Chapter 6. Development of a quantitative PCR assay for *pnpA* and *aprV2* mRNA in *D. nodosus*

6.1. Introduction

PnpA has been proposed to be a virulence repressor in *D. nodosus* (Whittle *et al.*, 1999). Using the PNPase assay (Chapter 4), it was shown that *pnpA* activity is slightly higher in benign strains compared to virulent strains. Additionally, *pnpA* knockouts in benign strains had an increase in twitching motility (Chapter 4), which was decreased when the *pnpA* knockouts were complemented to form a full length *pnpA* (Chapter 5). These results support the hypothesis that PnpA is a virulence suppressor.

It has been proposed that PnpA activity in *D. nodosus* is controlled by the integration of genetic elements into the *tRNA* gene immediately downstream from *pnpA* (Whittle *et al.*, 1999). Transcription of *pnpA* extends into the adjacent integrated element, so the 3' ends of *pnpA* transcripts differ in different strains according to which genetic element is adjacent to *pnpA*. This could alter the stability of the mRNA, or the ability of the mRNA to be translated. This could be investigated by measuring *pnpA* mRNA levels. However, previous attempts to do this using northern blotting (Whittle, 1999) were unsuccessful due to the low level of *pnpA* transcripts.

The thermostability of extracellular proteases from virulent stains is greater than that of benign strains. In the virulent strain A198, there are three protease genes, *aprV2*, *aprV5*

and *bprV*. AprV2 is the major thermostable protease, while AprV5 is the major thermolabile protease (Kennan, R. M., and Rood, I. J., personal communication). The benign strains have three highly-related genes, *aprB2*, *aprB5* and *bprB* (Lilley *et al.*, 1995; Riffkin *et al.*, 1993; 1995).

The difference in thermostability of secreted proteases could be due to differential expression of these three genes in virulent and benign strains. It has been proposed that PnpA and GlpA act as post-transcriptional regulators of virulence determinants in *D. nodosus*. However, the PnpA knockout studies in benign strains did not support a role for PNPase in controlling protease thermostability. GlpA may bind to protease mRNA and prevent translation and/or promote degradation. This could be investigated by comparing protease mRNA levels in virulent and benign strains.

6.1.1. Reverse-transcriptase quantitative PCR (RT-qPCR)

RT-qPCR is a technique for quantifying mRNA levels. The benefits of this procedure over standard methods for measuring mRNA levels include its sensitivity and large dynamic range (Dorak, 2006). In this chapter, RT-qPCR assays for measuring 16S rRNA (16S ribosomal RNA), *pnpA* and *aprV2* were developed and used to quantify *pnpA* and *aprV2* mRNA levels in benign and virulent strains and in the *pnpA* knockout mutants. 16S rRNA was used as the internal control to normalise the *pnpA* and *aprV2* mRNA levels.

6.2. Results

6.2.1. Construction of an internal standard (pSKrR1 - 16S rRNA plasmid) for the standard curve

Generally two strategies are employed to quantify the results obtained from real-time PCR; the standard curve method and the comparative threshold method. The standard curve method has been applied in these experiments (Dorak, 2006).

The reliability or repeatability of a real time RT-PCR analysis generally requires an internal control to normalise assay results between samples and to allow for errors in RNA quantification. This method assumes that the expression of housekeeping genes does not vary between different strains under the same conditions. Therefore 16S rRNA was chosen as an internal control. Dewhirst *et al.* (1990) sequenced and characterised 16S rRNA of *D. nodosus* and revised the phylogeny. La Fontaine and Rood (1996) studied 16S rRNA in *D. nodosus* strain A198 and found three copies of the rRNA locus, each `````comprising a 16S, 23S and 5S rRNA with a single rrn promoter present upstream of the 16S rRNA gene.

A set of primers (BCH316' and BCH317', Table A2) were designed based on the published 16S rRNA sequences (GenBank accession nos M35016 and DQ016291) to amplify 507 bp from the 5' end of the 16S rRNA of *D. nodosus* strain UNE64. The PCR product was cloned into a T-A cloning vector system using the pGEM^{T} easy vector system (Promega) to construct the plasmid pSKrR1. The identity of the insert was

confirmed by DNA sequencing (Fig. 6.1). There were no differences between the sequence of the cloned insert and the *D. nodosus* 16S rRNA sequence (GenBank accession no M35016), so no errors were introduced by Taq polymerase.

```
M35016   47 TTAACACATGCAAGTCGAACGGGGTTATGTAGCTTGCTATGTAACCTAGTGGCGGACGGG
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pSKrR1   89 TTAACACATGCAAGTCGAACGGGGTTATGTAGCTTGCTATGTAACCTAGTGGCGGACGGG

M35016  107 TGAGTAATATATAGGAATCTGCCTTATGGTGGGGGACAACGTATGGAAACGTACGCTAAT
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pSKrR1  149 TGAGTAATATATAGGAATCTGCCTTATGGTGGGGGACAACGTATGGAAACGTACGCTAAT

M35016  167 ACCGCATAAGATTGAAGAATGAAAGCGGGGGCTCGAAAGACCTCGCGCCGTAAGATGAGC
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pSKrR1  209 ACCGCATAAGATTGAAGAATGAAAGCGGGGGCTCGAAAGACCTCGCGCCGTAAGATGAGC

M35016  227 CTATATCGGATTAGCTAGTTGGTGGGGTAAGAGCCTACCAAGGCGACGATCCGTAGCTGG
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pSKrR1  269 CTATATCGGATTAGCTAGTTGGTGGGGTAAGAGCCTACCAAGGCGACGATCCGTAGCTGG

M35016  287 TTTGAGAGAATGATCAGCCACATCGGGACTGAGACACGGCCCGAACTCCTACGGGAGGCA
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pSKrR1  329 TTTGAGAGAATGATCAGCCACATCGGGACTGAGACACGGCCCGAACTCCTACGGGAGGCA

M35016  347 GCAGTGGGGAATATTGGACAATGGGGGGAACCCTGATCCAGCAATGCCGCGTGTGTGAAG
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pSKrR1  389 GCAGTGGGGAATATTGGACAATGGGGGGAACCCTGATCCAGCAATGCCGCGTGTGTGAAG

M35016  407 AAGGCCTTCGGGTTGTAAAGCACTTTTATTAGTGAAGAACGGTGCATGGTTAATACCCAT
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pSKrR1  449 AAGGCCTTCGGGTTGTAAAGCACTTTTATTAGTGAAGAACGGTGCATGGTTAATACCCAT

M35016  467 GCAATTGACATTAGCTAAGGAAAAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAT
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
pSKrR1  509 GCAATTGACATTAGCTAAGGAAAAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAT

M35016  527 ACGGAGGGTGCAAGCGTTATTCGGAAT 553
            :::::::::::::::::::::::::::
pSKrR1  569 ACGGAGGGTGCAAGCGTTATTCGGAAT 595
```

Fig. 6.1: Sequence alignment of 16S rRNA sequence (GenBank accession no M35016) with the cloned partial 16S rRNA sequence in the pSKrR1 plasmid, showing the qPCR primers in red and qPCR probe in pink.

6.2.2. Plasmid standards for *pnpA* and *aprV2*

The plasmid pGW19.2 which contains 650 bp from the 3' end of *pnpA*, nt 1 to 677 from GenBank accession no X98545, was cloned and sequenced by Whittle, (1999). This was chosen as a standard for determining *pnpA* mRNA levels. The plasmid pJH3.1 which contains 603 bp from nt 711 to 1314 of the *aprV2* gene (GenBank accession no X38395) was cloned and sequenced by Hyman (2006) and used as a standard for determining *aprV2* mRNA levels.

6.2.3. Design of probes and primers for RT-qPCR

For real-time PCR, the TaqMan system was used as it is more specific and reliable than the SybrGreen system, as SybrGreen can bind equally well to both non-specific products like primer dimers as well as the desired PCR product. TaqMan probes and primers for the RT-qPCR experiment were designed using the 16S rRNA sequence (GenBank accession no M35016) and *pnpA* sequence (Genebank accession no X98545). The primers and probes for the *aprV2* thermostable protease gene (Genebank accession no L38395) were designed by Hyman (2006). All probes and primers were designed using Beacon Designer software, such that the sense and anti-sense primers lay inside the 16S rRNA, *pnpA* and *aprV2* sequences present in the plasmids pSKrR1, pGW19.2 and pJH3.1, respectively and the probe lay between the sense and anti-sense primers. The primers and probes for the *pnpA* gene were from the 3' end of the *pnpA* gene at the PNPase KH and S1 RNA binding domains. The 16S rRNA and *aprV2* probes were labelled with the fluorophore FAM, while the *pnpA* probe was labelled with HEX. Sequences of all probes and primers are tabulated in Tables A4 and A5, and the positions of rRNA probe and primers are shown in Fig. 6.1.

6.2.4. Standard curve analysis for 16S rRNA

Analyses of the standard control plasmids are important for reliable mRNA expression analysis studies. The respective standard plasmid concentrations were evaluated on an agarose gel (Fig. 6.2) after linearising the plasmids at various dilutions with their respective restriction enzymes. The plasmid copy numbers were then calculated using the molecular weight of the plasmid, according to the equation in section 2.16.1.

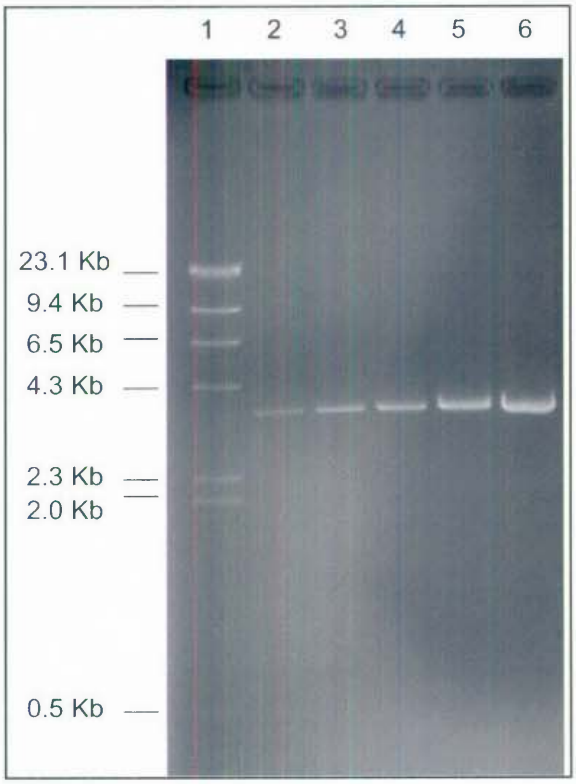


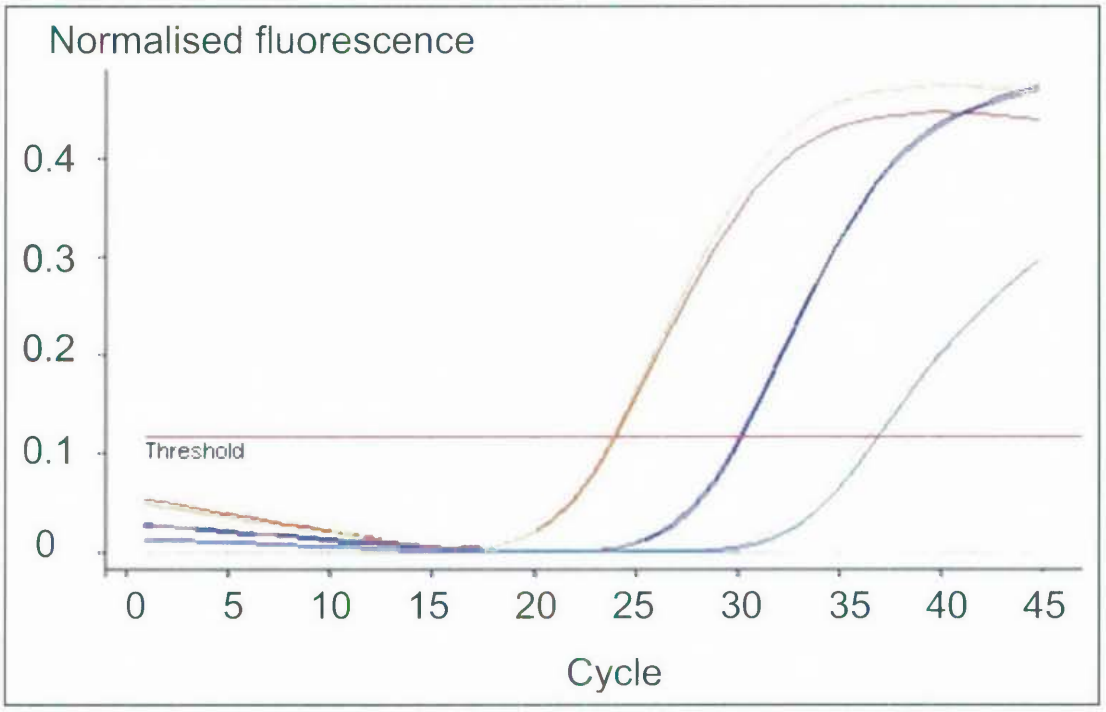
Fig. 6.2: A 1% agarose gel showing the diluted plasmid pSKrR1 linearised with *Pst*I. (Lane 1: λ/*Hind*III standard, lane 2: pSKrR1 0.125 µl, lane 3: pSKrR1 0.25 µl, lane 4: pSKrR1 0.5 µl, lane 5: pSKrR1 1 µl, and lane 6: pSKrR1 2 µl).

Once copy numbers were calculated, standard plasmid controls were run on the Rotor-Gene 3000 (Corbett Scientific) by setting up duplicates of 100 fold dilutions of the plasmid molecules over the range of 10^2 to 10^6 copies. The reliability of the experiments

can be rated against a set of values defined by Dorak (2006), as shown in Table 6.1. The efficiency of the reaction was confirmed after several trials and cross checking of the Ct values for each dilution between the assays. A sample electrograph of the 16S rRNA plasmid standard curve is shown in Fig. 6.3 and the results are tabulated in Table 6.2. This result shows that the variation between duplicate determinations is low and the assay is linear with respect to C_t value. The slope, correlation co-efficient (R^2) and reaction efficiency are within the desired ranges.

	Desired Range of values
Slope	-3.58 to -3.1
Efficiency	90% to 110%
R^2	> 0.99

Table 6.1: Desired range for qPCR parameters (Dorak, 2006).



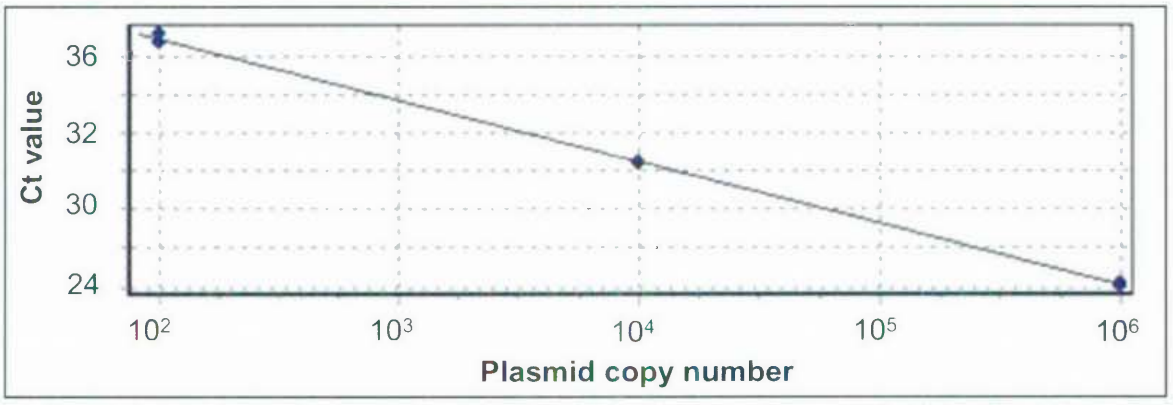


Fig. 6.3: Raw data from qPCR showing the 16S rRNA plasmid (internal) standard curve electrogram and standard curve.

	Desired Range of values	16S rRNA standard plasmid assay results
Slope	-3.58 to -3.1	-3.18313
Efficiency	90% to 110%	106%
R^2	> 0.99	0.99916

(a)

Name	Type	Ct value	Given conc Copies/reaction	Calculated conc Copies/reaction	% Var
pSKrR1 16S rRNA 10^6	Standard	24.08	1,000,000.00	1,024,621.81	2.5%
pSKrR1 16S rRNA 10^4	Standard	30.44	10,000.00	10,267.38	2.7%
pSKrR1 16S rRNA 10^2	Standard	36.67	100.00	113.43	13.4%
16S rRNA no template control	NTC	-	-	-	-

(b)

Table 6.2: qPCR assay of rRNA plasmid standard (a) parameters for the assay and (b) the results with the given and calculated concentration of the plasmid DNA copy numbers with the % variation and their Ct value. NTC = no template control.

The internal control plasmids pSKrR1, pGW19.2 and pJH3.1 were checked using Taq polymerase instead of the RT / Platinum® Taq mix (Invitrogen). Values were very similar using both sets of enzymes (data not shown). Taq Pol was used instead of the reverse transcriptase enzyme to assess DNA contamination of RNA samples.

6.2.5. RNA isolation and quantitation

RNA from all the strains used in the work was isolated as previously described (Section 2.15.1). The RNA samples were analysed by agarose gel electrophoresis (Section 2.15.2), as shown in Fig. 6.4. Agarose gel electrophoresis showed that the RNA was not degraded as the ribosomal RNA bands were clear. The RNA concentration was determined using the NanoDrop 1000 spectrophotometer.

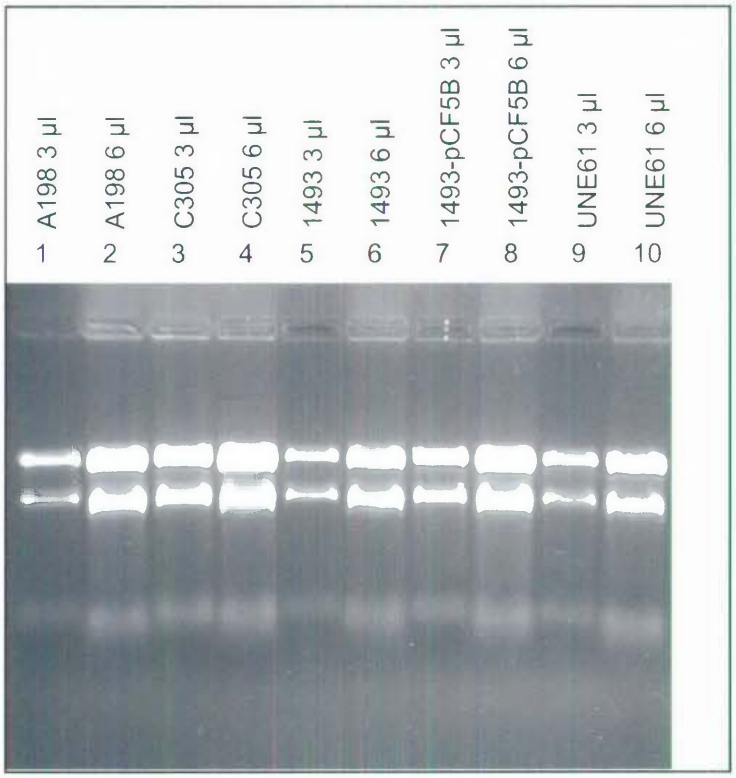


Fig. 6.4: An RNA gel showing total RNA isolated from strains A198, C305, 1493, 1493-pCF5B and UNE61, with 3 µl and 6 µl samples loaded adjacent to each other.

6.2.6. Expression analysis of 16S rRNA in the experimental strains

The RNA samples stored at -80°C were thawed and diluted to 1 ng/µl. 10 pg of RNA was then used for the 16S rRNA qPCR assays. Samples were assayed in triplicate against a plasmid standard curve covering the range of 10^4 to 10^8 molecules/reaction. A sample electrograph is shown in Fig. 6.5, together with the reaction parameters and the standard curve. Results from the assay of the samples are shown in Tables B1, B2, B4 and B5, appendix B.

The reaction efficiency, slope and Ct values were within the acceptable limits for analysis. There was 5 – 30% variation in the levels of 16S rRNA between the samples, even though all contained 10 pg of RNA. This variation could be due to variation in the purity of the RNA preparations. The concentration and the Ct values of the triplicates were in good agreement with each other. The Ct values did not show much variation within the same run, or between the two runs. This showed that the RNA samples were stable under storage and during sample handling to set up the assays. The samples were normalised against strain A198. The mean of the values obtained from the 16S rRNA determinants was used to normalise the results from the *pnpA* and *aprV2* assays.

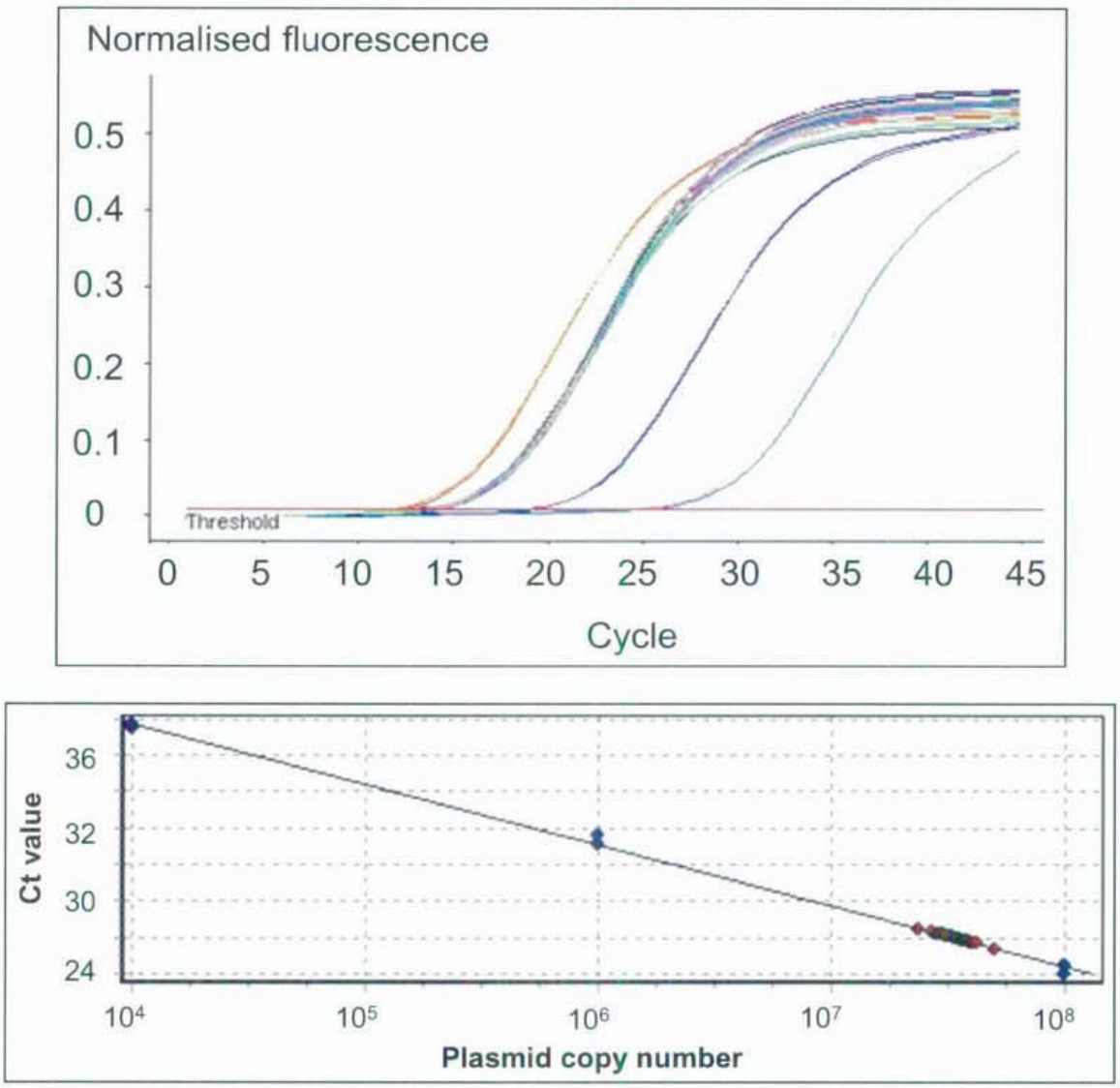


Fig. 6.5: The raw data of the 16S rRNA qPCR assay with pSKrR1 standard and RNA from the experimental stains, together with the standard curve. Plasmid standards are shown in blue, and RNA samples in red.

6.2.7. Expression analysis of *pnpA*

The *pnpA* mRNA levels in virulent and benign strains, and their respective *pnpA* knockouts were measured using 100 pg of total RNA in triplicate, using the pGW19.2 plasmid standard. The plasmid was diluted from 10^6 to 10^2 copies per reaction. The results were normalized with the 16S rRNA measurement levels for each sample. The assay was repeated three times and the results were reproducible with good agreement between the standards and the sample Ct values. The mean normalised round numbers are shown in Table 6.3 and the original results are given in Tables B3 and B6. Studies using Taq polymerase instead of reverse transcriptase showed that almost all RNA samples had undetectable DNA contamination, except strain 2483, where DNA contamination was estimated at less than 0.1%.

The prototype virulent strain A198 and virulent strains UNE61 and UNE64 had lower mRNA levels of *pnpA* than in the prototype benign strain C305, which is consistent with higher PNPase activity in the benign strain C305 than the three virulent strains. However, the benign strains 819, 1493 and 2483 had half the levels of *pnpA* mRNA found in strain C305. Overall, there was no correlation between *pnpA* mRNA level and virulence, or between *pnpA* mRNA level and PNPase activity.

In general, *pnpA* knockout strains had lower *pnpA* mRNA levels than the parent strains. However, the *pnpA* knockout of strain 819 had a higher *pnpA* mRNA, although the PNPase activity of the knockout was lower.

Results for *pnpA* mRNA levels measurement			
Strain	**Type**	**Ave of means**	**Standard Deviation**
A198	virulent	12,400	150
C305	benign	19,700	2,300
819	benign	7,600	230
819S1B	*pnpA*$^{\Delta}$	16,200	400
1493	benign	5,600	2,150
1493-pCF5B	*pnpA*$^{\Delta}$	5,800	1,400
2483	benign	3,500	850
2483-pCF5D	*pnpA*$^{\Delta}$	720	160
UNE61	virulent	7,200	1740
UNE61-1D8	*pnpA*$^{\Delta}$	3,900	630
UNE64	virulent	15,500	2,900
UNE64-1B	*pnpA*$^{\Delta}$	1,900	180

Table 6.3: Copy number of *pnpA* mRNA in 100 pg of RNA samples from benign, virulent and *pnpA* knockout strains (*pnpA*$^{\Delta}$) of *D. nodosus*. Raw data from Table B6 were normalised against relative 16S rRNA concentrations of the 10 pg RNA samples.

6.2.8. Measurement of *aprV2* levels in *D. nodosus*

Hyman (2006) analysed *aprV2* expression using RT-qPCR and found that the virulent strain UNE61 had significantly higher levels of *aprV2* mRNA than the benign strain 1493. However, it was later found that the part of the *aprV2* gene to which the forward primer binds spans a mutation site described by Riffkin *et al.* (1995) who showed that the genes *aprV2* (L38395) from the virulent strain A198 and *aprB2* from the benign strain C305 were identical except for mutations at 883 and 884 nt. These mutations change nt 883 and 884 from TA to CG, and thus the forward primer may bind more tightly to *aprV2* cDNA than to *aprB2*, which may explain the measured difference in expression. To avoid this problem, a new forward primer was designed, closer to the probe binding site, in a region which is identical in *aprV2* and *aprB2*.

A test assay was performed using the standard plasmid to estimate the copy number of *aprV2* mRNA present in a 100 pg sample of UNE64 RNA, a virulent strain expected to express high levels of *aprV2* (data not shown). This assay showed that the level was within the range of the standard curve, from 10^2 to 10^6 plasmid molecules/reaction. Subsequently the mRNA levels of the *aprV2* gene from virulent and benign strains and their respective *pnpA* knockouts were measured using 100 pg of total RNA in triplicates against pJH3.1 plasmid standards. The raw data are given in Tables B9 and B10. The results were normalized with the 16S rRNA measurement levels for each sample (Table 6.4). There were no results obtained for strain C305, which requires further study.

Results for *aprV2* mRNA levels measurement					
Strain	**Type**	**Copy number (Average of Means)**	**SD**	**% relative to A198**	**SD**
A198	virulent	346,700	113,000	100	0
C305	benign	-	-	-	-
819	benign	532,800	180,600	132.28	16.63
819S1B	*pnpA*$^{\Delta}$	250,400	74,100	53.74	9.63
1493	benign	428,800	166,400	97.58	29.12
1493-pCF5B	*pnpA*$^{\Delta}$	658,800	201,600	155.40	31.25
2483	benign	386,400	101,900	75.63	14.68
2483-pCF5D	*pnpA*$^{\Delta}$	113,200	45,700	22.88	3.32
UNE61	virulent	338,500	122,400	75.65	27.61
UNE61-1D8	*pnpA*$^{\Delta}$	167,100	55,`````600	37.98	4.64
UNE64	virulent	895,500	314,700	206.62	66.33
UNE64-1B	*pnpA*$^{\Delta}$	226,200	105,900	51.42	18.09

Table 6.4: Copy number of *aprV2* mRNA in 100 pg of RNA samples from benign, virulent and *pnpA* knockout (*pnpA*$^{\Delta}$) strains of *D. nodosus*. Raw data from Table B9 and B10 were normalised against relative 16S rRNA concentrations of the 10 pg RNA samples.

When the copy number was calculated from the plasmid standard curve, the variation between values obtained in the four different experiments for the same RNA samples was high, with standard deviations of 30 – 47% of the mean. However, analysis of the data (Tables B9 and B10) shows that Ct values for the RNA samples are quite constant, while Ct values for the plasmid standards vary quite widely. For example, the Ct values for 10^6 molecules of the plasmid standard vary from 20.6 – 23.2, while the Ct values for A198 RNA vary from 23.5 to 24.5. This suggests that there may be some instability in the plasmid standard DNA. To overcome this problem, copy number for all RNA samples was expressed as a percentage of the copy number in strain A198. The percentage values for the four experiments were averaged, to give the values in Table 6.4.

Variation between *aprV2* mRNA levels between the six parent strains was not large, and there was no correlation between the amount of *aprV2* mRNA and virulence. The *aprV2* mRNA level was decreased in all knockout strains compared to the parent strains, except for strain 1493. PNPase knockouts had no effect on protease thermostability in all strains except UNE64, where protease thermostability was decreased. The *pnpA* knockout in strain UNE64 produced the largest decrease in *aprV2* mRNA, which is consistent with a reduction in protease thermostability.

Overall, the results do not show much variation in *aprV2* mRNA levels between virulent and benign strains. These results suggest that the increased thermostability of secreted proteases from virulent strains is not due to increased transcription of *aprV2*, or to increased stability of *aprV2* mRNA. However, translation of *aprV2* mRNA may be inhibited in benign strains. CsrA specifically inhibits the translation of certain mRNAs in

E. coli (Liu and Romeo, 1997) and GlpA may act as a virulence repressor in *D. nodosus* by inhibiting translation of *aprV2* mRNA, and mRNA from other virulence determinants.

6.2.9. Sequence analysis of *aprV2* and *aprB2* genes from virulent and benign experimental strains

aprV2 codes for an extracellular acidic protease gene in *D. nodosus* and is present in both benign and virulent strains. In benign strains the gene is called *aprB2* and Riffkin *et al.* (1995) identified only two bp changes between *aprV2* and *aprB2* at nucleotide positions 883 and 884. This results in a single amino acid substitution of tyrosine in strain A198 to arginine in strain C305. This is a non-conservative substitution which could affect the activity or thermostability of the proteases. Only the *aprV2* sequence of strains UNE64 (Genebank accession number CP000513), A198 (L38395) and C305 (Riffkin *et al.*, 1995) were already known, so benign strains 819, 1493 and 2482 and the virulent strain UNE61 were sequenced. The mutations shown by Riffkin *et al.* (1995) were present in all the benign strains 819, 1492 and 2483 but not in virulent strains UNE61 and UNE64. The results are shown in Fig. 6.6.

Virulent strains

```
A198  (L38395)      TTGATAACTGGGCATGCGGTGGTTATCCTGATCCACGCAAAGAAAGAA
                    ::::::::::::::::::::::::::::::::::::::::::::::::
UNE61 sequenced     TTGATAACTGGGCATGCGGTGGTTATCCTGATCCACGCAAAGAAAGAA
                    ::::::::::::::::::::::::::::::::::::::::::::::::
UNE64 (CP000513)    TTGATAACTGGGCATGCGGTGGTTATCCTGATCCACGCAAAGAAAGAA
Benign strains      ::::::::::::::::::::::  ::::::::::::::::::::::::
819   sequenced     TTGATAACTGGGCATGCGGTGGTCGTCCTGATCCACGCAAAGAAAGAA
                    ::::::::::::::::::::::::::::::::::::::::::::::::
1493  sequenced     TTGATAACTGGGCATGCGGTGGTCGTCCTGATCCACGCAAAGAAAGAA
                    ::::::::::::::::::::::::::::::::::::::::::::::::
2483  sequenced     TTGATAACTGGGCATGCGGTGGTCGTCCTGATCCACGCAAAGAAAGAA
                    ::::::::::::::::::::::::::::::::::::::::::::::::
C305                TTGATAACTGGGCATGCGGTGGTCGTCCTGATCCACGCAAAGAAAGAA
C305 sequence shown by Riffkin et al. (1995)
```

Fig. 6.6: Sequence of part of *aprV2* from *D. nodosus* strains.

During the course of this work, the complete sequences of *aprV2* from the virulent strain A198 (GenBank accession no L38395) and the virulent strain UNE64 (CP000513) *aprV2* were aligned, and five differences were found, which are tabulated in Table 6.5.

Position of mutation identified in *aprV2* (L38395)	Strain A198 (L38395)	Strain UNE64 (CP000513)	Possible amino acid change
286 -287	CG	GC	Alanine to arginine (A-R)
1284	C	G	Threonine to serine (T-S)
1314	G	C	Silent mutation
1955	C	-	Frame shift
1985	-	C	Frame shift

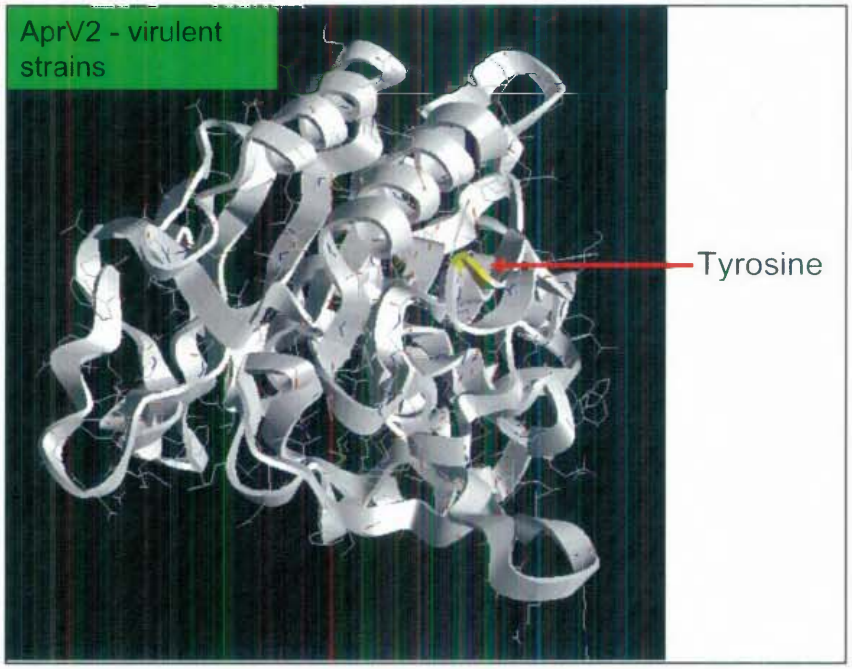
Table 6.5: The mutations identified in strain A198 (L38395) and UNE64 (CP000513) *aprV2* gene "these two strains are virulent strains".

These sequence differences could have a significant effect on the structure of AprV2, particularly the frame shift mutation, which would alter ten amino acids. To investigate these differences, the sequence of this part of *aprV2* from strains A198, 819, 1493, 2483 and UNE61 was determined, using the primers BCH332, BCH359, BCH395, BCH396, BCH398 and BCH399 (Table A1). The sequencing results showed that the sequence differences reported in strain A198 (L38395) were due to sequencing errors. All strains used in this work had an identical *aprV2* sequence to UNE64 (CP000513) at bp position 286, 287, 1284, 1314, 1955 and 1985.

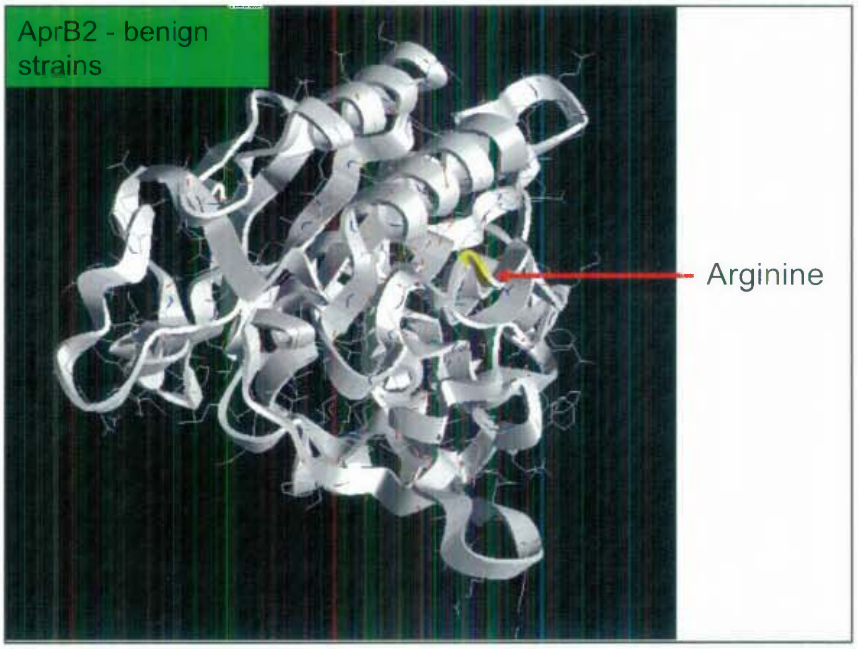
6.2.10. Protein structure prediction analysis of *aprV2* and *aprB2* protein from virulent and benign strains

Riffkin *et al.* (1995) found the tyrosine to arginine difference in sequence between AprV2 and AprB2 gene and confirmed that there is a higher pI value for AprB2 protein than

AprV2. This has not been explored further. The sequences of AprV2 and AprB2 were aligned with similar proteins, which showed that the amino acid at this position is not conserved (data not shown). To analyse the effect of this amino acid substitution on protein structure, the protein structure prediction tool SwissProt and MVM protein structure viewing tool were used. The predicted protein structures of AprV2 and AprB2 are shown in Figs 6.7 and 6.8. No differences in the overall folding pattern were seen (Fig. 6.7). However, minor changes to side chains are seen close to the amino acid substitution site (Fig. 6.8).

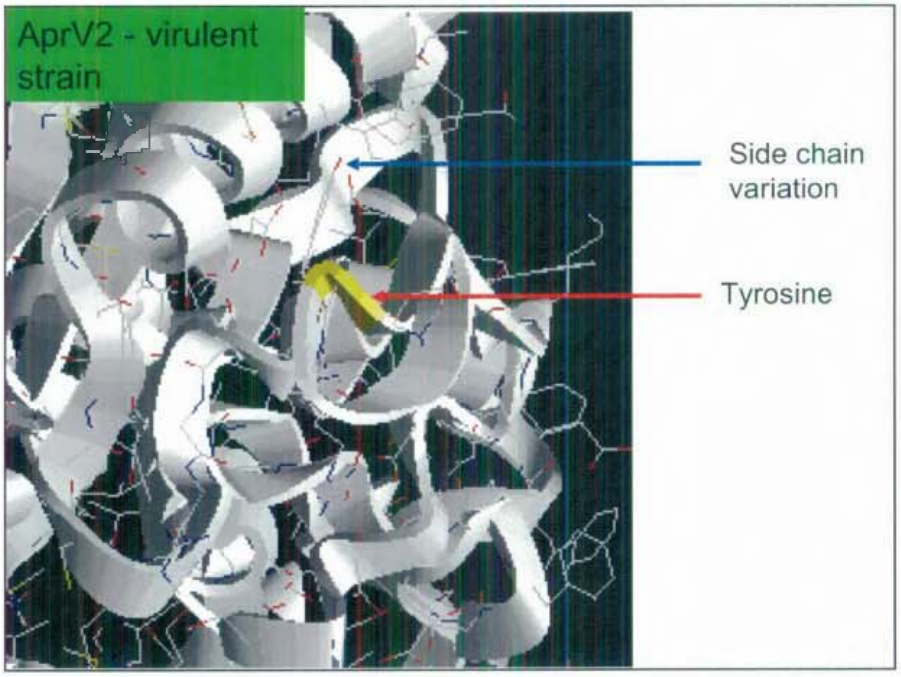


(a)

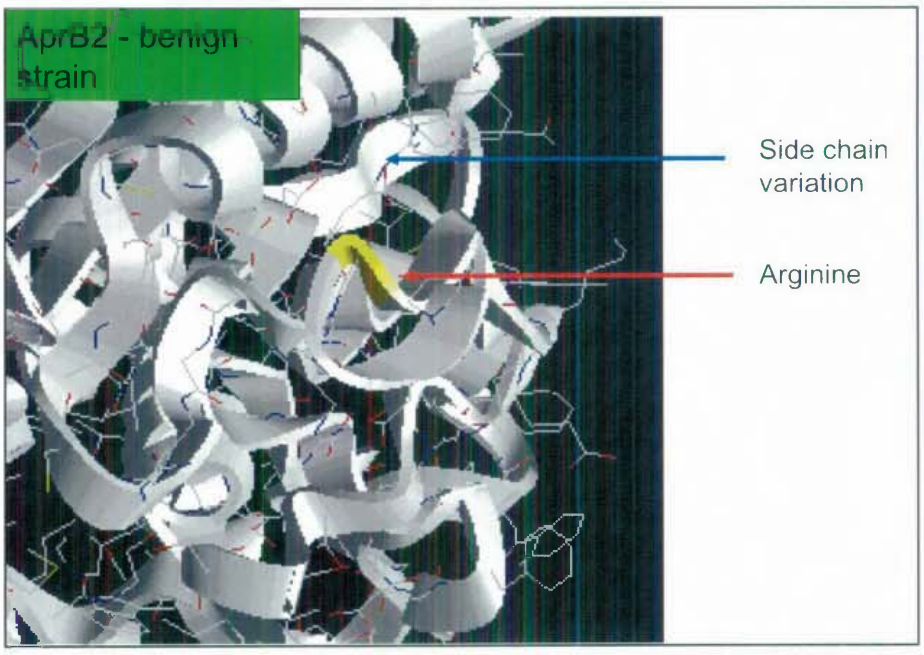


(b)

Fig. 6.7: Predicted structures of AprV2 (a) and AprB2 (b) showing arginine for tyrosine at position 219.



(a)



(b)

Fig. 6.8: Close up of predicted structures of AprV2 (a) and AprB2 (b) showing the effect of substituting arginine for tyrosine at position 219.

6.3. Discussion

Methods were developed for quantifying the levels of 16S rRNA, *pnpA* mRNA and *aprV2* mRNA. The assays were sensitive and could detect less than 1000 copies of the RNA molecules. The rRNA and *pnpA* mRNA assays were highly reproducible. However, there was a lot of variation between repeated assays of the plasmid standard for the *aprV2* assay, suggesting that the DNA preparation was unstable.

There was no correlation between the *pnpA* mRNA level and the virulence of the strains. These results do not support the hypothesis that the integrated elements downstream from *pnpA* alter the stability of *pnpA* mRNA. However, it is still possible that the different integrated elements alter the ability of *pnpA* mRNA to be translated.

The *pnpA* knockout strains generally had lower levels of *pnpA* mRNA, but in strain 819, the *pnpA* knockout has a 2 fold increase in *pnpA* mRNA level. In all strains except 819, *tetM* was inserted by a double crossover. In strain 819, a single crossover occurred. This difference could result in a difference in the 3' end of *pnpA* transcripts, which could alter their stability.

There was no correlation between *aprV2* mRNA levels and the virulence of the strains. Levels of *aprV2* mRNA were generally lower in *pnpA* knockout strains, except strain 1493. These findings are consistent with the protease thermostability assays for PNPase knockouts, and do not support the hypothesis that PNPase acts as a negative regulator of *aprV2*. The strain UNE64 with the greatest decrease in *aprV2* mRNA levels also shows the only significant decrease in protease thermostability.

The *aprV2* sequence analyses of the benign and virulent strains have shown that all the benign strains have the mutation identified by Riffkin *et al.* (1995) which changes tyrosine to arginine. The protein folding analysis in-silico (Marchler-Bauer *et al.*, 2007) has not identified any significant difference between the benign (*aprB2*) and virulent (*aprV2*) proteins and also the alignment study of AprV2 with other bacterial strains showed that the mutation of Y to R is not conserved. It is possible that the reduced thermostability of extracellular proteases in benign strains is due to this mutation, but studies of a larger number of strains would be needed to verify this.

Chapter 7. Discussion and future directions

7.1. Model for virulence

Whittle *et al.* (1999) proposed that in *D. nodosus* PnpA and GlpA are global repressors of virulence, which act as post-transcriptional repressors of genes encoding virulence determinants, such as genes involved in fimbrial biogenesis and extracellular protease production. According to this model, the integration of genetic elements downstream from *pnpA* and *glpA* alters the expression of these two genes, by altering the 3' ends of their transcripts. This could result in a change in stability of *pnpA* and *glpA* mRNA, or alter the ability of the mRNA to be translated. Thus, PnpA and GlpA activity is modulated by the integration of genetic elements, and PnpA and GlpA in turn modulate the expression of virulence determinants. In the present study, this model was tested by measuring PNPase activity in benign and virulent strains, creating *pnpA* knockouts and determining their effects on characteristics associated with virulence, and using qPCR to measure the levels of *pnpA* and *AprV2* transcripts.

7.2. Differences in PNPase activity between strains

A small increase in PNPase activity was seen in benign strains compared to virulent strains. However, it would be necessary to measure PNPase activity in a larger number of benign and virulent strains to determine whether this difference in activity is significant, as well as *D. nodosus* strains from a range of environmental conditions may also need to

be investigated. These could include culture media, incubation temperature etc, all of which may impact global gene expression. The *pnpA* knockout strains, which had lost the KH and S1 domains, had a small reduction in PNPase activity, similar to *E. coli* PNPase with the KH and S1 domain deleted (Briani *et al.*, 2007). The S1 domain may not be involved in the catalytic activity or in the autoregulation of PnpA (Jarrige *et al.*, 2002; Amblar *et al.*, 2007), but the KH domain may affect autoregulation of PnpA (Garcia-Mena *et al.*, 1999), which could reduce the catalytic activity of PNPase. The creation of a mutant lacking only the S1 domain could determine the relative roles of the KH and S1 domains in PNPase activity.

PnpA is an autoregulatory protein and is also controlled by an RNase III dependent mechanism (Robert-Le Meur and Portier, 1992; 1994; Jarrige et al., 2001). *D. nodosus* does have an RNase III gene (GenBank accession no CP000513), thus an RNase III deletion in a *D. nodosus* virulent strain could be constructed. This may increase the PnpA level and could affect the virulence of the strain.

7.3. Effect of *pnpA* knockout on twitching motility

In agreement with the hypothesis that PnpA is a virulence repressor in *D. nodosus*, *pnpA* knockouts in benign strains resulted in an increase in twitching motility. The C-terminal RNA binding domains of PnpA could be involved in the regulation of genes required for fimbrial biogenesis. The S1 domain in PnpA may not be just involved in providing a base for RNA binding. It might also be involved directly in the regulation of fimbrial biogenesis, similar to the situation in *Yersinia spp.* where the PnpA S1 domain plays a direct role in the regulation of the type three secretion system (Rosenzweig *et al.*, 2005).

The invidual roles of the KH and S1 domains of PNPase in twitching motility could be studied by restoring the domains independently in the *pnpA* knockout strains.

7.4. Effect of *pnpA* knockout on protease thermostability

These studies suggest that PnpA is not involved in the regulation of secretory proteases, as the *pnpA* knockouts did not alter protease thermostability, except in strain UNE64 which requires further study. The model for virulence (Whittle *et al.*, 1999) proposes that both PnpA and GlpA regulate genes encoding virulence determinants. Thus, in *D. nodosus* PnpA may regulate fimbrial biogenesis and GlpA may regulate protease production. However, Kenny (2004) produced a double crossover *glpA* knockout in the benign strain 819 which did not show increased protease thermostability (Locke, 2006). This finding needs to be repeated in a range of benign and virulent strains, but it suggests that GlpA does not repress thermostable protease production. It is possible that thermostable protease production is modulated by a third regulator, tmRNA.

tmRNA is a positive regulator of the virulence-associated YoP proteins of *Yersinia pseudotuberculosis*, as tmRNA knockouts reduce YoP protein production, and abolish the ability of *Y. pseudotuberculosis* to survive in macrophages and cause mortality in mice (Okan *et al.*, 2006). tmRNA mediates mRNA decay, a process involving RNase R, and independent of PNPase. In *D. nodosus*, the *vrl* integrates into *ssrA*, which encodes tmRNA. It is possible that tmRNA regulates protease secretion in *D. nodosus* independently of PNPase, and that integration of the *vrl* alters tmRNA activity.

However, a role for PnpA in regulating thermostable proteases production cannot be ruled out, as the knockout strains show only a 40% reduction in phosphorolytic activity.

The remaining activity, or the remaining functional domains of the protein, could be sufficient for regulation of thermostable protease production.

7.5. Expression of *pnpA* mRNA

Transcription of *pnpA* extends into the adjacent integrated elements. Thus, the 3' ends of the transcripts are different in strains with different genetic elements integrated next to *pnpA*. This could alter the stability of the mRNA, or its ability to be translated. qPCR measurements of *pnpA* mRNA levels in different strains showed no correlation between the amount of transcript and the integrated element next to *pnpA*, suggesting that the different integrated elements present at the 3' ends do not affect the stability or expression of the mRNA. However, the ability of these transcripts to be translated could be altered.

Transcript levels in *pnpA* knockouts were generally lower than in the parent strains. This is of interest, because disruption of the *pnpA* gene should not affect the promoter, and thus the rate of transcription should be unaltered. This would suggest that the reduced level of *pnpA* mRNA in *pnpA* knockout strains is due to reduced stability of the mRNA. The 3' end of the transcript would be different in *pnpA* knockout strains. The only knockout strain to show an increase in *pnpA* mRNA levels was strain 819, which is generated by a single, rather than a double, crossover event. Thus, in this strain, the 3' end of the transcript would again be different, and may result in increased transcript stability.

The production of *pnpA* transcripts is likely to be very complex, as the *pnpA* stop codon is only six nucleotides from the 5' end of the adjacent *tRNA*. Thus, *pnpA* transcripts contain this *tRNA*, which is essential as it is present in only one copy in the genome

(Myers *et al.*, 2007). This suggests that *pnpA* transcripts are processed to remove the tRNA molecules. Differences in the composition of the 3' end of the *pnpA* transcripts due to the presence of different genetic elements integrated into the tRNA could therefore affect processing of the transcript, as well as its ability to be translated.

7.6. Expression of *aprV2* mRNA

The virulent strains and benign strains in general did not show much variation in the *aprV2* mRNA level. One possibility is that the other two proteases, *aprV5* and *hprV* are regulated, and not *aprV2*. This is unlikely, but could be investigated by setting up similar assays to measure mRNA levels for these two genes. Alternatively, there may be a post-transcriptional control method to regulate translation of these transcripts. This is consistent with the finding that the benign and virulent *pnpA* knockout strains showed a reduction in *aprV2* expression, but the thermostability of the proteases were not changed in the protease assays. Post-transcriptional regulation could occur by folding of transcripts to prevent translation, binding of proteins, such as GlpA, to prevent translation, or differential termination of translation by tmRNA.

7.7. Mutations in PnpA

During the course of this work, it was found that there was an amino acid difference between PnpA from benign and virulent strains. Although this amino acid is not found at a conserved position, it may affect the thermostability or activity of PnpA.

To investigate the significance of this amino acid change, it would first be necessary to survey a much larger number of virulent and benign strains. If all benign strains had

proline at this position and all virulent strains had leucine, then the role in virulence could be investigated by creating mutants of benign strains with leucine at this position, and mutants of virulent strains with proline at this position, and determining the effects of these mutations on virulence characteristics. If the amino acid position is important for virulence, then integrated genetic elements may not control PNPase, in contrast to the model proposed by Whittle *et al.* (1999).

7.8. Mutations in AprV2

An amino acid difference was also seen in the sequence of AprV2 from benign and virulent strains. As with the mutation in PnpA, the significance of this could be determined by analysing a much larger number of benign and virulent strains. If the results show that arginine is always found at position 219 in benign strains and tyrosine at this position in virulent strains, then point mutations could be created to determine whether this amino acid change alters the thermostability of AprV2. Alternatively, this question could be answered by cloning *aprV2* and *aprB2* into an expression vector, and determining the thermostability of the expressed proteases.

7.9. Lineage of benign strains

Specific amino acid changes in PnpA and AprV2 have been associated with benign strains. It is possible that all benign strains arose from a common ancestor with these mutations. However, all the serogroups of *D. nodosus* contain both virulent and benign stains. This implies that the mutations identified in benign strains must have arisen before the serogroups diverged, and that serogroups must have arisen independently in both a benign and a virulent lineage, which seems unlikely. An alternative explanation is that a

mechanism exists for fimbrial variation so that serogroups may be altered. Although transformation of *D. nodosus* with *fimA* has been shown to cause serogroup conversion (Kennan *et al.*, 2003), there is no evidence that such a process occurs in nature. By contrast, the model proposing that virulence is modulated by integrated genetic elements easily accounts for benign and virulent phenotypes in all serogroups, since these genetic elements can be lost, as has been seen in culture, and mechanisms exist, such as conjugation and transduction for their acquisition. Thus, the virulence of strains can be easily altered by changing the pattern of integrated genetic elements.

7.10. Modification of integrated elements in strains

The current study did not show a correlation between *pnpA* mRNA levels and the genetic element integrated next to *pnpA*. However, this does not rule out a model where the adjacent integrated elements affect the ability of *pnpA* mRNA to be translated. To fully investigate the potential role of integrated elements in virulence, mutant strains could be constructed in which the integrated elements (or part of these elements) next to *pnpA* were altered and the effects of these alterations on virulence characteristics could be determined.

7.11. Conclusions

These studies have shown that PnpA acts as a virulence repressor in *D. nodosus*, since *pnpA* knockouts in benign strains result in increased twitching motility, a virulence characteristic. However no change was seen in protease thermostability in *pnpA* knockout strains. Amino acid differences between the sequences of PnpA and AprV2 in benign and virulent strains were identified. Further studies would be needed to determine the role, if

any, of these amino acid differences in virulence. Further work to test the proposed modulation of virulence by integrated genetic elements (Whittle *et al.*, 1999) could include mutation studies to delineate the roles of *glpA* and *ssrA* in virulence, and the construction of strains with altered patterns of integrated genetic elements.