</td>
<td>6.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G91</td>
<td>12.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G92</td>
<td>11.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N96</td>
<td>11.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q97</td>
<td>17.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K105</td>
<td></td>
<td></td>
<td>10.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H110</td>
<td></td>
<td>28.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y127</td>
<td></td>
<td>38.0%</td>
<td>4.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V128</td>
<td>99.6%</td>
<td>100.0%</td>
<td>18.5%</td>
<td>20.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ins-G3</td>
<td></td>
<td></td>
<td>18.4%</td>
<td>20.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ins-L4</td>
<td></td>
<td>12.2%</td>
<td>4.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ins-G5</td>
<td></td>
<td>6.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ins-V8</td>
<td></td>
<td>79.1%</td>
<td>100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L129</td>
<td>92.2%</td>
<td>92.8%</td>
<td>79.1%</td>
<td>100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G130</td>
<td>67.0%</td>
<td>78.0%</td>
<td>78.7%</td>
<td>97.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S131</td>
<td>19.3%</td>
<td>68.6%</td>
<td>27.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A132</td>
<td>48.9%</td>
<td></td>
<td></td>
<td>22.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M133</td>
<td>34.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N158</td>
<td>50.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q159</td>
<td>19.3%</td>
<td>68.6%</td>
<td></td>
<td>22.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V160</td>
<td>67.0%</td>
<td>78.0%</td>
<td>78.6%</td>
<td>97.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y161</td>
<td>92.2%</td>
<td>92.8%</td>
<td>78.9%</td>
<td>100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y162</td>
<td>99.6%</td>
<td>100.0%</td>
<td>78.9%</td>
<td>100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D166</td>
<td>3.9%</td>
<td></td>
<td>4.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N170</td>
<td>3.8%</td>
<td></td>
<td>4.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T198</td>
<td></td>
<td></td>
<td></td>
<td>10.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A-4. Time-averaged β-content occupancies in four MD trajectories
**Table A-5.** Per residue solvent accessible surface area (SASA) Per residue solvent accessible surface area (SASA) was calculated from the last 2 ns of the four 20 ns trajectories. Difference in per-residue SASA for models M128V and HRDdup were averaged over two trajectories each; total per-residue SASA; and difference in weighted average per-residue SASA for models M128V and HRDdup. In the first two columns, hydrophobic residues are highlighted in blue, and charged residues are highlighted in red. In the last column, residues that exhibit more than 25% differences in SASA are highlighted in dark red.

| Residue | SASA, nm² | Difference in average SASA(\(\Delta\)), nm² | Maximal SASA(\(\bar{\text{SASA}}\)), nm² | Difference in weighted average SASA, %(|\(\Delta\)|) |
|---------|-----------|---------------------------------|-----------------|-----------------|
| G89     | 0.921     | 0.989, 0.878, 0.806, -0.113 | 0.813, -13.89   |
| Q90     | 1.405     | 0.993, 0.916, 0.951, -0.265 | 1.812, -14.63   |
| G91     | 0.282     | 0.330, 0.569, 0.626, 0.292 | 0.813, 35.92    |
| G92     | 0.554     | 0.249, 0.785, 0.534, 0.258 | 0.813, 31.69    |
| G93     | 0.602     | 0.603, 0.738, 0.745, 0.139 | 0.813, 17.09    |
| T94     | 1.387     | 0.555, 0.699, 1.002, -0.120 | 1.401, -8.59    |
| H95     | 1.282     | 1.752, 1.224, 1.330, -0.240 | 1.845, -13.02   |
| N96     | 0.850     | 0.856, 1.073, 0.641, 0.004 | 1.566, 0.23     |
| Q97     | 1.136     | 0.479, 1.127, 1.232, 0.372 | 1.819, 20.45    |
| W98     | 1.318     | 1.074, 1.321, 2.012, 0.470 | 2.538, 18.53    |
| N99     | 1.172     | 0.946, 1.255, 0.743, -0.060 | 1.566, -3.82    |
| K100    | 1.416     | 1.489, 1.161, 1.578, -0.083 | 2.049, -4.03    |
| P101    | 0.496     | 0.901, 0.747, 0.769, 0.060 | 1.364, 4.37     |
| S102    | 0.843     | 1.079, 0.920, 0.672, -0.165 | 1.159, -14.23   |
| K103    | 1.280     | 1.566, 1.469, 1.521, 0.072 | 2.049, 3.51     |
| P104    | 0.784     | 1.172, 0.675, 0.540, -0.371 | 1.364, -27.18   |
| K105    | 1.729     | 1.730, 1.409, 1.418, -0.316 | 2.049, -15.40   |
| T106    | 1.152     | 0.859, 0.791, 0.292, -0.464 | 1.401, -33.12   |
| N107    | 0.992     | 1.270, 0.523, 1.053, -0.343 | 1.566, -21.87   |
| L108    | 1.659     | 0.798, 0.609, 0.790, -0.529 | 1.821, -29.07   |
| K109    | 1.378     | 1.294, 1.074, 0.889, -0.354 | 2.049, -17.29   |
| H110    | 0.993     | 1.465, 0.962, 1.334, -0.081 | 1.845, -4.37    |
| V111    | 0.572     | 0.369, 0.975, 0.863, 0.448 | 1.527, 29.36    |
| A112    | 0.342     | 0.727, 0.470, 0.549, -0.025 | 1.085, -2.29    |
| G113    | 0.455     | 0.611, 0.758, 0.514, 0.103 | 0.813, 12.64    |
| A114    | 0.554     | 0.373, 0.625, 0.443, 0.071 | 1.085, 6.51     |
| A115    | 0.319     | 0.775, 0.961, 0.949, 0.408 | 1.085, 37.58    |
### Table A-6. Total SASA for two trajectories of M128V and HRdup PrP alleles

Data are calculated from the last 2 ns of the four MD trajectories, and per-residue SASA for models M128V and HRdup in average for two trajectories each.
<table>
<thead>
<tr>
<th>PrP allele</th>
<th>S1-S2</th>
<th>S2-H2</th>
<th>H2-H3</th>
<th>S2-H3</th>
<th>N termini-insert</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G130.Q159</td>
<td>V160.C178</td>
<td>V179.C213</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y127.R163</td>
<td>Y161.I181</td>
<td>V175.E210</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>V128.Y162</td>
<td>C178.V209</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y127.Y161</td>
<td>V175.V214</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S131.Q159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.Y162</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>V128.R163</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.Q159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.V160</td>
<td>V160.V179</td>
<td>V175.V214</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S131.Q159</td>
<td>T182.V209</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.Y162</td>
<td>C178.E210</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G130.Q159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G130.Y161</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.Q159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S131.N158</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>V128.R163</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>insV8.Y161</td>
<td>R163.C178</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>insY7.Y161</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>insV8.R163</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.Y162</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G130.Q159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G130.Y161</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.Q159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S131.Q159</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>insY7.R163</td>
<td>Y161.I181</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>insY7.Y161</td>
<td>Y162.I181</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.V160</td>
<td>Y162.T182</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G130.V160</td>
<td>Y162.C178</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L129.Y162</td>
<td>V160.I181</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A-7. Residue pairs with robust inter-correlations for Cα atoms. This indicates PrP regions of the strongest dynamic correlations

<table>
<thead>
<tr>
<th>M128V-I homodimer</th>
<th>M128V-II homodimer</th>
<th>HRdup-I homodimer</th>
<th>HRdup-II homodimer</th>
</tr>
</thead>
<tbody>
<tr>
<td>L124 Y127 H176 N180 Y168</td>
<td>Y127 V128 P164 S131 M133</td>
<td>W98 K100 W144 Y148 Y148</td>
<td>R135 K100 Q159 Q159</td>
</tr>
<tr>
<td>T200 G126 G125 K100 W144</td>
<td>M133 V165 A132 N158 P164</td>
<td>Y127 W98 G126 K184</td>
<td>R135 Y127 M133 Y154</td>
</tr>
</tbody>
</table>

Table A-8. Residues in M128V and HRdup monomers predicted to be involved in dimer contacts
Data is based on the consensus of top 10 models from three methods (InterEVScore, SOAP_PP, and FRODOCK) with the InterEVDock. Residues highlighted in red are only involved in HRdup contacts.

SUPPLEMENTARY METHODS

Expression in mammalian cells
pcDNA3.moPrP.wt was subjected to site directed mutagenesis to create pcDNA3.moPrP128V and pcDNA3.moPrPHRdup128V using the Agilent QuikChange II XL kit. 128V primers:

SW132 5’-GGGGGCGCTTGGTGCTACGTGCTGGGAGCAGCGCATGACG-3’,
SW133 5’-GCTCATGGCGCTCCCCAGCACGTAGCCACCAAGGCCCCCC-3’.
The sequences were then verified using the universal T7 and SP6 primers before excision and religation into pcDNA3 that had not been subjected to site directed mutagenesis. Rabbit kidney epithelial (RK13) cells were maintained in Dulbecco’s Modified Eagles Medium (DMEM; Invitrogen) that had been supplemented with 10% fetal bovine serum (Invitrogen) and 1% penicillin/streptomycin (Invitrogen) at 37°C and 5% CO2. Cells were transfected at a confluence of 80-90% using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions.

**Immunocytochemistry**

RK13 cells were transfected with the indicated pcDNA3 PrP vector. 24 h after transfection, cells were re-plated on glass coverslips and given 24 h to recover. After rinsing twice with PBS and fixation with 4% paraformaldehyde, cells were washed 3 times with PBS before permeabilization with 0.2% Triton X-100 (Sigma) or were left un-permeabilized. Cells were rinsed 3 times with PBS and incubated in Sha31 (1:5000) overnight at 4 °C with rocking. Following 3 PBS washes, cells were blocked with 2% goat serum (Invitrogen) and incubated with goat α-mouse Alexa Fluor 488 (Invitrogen, 1:300) for 2 h at room temperature. Nuclei were stained with 1 μg/mL Hoechst and visualized using a Eclipse 90I microscope (Nikon) and a CFI PL 40X/ N.A. 0.75 lens (Nikon) using the following excitation/emission filter properties: 325 – 375/500 – 575 nm with a 495 nm long-pass filter (blue channel) and 440-510/475-575 nm with a 495 nm long-pass filter (green channel). Images were acquired at room temperature with a Retiga 2000R mono cooled camera, fast 1394 using NIS-Elements AR advanced research software.

**Cell Surface Biotinylation Assay**

The Pierce cell surface protein isolation kit was scaled down to a 6-well format and performed according to manufacturer's instruction.
Fig. B-1. 2D $^1$H, $^{15}$N-HSQC NMR spectrum of WT mPrP.
Fig. B-2. 2D $^1$H, $^{15}$N-HSQC NMR spectrum of V128 mPrP.
Fig. B-3. 2D $^1$H, $^{15}$N-HSQC NMR spectrum of HRdup mPrP.
Fig. B-4. A strip-plot of 3D NMR experiment. This diagram shows $^1$H-$^1$HN slices from an overlay of 3D CACBCONH and 3D NHCACB NMR spectra collected at 600 MHz ($^1$H) and 30°C. The resonances labeled in red are from Cα and Cβ (i-1) in CBCACONH spectrum; the blue and orange resonances are from Cα and Cβ (i, i-1) in NHCACB spectrum, respectively. Each strip corresponds to a residue in the N-terminal repetitive part of the sequence. Sequential residues are connected by solid lines, and the assignments are shown at the top of the strips. Hydrogen chemical shifts not shown. The assignments were deposited in BMRB, ID 27835.
Fig. B-5. Diagrams of the backbone structure of the proposed type 2 β-turn. 

- **a)** The \( d_{\text{oii}(i, i+1)} \) NOEs that can be observed in 3D NHSQC-NOESY NMR spectrum are indicated.
- **b)** Regions of 3D NHSQC-NOESY NMR spectrum on the nitrogen planes at chemical shift of 125.56 ppm and 110.59 ppm, which correspond to the NH of residues LX3 (left) and GX4 (right), respectively. The same vertical scale was used to plot these two spectra.
**Fig. B-6.** Two ways the β1 strand can be formed. VX2 and LX3 with high β-propensity in TALOS are colored in red.
**Fig. B-7.** 1D $^1$H-NMR spectra of V128 at 500, 600 and 700MHz. One of the two resonances of G130 is indicated.
Fig. C-1. The downfield region of the overlay of 30 high resolution 500 MHz 1D $^1$H-NMR spectra. This also demonstrates the high sufficiency of the novel extraction method in removing the fat and lipid molecules.
Fig. C-2. The explained variance by each component in the PCA analysis. The green bars are the accumulated explained variance. For each individual component, the $R^2$ value is 0.412, 0.228, 0.083, 0.059, 0.034.
Fig. C-3. Spectra of outliers compared with other samples. A) The upper field of the spectra of sample 82806 aligned with 82805, a sample from the same group, processed on the same day in the same batch.
B) The down field of the spectra of sample 82806 aligned with 82805, a sample from the same group, processed on the same day in the same batch.
C) The upper field of the spectra of sample 82864 aligned with 82863, a sample from the same group, processed on the same day in the same batch.
D). The down field of the spectra of sample 82864 aligned with 82863, a sample from the same group, processed on the same day in the same batch.
Fig. C-4. More concentrations of metabolites at different timepoints. The concentrations of the metabolites of NBH at 30, 60, 90, 120, 150 dpi timepoints (NBH1 to NBH5, solid circles) plotted next to their concentrations of RML at corresponding timepoints (RML1-RML5, solid boxes). Each dot represents one sample. The outlier samples 82806 from RML4 and 82864 from RML5 are specially indicated with triangles. The big variation is caused by the lack of enzymatic inhibitor during sample processing.
<table>
<thead>
<tr>
<th>Batch #</th>
<th>Timepoints (days)</th>
<th>NBH</th>
<th>RML</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>6</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
<td>7</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>120</td>
<td>8</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>150</td>
<td>7</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>Sum</td>
<td>-</td>
<td>34</td>
<td>39</td>
<td>73</td>
</tr>
</tbody>
</table>

**Table C-1.** Experimental design of the number of brains harvested at each timepoint.
<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Metabolite</th>
<th>Metabolite</th>
<th>Metabolite</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-aminobutyrate</td>
<td>Formate</td>
<td>Leucine</td>
<td>Putrescine</td>
</tr>
<tr>
<td>Acetate</td>
<td>Fumarate</td>
<td>Lysine</td>
<td>Pyroglutamate</td>
</tr>
<tr>
<td>Alanine</td>
<td>Glucose</td>
<td>Malate</td>
<td>Serine</td>
</tr>
<tr>
<td>Arginine</td>
<td>Glutamate</td>
<td>Mannose</td>
<td>Taurine</td>
</tr>
<tr>
<td>Ascorbate</td>
<td>Glutamine</td>
<td>Methanol</td>
<td>Threonine</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Glycerophosphocholine</td>
<td>Methionine</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>Aspartate</td>
<td>Glycine</td>
<td>Myo-Inositol</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>Carnosine</td>
<td>Histidine</td>
<td>N-Acetylaspartate</td>
<td>Uracil</td>
</tr>
<tr>
<td>Choline</td>
<td>Hypoxanthine</td>
<td>Nicotinurate</td>
<td>Uridine</td>
</tr>
<tr>
<td>Creatine</td>
<td>Inosine</td>
<td>O-Phosphocholine</td>
<td>Valine</td>
</tr>
<tr>
<td>Cytidine</td>
<td>Isocitrate</td>
<td>O-Phosphoethanolamine</td>
<td>Xanthine</td>
</tr>
<tr>
<td>Dimethylamine</td>
<td>Isoleucine</td>
<td>Phenylalanine</td>
<td></td>
</tr>
<tr>
<td>Ethanolamine</td>
<td>Lactate</td>
<td>Propylene glycol</td>
<td></td>
</tr>
</tbody>
</table>

Table C-2. In total of 50 metabolites are identified on the spectra of all the samples. For PCA, creatine was used for normalization and lactate was removed due to artificial error.
Table C-3. Two-tail t-test assuming unequal variance of the metabolite concentrations between RML samples and NBH samples at five different timepoints.

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>30 d.p.i.</th>
<th>60 d.p.i.</th>
<th>90 d.p.i.</th>
<th>120 d.p.i.</th>
<th>150 d.p.i.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GABA</td>
<td>0.145</td>
<td>0.210</td>
<td>5.05×10⁻⁴</td>
<td>3.00×10⁻⁵</td>
<td>2.24×10⁻⁸</td>
</tr>
<tr>
<td>Acetate</td>
<td>0.313</td>
<td>0.606</td>
<td>2.11×10⁻³</td>
<td>4.75×10⁻⁵</td>
<td>3.81×10⁻⁶</td>
</tr>
<tr>
<td>Aspartate</td>
<td>0.946</td>
<td>0.839</td>
<td>7.15×10⁻⁴</td>
<td>1.63×10⁻⁵</td>
<td>1.05×10⁻⁸</td>
</tr>
<tr>
<td>Taurine</td>
<td>0.677</td>
<td>0.759</td>
<td>2.79×10⁻²</td>
<td>9.73×10⁻⁴</td>
<td>1.27×10⁻⁴</td>
</tr>
<tr>
<td>Myo-inositol</td>
<td>0.372</td>
<td>0.880</td>
<td>0.334</td>
<td>2.51×10⁻²</td>
<td>4.24×10⁻⁵</td>
</tr>
<tr>
<td>Glutamine</td>
<td>0.493</td>
<td>0.508</td>
<td>0.807</td>
<td>4.94×10⁻²</td>
<td>1.23×10⁻⁴</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.725</td>
<td>0.198</td>
<td>0.137</td>
<td>2.14×10⁻⁴</td>
<td>4.05×10⁻²</td>
</tr>
<tr>
<td>NAA</td>
<td>0.307</td>
<td>0.431</td>
<td>8.54×10⁻²</td>
<td>0.162</td>
<td>0.625</td>
</tr>
</tbody>
</table>