strain (Figure 7.12). Most benign strains hybridise to a 4.1 kb *Hin*dIII fragment, which corresponds to a copy of *pnpA* that is adjacent to $intB_N$ (C3052, 1169, 2483, 1493, 1469). In strain 819, the *intD* element is integrated adjacent to *pnpA*, and consequently this strain hybridises to a *Hin*dIII fragment of 5.9 kb.

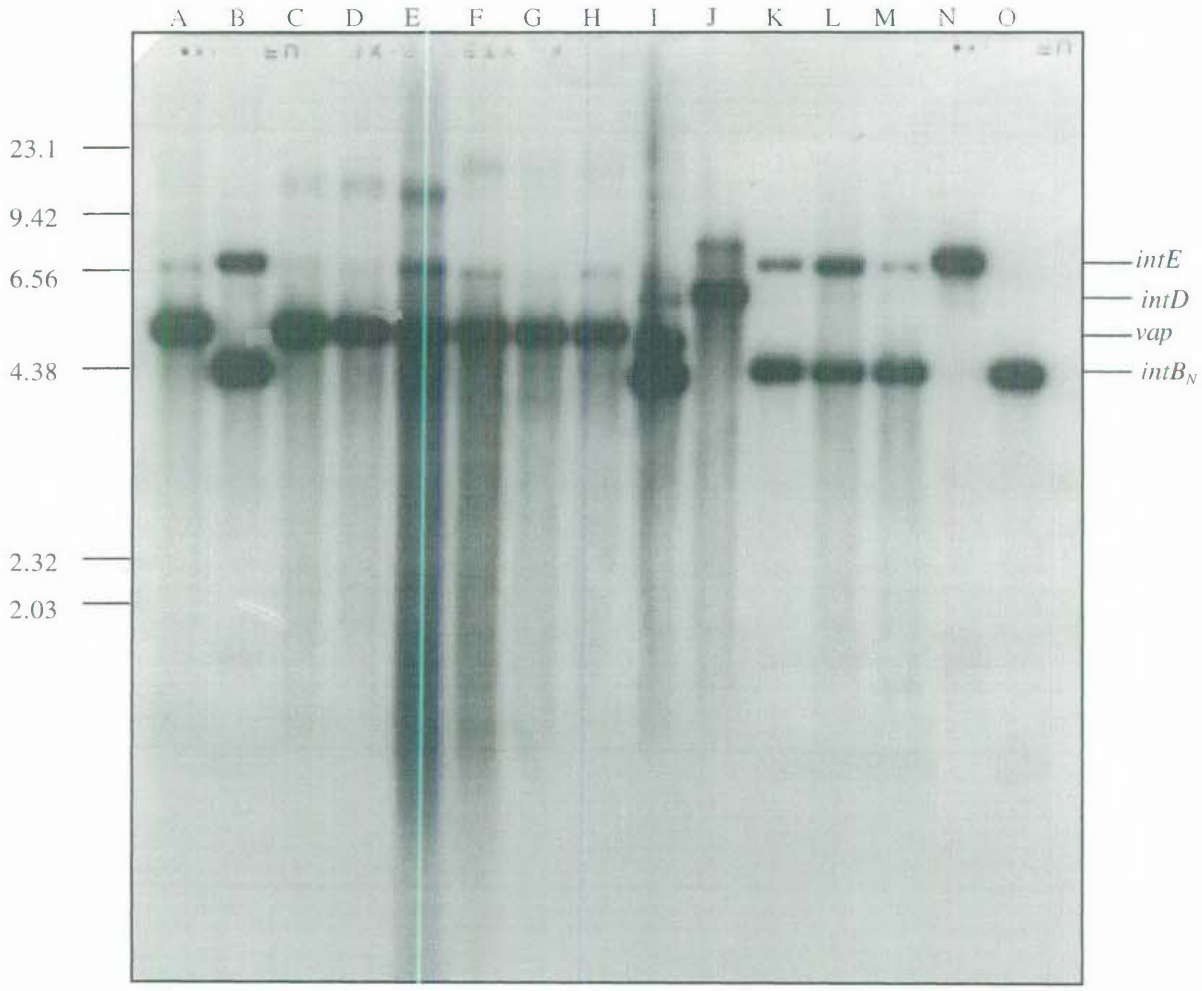

Figure 7.12: Southern blot analysis of *pnpA* integration site in seventeen strains of *D. nodosus* as follows: A198 (lane A), C305 (lane B), 1311 (lane C), 1311A (lane D), AC3577 (lane E), B1006 (lane F), G1220 (lane G), H1204 (lane H), H1215 (lane I), 819 (lane J), 1169 (lane K), 2483 (lane L), 1493 (lane M), 3138 (lane N), and 1469 (lane O). Genomic DNA was digested with *Hin*dIII. Lambda *Hin*dIII standard sizes are indicated at the left of the panel in kb. DNA fragments specific for *pnpA* were used as a probe (Appendix 8). On the right of the above panel, the position of restriction fragments that result from the integration of *intE* (7.1 kb), *intD* (5.9 kb), *vap* (5.0 kb), and $intB_N$ (4.1 kb) elements next to *pnpA* are indicated.

In strain 3138, the *pnpA* probe hybridises strongly to a 7.1 kb *Hin*dIII band. It is currently unknown what is integrated adjacent to *pnpA* in strain 3138, but Southern blot analyses (Appendices 4-10) show that it is not the *vap*, *intB*, *intC* or *intD* elements. Benign strains C3052, 1169, 2483 and 1493 also hybridise faintly to this same 7.1 kb band in addition to a strong band of 4.1 kb (Figure 7.12).

The presence of this 7.1 kb band is not attributable to partial digestion since this same membrane was stripped and reprobed with *gepA* (Appendix 5), and only single restriction fragments hybridised to the *gepA* probe; also it is a fragment of this exact size that is fixed in the *pnpA* position in strain 3138 (Figure 7.13). In addition, the presence of this 7.1 kb fragment is not the result of the loss of $intB_N$, since 3138 still contains $intB_N$, which is integrated upstream of the *intD* element in this strain (Section 6.2.2 and Figure 7.13). It could also be hypothesised that site-specific inversion of $intB_N$ and downstream sequences could generate this 7.1 kb band. However, this seems unlikely because all strains of *D. nodosus* studied in this work contain a copy of $intB_N$ and downstream sequences, thus it seems unlikely that $intB_N$ can only invert in four of seventeen strains analysed. The most plausible explanation is that the 7.1 kb band may correspond to the integration of another genetic element that has not yet been isolated from *D. nodosus*, that will be herein referred to as the *intE* element.

The presence of the faint 7.1 kb *Hin*dIII bands could be explained if the *intE* element were present in these strains, and able to excise at high frequency, but be maintained either as a circular molecule or integrated in a different location. In strains C3052, 1169, 2483, and 1493, this *intE* element may be present elsewhere in the *D. nodosus* genome and thus is able to integrate and excise site-specifically downstream of *pnpA*. Since it seems that these putative *intE* element sequences are permanently integrated next to *pnpA* in strain 3138 (Figure 7.12), this hypothesis could be tested by using I-PCR to amplify sequences adjacent to *pnpA* in *D. nodosus* strain 3138. Alternatively, since this *intE* element would also be

present in the C305 lambda library constructed previously (Bloomfield, 1997), *pnpA* could be used as a probe to identify lambda clones that contain *pnpA* and hybridise to 7.1 kb *Hin*dIII, 7.1 kb *Hin*dIII/*Eco*RI and 13.8 kb *Eco*RI fragments, like the putative *intE* element from *D. nodosus* strain 3138.

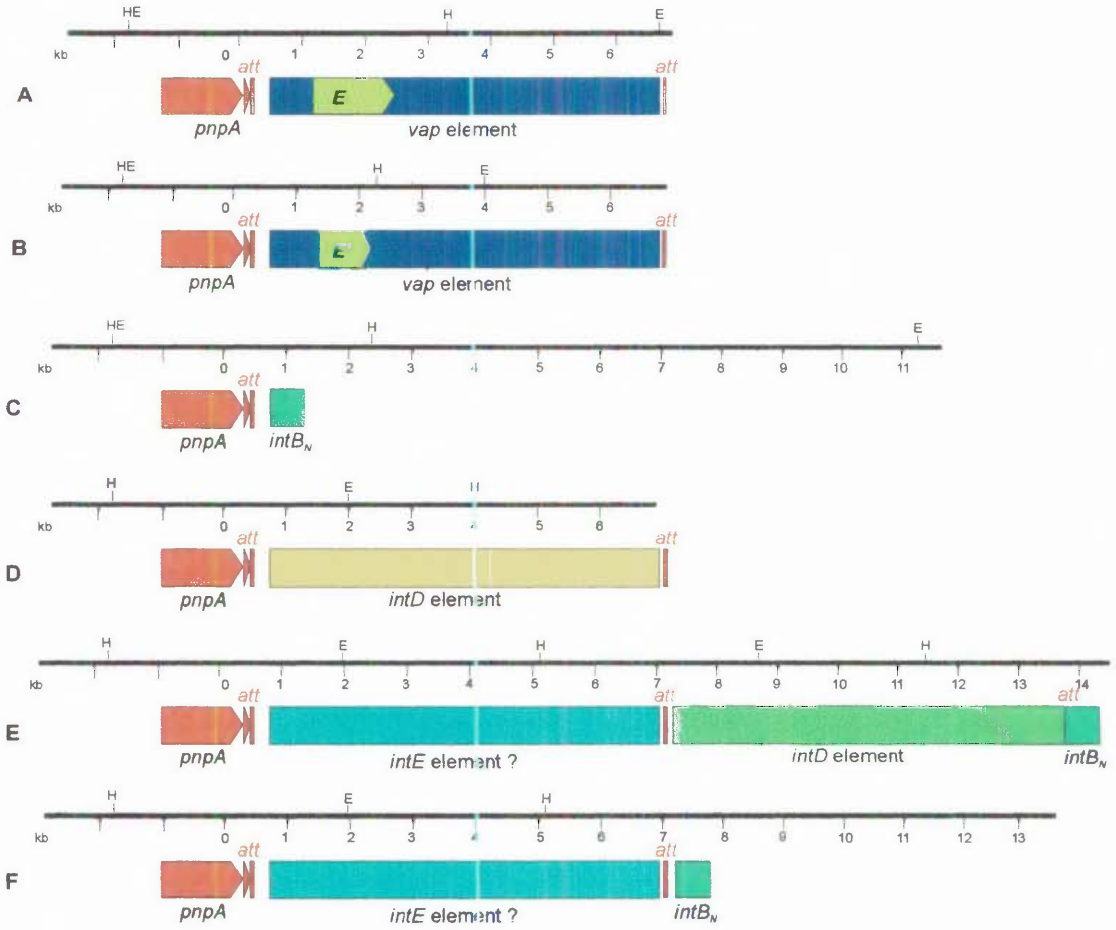

Figure 7.13: The variation in the *pnpA* integration site in seventeen strains of *D. nodosus*. In Southern blot analyses hybridisation to the *pnpA* probe results in seven restriction fragment polymorphisms, and are dependent upon the sequences integrated downstream from the *pnpA* gene. The different arrangements are as follows: (**A**) a copy of the *vap* element containing *vapE* is adjacent to *pnpA* (strains A198, 1311, 1311A, AC3577, B1006, G1220, H1204, D1172, AC390); (**B**) a copy of the *vap* element containing a copy of *vapE'* instead of *vapE* is adjacent to *pnpA* (strain H1215); (**C**) $intB_N$ and *pnpA* are adjacent (strains C3052, 1169, 2483, 1493, 1469); (**D**) the *intD* element and *pnpA* are juxtaposed (strain 819); (**E**) the putative *intE* element and *pnpA* are juxtaposed (strain 3138); (**F**) the putative *intE* element integrates between *pnpA* and $intB_N$ in a small population of cells (strains C3052, 1169, 2483, 1493). The features indicated are as follows: putative *att* sites (narrow red boxes), $tRNAser_{GGA}$ (red triangles). Genes and elements are labelled appropriately. Numbers indicate the distance in kb. Restriction sites shown include *Eco*RI (E) and *Hin*dIII (H).

In addition, the *intD* element must be different to the putative *intE* element because only six strains (1311, AC3577, B1006, D1172, 819 and 3138) contain the *intD* element sequences, and the putative *intE* element appears to be present in strains (C3052, 1169, 2483 and 1493) that do not contain the *intD* element sequences (Section 6.2.2). In addition, the *intD* element in strain 819 is adjacent to *pnpA* and probes specific for *intD* (Shaw & Cheetham, unpublished) and *pnpA* hybridise strongly to an *Eco*RI fragment of 6.3 kb. However, in strain 819, the *pnpA* probe also hybridises faintly to a 17 kb *Eco*RI fragment, which suggests that the *intD* element is lost in a small number of 819 cells which results in the relocation of $intB_N$ adjacent to *pnpA* (Figure 7.14). The *intD* element is not adjacent to *pnpA* in strain 3138, and this supports the hypothesis that a different genetic element, the *intE* element, is integrated adjacent to *pnpA* in 3138, and may be present elsewhere in the genome of other strains of *D. nodosus* (Figure 7.13). Although the excision of the *intD* element may be occurring in strain 3138, it is not detectable in *pnpA* blots, since the putative *intE* element separates *pnpA* and the *intD* element in this strain.

7.2.10 Comparison of sequences at the beginning of the *vap*, *intB* and *intC* elements

Results from the analysis of the integration sites of the *vap*, *intB* and *intC* elements in benign and virulent strains of *D. nodosus* suggest that, although there is no direct correlation between the presence of the *intB* and *intC* elements and virulence (Section 7.2.8), the site at which these elements are integrated may affect the ability of the isolate to cause disease by affecting the production or secretion of thermostable proteases (Section 7.2.6). In addition, Southern blot results indicate that virulent strains have either the *vap* element or the *intC* element next to *askA* and the *vap* element integrated next to *pnpA*.

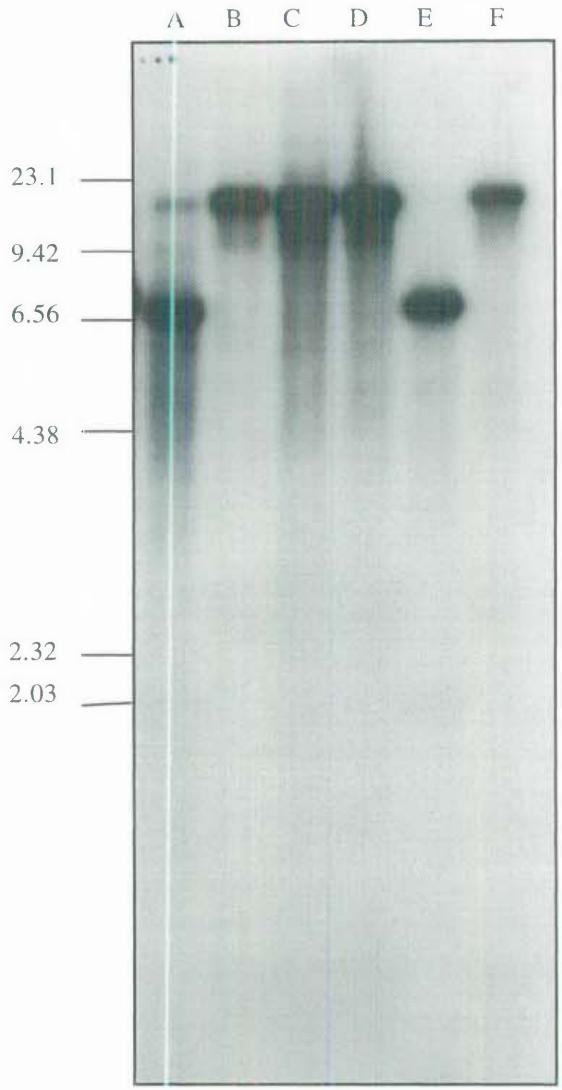

Figure 7.14: Southern blot analysis of *D. nodosus* benign strains 819 (lane A), 1169 (lane B), 2483 (lane C), 1493 (lane D), 3138 (lane E), 1469 (lane F). Genomic DNA was digested with *Eco*RI. DNA fragments specific for *pnpA* were used as a probe (Appendix 8). Lambda *Hin*dIII fragment sizes are shown at the left of the panel in kb. Strain 1169, 2483, 1493 and 1469 all hybridise to an intense band of ~17 kb, all of which have a copy of $intB_N$ adjacent to *pnpA*.

Between *askA* and the $tRNAser_{GCU}$ gene a new gene designated *glpA* has been identified (Section 1.6.8.1). GlpA (64 aa) has very high similarity (74% over the first 50 aa) to the RsmA protein from plant pathogen *Erwinia carotovora* (Mukherjee *et al.*, 1996) that acts as a global repressor of virulence. RsmA is known to suppress the synthesis of extracellular pectate lyase, polygalacturonase, cellulases, proteases, flagellum formation and motility and thus pathogenicity of soft-rotting *Erwinia* sp (Mukherjee *et al.*, 1996). The very high degree of similarity between RsmA and GlpA suggests that GlpA may have an analogous role to RsmA in *D. nodosus*.

To determine whether there were any sequence similarities or differences of significance upstream from the *intA*, *intB* and *intC* genes of the *vap*, *intB* and *intC* elements respectively, the DNA sequences starting from the *glpA* TGA stop codon for strains C305, A198 and 1311A were compared (Figure 7.15). The sequences between *askA-intC* in strain C305 had been determined previously (Bloomfield, 1997), as was the sequence between *askA-intA* in strain A198 (Cheetham *et al.*, 1995). However, in order to obtain the *askA-intB* 1311A sequence, a 1.1 kb *askA-intB* PCR product (Figure 7.5, panel II) was cloned and sequenced (Figure 7.15).

The sequences upstream of *intA*, *intC* and *intB* in the *vap*, *intA* and *intB* elements respectively, were scanned for putative promoter sequences which could potentially produce an antisense molecule that could sequester the transcripts produced on the opposite strand and hence affect transcript level of genes upstream of the integrated element. Sequences weakly resembling the consensus sequence for the *E. coli* σ^{70} -35 and -10 promoter sequences (Hawley & McClure, 1983) were identified in regions upstream of *intA* and *intC*, but were not present upstream of *intB* (Figure 7.15). Whether these are functional promoter sequences is uncertain, and primer extension or RNase protection studies would need to be undertaken to determine whether they correspond to a functional promoter region.

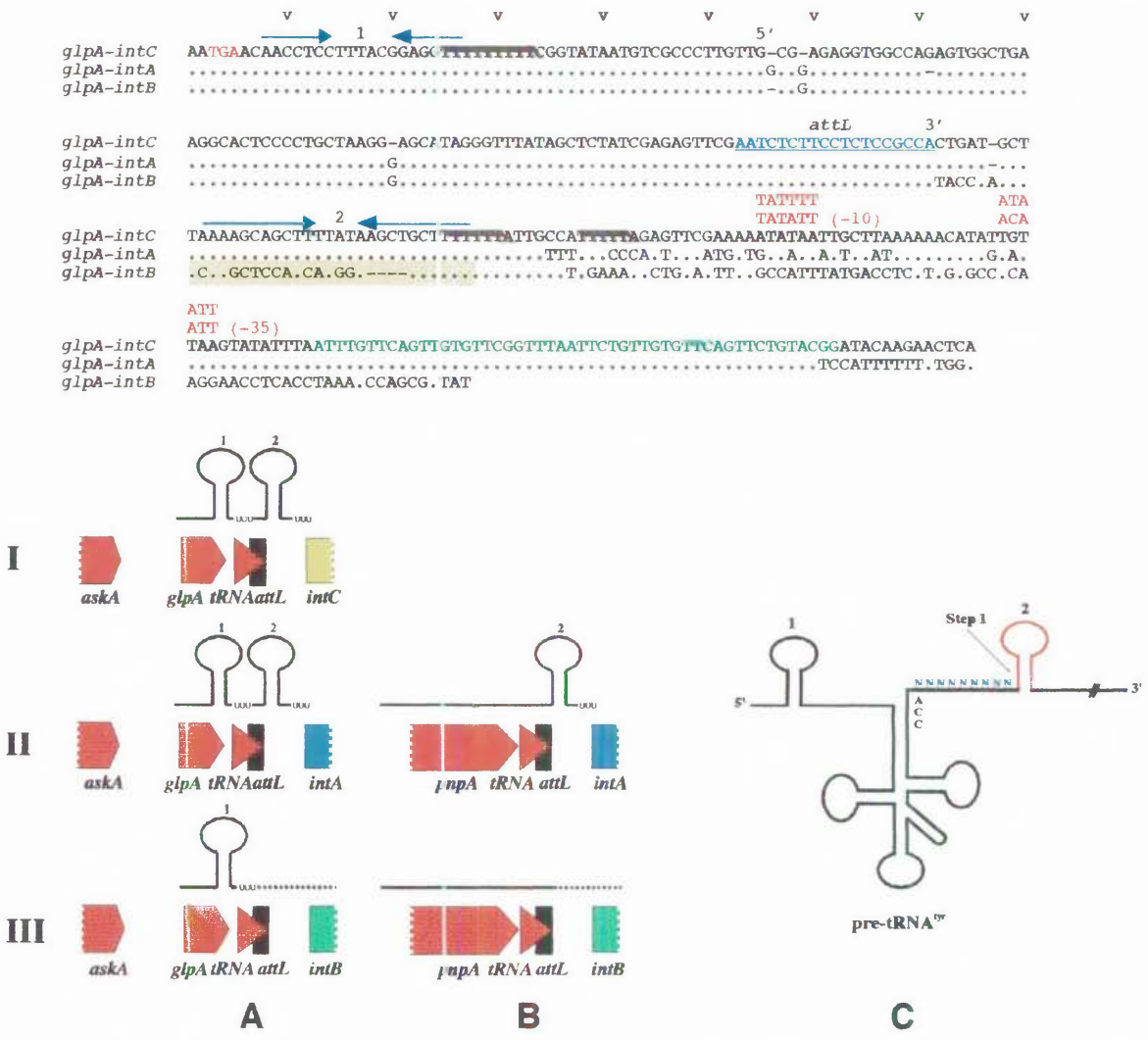

Figure 7.15: Dyad symmetry present in upstream region of the *vap* element and *intC* element, but absent from the sequences upstream of *intB* in the *intB* element.

(**Top**) Alignment of sequences downstream of *glpA* in *D. nodosus* strains C305 (*glpA-intC*), A198 (*glpA-intA*) and 1311A (*glpA-intB*). Nucleotides that are identical to nucleotides in the C305 sequence (top) are indicated by a fullstop. The TGA stop codon at the end of *glpA* is shown in red type. The two regions of dyad symmetry are indicated by inverted arrows and labelled 1 or 2. The *att* site that is present in all three sequences is shown in blue and underlined. The sequence that replaces the dyad symmetry in the *intB* element is shown in a box shaded yellow. The 3'-end of the *tRNA* gene in each sequence corresponds with the end of the *att* sequence, and is indicated. The five imperfect direct repeats that are present in the *intC* and *vap* elements are indicated in green type. Potential -35 and -10 promoter sequences that are present upstream of *intC* and *intA* are shown in red type above the sense strands.

(**Bottom left**) Diagrammatic representation (not to scale) of elements integrated next to *askA* (panel **A**) and *pnpA* (panel **B**) in *D. nodosus* strains. The *intC* element (**I**) may integrate next to *askA*, but has not been found next to *pnpA*. The *vap* element (**II**) and the *intB* element (**III**) or part thereof may integrate in either the *askA* or *pnpA* positions. In general virulent strains have *intA* or *intC* next to *askA*, and *intA* next to *pnpA*. Stem-loop structures 1 and 2 in the predicted mRNA species are shown above the genes. (panel **C**) The first step in the maturation of *E. coli* pre-tRNAtyr involves the endonucleolytic cleavage of the pre-tRNA transcript at a stem-loop structure at the 3'-end of the tRNA sequence.

Further analyses of the upstream sequences led to the identification of a region of dyad symmetry (loop 1, Figure 7.15) within the first thirty nucleotides that is present in all three sequences, and may act as a transcriptional terminator for *glpA*. The alignment between these sequences is almost identical up to the 3'-end of the *tRNA* gene in all three strains, after which there is a 42 bp sequence that is identical in the *vap* element and *intC* element sequences, but is not present in the *intB* element (Figure 7.15). Within this 42 bp sequence there is a region of dyad symmetry (loop 2, Figure 7.15) that is present in both the *vap* and *intC* elements, that is not present in the *intB* element sequence.

The presence of such a stem-loop may have the potential to alter the expression of genes situated upstream of it (loop 2, Figure 7.15) by interfering with the processing of transcripts from these genes or by a post-transcriptional mechanism such as retroregulation. Thus, if genes upstream of loop 2 have a direct or indirect role in virulence, the presence of this stem-loop structure may affect the expression of virulence determinants (Figure 7.15) such as thermostable proteases. Between *askA* and the integrated element there are two genes, *glpA* and a $tRNAser_{GCU}$ gene. At the other integration site, a $tRNAser_{GCU}$ gene separates *pnpA* from the integrated element.

The stem-loop (loop 2, Figure 7.15) structure might be required for the correct processing of the tRNAser transcripts that are adjacent to the integrated elements (Figure 7.15) in both the *askA* and *pnpA* positions. The first step in the maturation of the well characterised pre-tRNAtyr starts with the endonucleolytic cleavage by Pnpase or RNaseII (Zhongwei & Deutscher, 1994) of a stem loop structure from the 3'-end of the pre-tRNA sequence, which is located approximately nine nucleotides from the 3'-end of the mature tRNA molecule (Figure 7.15) (Sekiya *et al.*, 1979; Zhongwei & Deutscher, 1994). The stem-loop structure that is present in the sequences of the *vap* and *intC* elements is not present when the *intB* element is integrated and is also exactly nine nucleotides from the 3'-end of the tRNAser molecule (Figure 7.15). Thus, it seems likely that the stem-loop 2

(Figure 7.15) might be required for the correct processing of the tRNAser transcripts that are produced by the tRNAser genes that comprise the target site for integration of *vap*, *intB* and *intC* elements in *D. nodosus*. Consequently, the absence of the appropriate stem-loop structure when the *intB* element is integrated may alter the levels of functional tRNAser and in this way may affect the production of thermostable protease either directly, or *via* a regulatory gene that requires these particular tRNAser molecules. A role for *tRNA* genes in virulence has been described previously (Blum *et al.*, 1994; Dobrindt *et al.*, 1998; Gray, Wang & Gelvin, 1992; Hacker *et al.*, 1990; Hromockyj, Tucker & Maurelli, 1992; Morschhauser *et al.*, 1994; Ritter *et al.*, 1995) and, tRNAs are known to be involved in the regulation of other prokaryotic genes including antibiotic export genes (Fernandez-Moreno *et al.*, 1991) and genes involved in solvent formation (Sauer & Durre, 1992), and is the subject of further discussion in Section 7.3.

Unfortunately, in the absence of a transformation system for *D. nodosus*, it is not possible to do complementation studies to determine whether the *tRNAser* genes do affect the expression of virulence determinants (such as thermostable proteases), as they do in other pathogenic bacteria.

The stem-loop structure (loop 2, Figure 7.15) might also have a role in the retroregulation of upstream genes. Retroregulation can be defined as the regulation of a transcript by a downstream sequence, and is a novel post-transcriptional regulatory mechanism that involves sequences located at the distal rather than the proximal end of a gene (Birge, 1994; Schindler & Echols, 1981). The expression of lambda integrase is regulated by a retroregulatory mechanism (Court, Huang & Oppenheim, 1983). The *int* gene can be transcribed from the either the P_I promoter or the $P_{I'}$ promoter. P_I is positively regulated by *c*II/*c*III proteins which are involved in the repression of the lytic cycle, whilst the $P_{I'}$ is positively regulated by the antitermination protein *N* that promotes the lytic cycle. Post-transcriptional control of *int* is elicited by a sequence within the *b* region of lambda that

is able to form a stem-loop structure that is called *sib* (sitio inhibidor en *b*) (Mascarenhas *et al.*, 1983; Schindler & Echols, 1981). The *sib* structure formed by the P_L transcript has been demonstrated to be a signal for RNase III processing *in vivo* and *in vitro*, since *sib*-mediated regulation is lost in an RNase III$^-$ host (Court *et al.*, 1983; Schindler & Echols, 1981; Schmeissner *et al.*, 1984). Thus during the lytic cycle when transcription is predominantly from P_L, *sib* inhibits the production of Int. Transcription from P_I, on the other hand terminates at a different position and is unable to form the same stem-loop structure, and during lysogeny *sib* is separated from *int* and so does not inhibit the production of Int (Schindler & Echols, 1981).

Thus, in bacteriophage lambda, *sib* is involved in a type of recombinational control of genes required for lysogeny (Court *et al.*, 1983; Guarneros, 1988; Mascarenhas *et al.*, 1983; Schindler & Echols, 1981; Schmeissner *et al.*, 1984). It has been reported previously that, in lambda, the relocation of the *sib* sequence downstream of other genes is also able to result in retroregulation of these 'other' genes (Schindler & Echols, 1981; Schmeissner *et al.*, 1984).

There are two models that have been proposed to explain how the RNase III processing of the *int* transcript might reduce production of Int: *(i)* RNase III processing at the *sib* site leaves the 3'-end of the message susceptible to exonuclease degradation, thus preventing translation of the *int* transcript (Guarneros, 1988); or *(ii)* when transcription beyond the *sib* region occurs a duplex is formed between the *sib* site and a homologous sequence at the end of *int*. The duplex is an RNase cleavage structure, and processing at this site removes part of the *int* gene, thus preventing efficient translation (Schindler & Echols, 1981). It has not yet been determined which of these two models is the correct one. If the stem-loop structure that is present in the *D. nodosus* were regulating genes upstream *via* a retroregulatory mechanism, this might be visible as reduced expression of upstream genes due to an increase in the rate of RNA degradation or as a reduction at the level of translation

if the mechanism followed model *(i)* or *(ii)* respectively.

In *D. nodosus*, the analogous stem-loop structure (loop 2, Figure 7.15) could be involved in the recombinational control of expression of virulence determinants. That is, the integration of the *vap* of *intC* element results in the retroregulation of sequences external to the integrated element, that does not occur when the *intB* element is present in the same position, due to the absence of a stem-loop structure. It is also interesting that in both lambda and the *vap* and *intC* elements this stem-loop structure is adjacent to an integrase gene.

It does seem unlikely that the rate of degradation of loop 2 by RNase III would decrease the half-life of transcripts generated from upstream genes, because the secondary structure of *tRNA* genes are thought to protect tRNA molecules from degradation (Arriano, 1993). A consensus sequence for RNase III has not yet been established (Alifano, Bruni & Carlomagno, 1994). However, short stem-loop structures with a few base pair mismatches are thought to be good substrates for RNase III (Stern *et al.*, 1984), which can trigger RNA degradation either upstream or downstream of cleavage sites (Portier *et al.*, 1987; Schmeissner *et al.*, 1984). The stem-loop 2 that has been identified within the *vap* and *intC* elements (Figure 7.15) meets the criteria for a good RNase III substrate as described above.

In addition there is a 61 bp sequence which is located downstream of the of loop 2 (Figure 7.15) that is conserved in both the *vap* and *intC* elements but is not present in the *intB* element sequences. Within this region are five imperfect direct repeats of 10 bp with the consensus sequence 5′ -TT(T/G)TGTTC(A/G)G-3′ . The significance of this conserved region and the directly repeated region is unknown. It could be speculated, that like the directly repeated sequences which comprise the iterons at origins of replication regions, these imperfect direct repeats may be involved in protein binding that could as a consequence sterically hinder the processing of transcripts from upstream genes.

It has been proposed that the transcription of genes upstream of *vap* element and *intC* elements may be differentially regulated post-transcriptionally by an antisense mechanism, by affecting the processing of the upstream transcripts or by a retroregulatory mechanism. Furthermore, such post-transcriptional processes may be responsible for the absence of virulence-associated thermostable protease secretion in *D. nodosus* strain 1311A. This raises the question as to whether it is possible that any of the upstream genes other than the *tRNAser* genes which have already been discussed are, or could be, involved in virulence-associated processes.

The very high similarity between GlpA and the *E. carotovora* global repressor of virulence, RsmA, suggests that it is conceivable that if expression of the *glpA* gene is affected by the integration of a genetic element, then virulence in *D. nodosus* may be modulated accordingly.

RsmA is known to suppress the expression of virulence factors by repressing the levels of homoserine lactone synthase (*hslI*) (Swift *et al.*, 1993), and is thought to achieve this by binding to the *hslI* transcript and consequently accelerating the rate of mRNA degradation (Mukherjee *et al.*, 1996). This is interesting because the binding to the *hslI* transcript is thought to be mediated by the KH RNA-binding domain. Polynucleotide phosphorylases (*pnp*) also contain a KH RNA-binding domain (Liu, Yang & Romeo, 1995), and similarly affects the processing and decay of mRNA as one of the two principal enzymes involved in the degradation of bacterial mRNA (Alifano *et al.*, 1994; Li & Deutscher, 1994). Thus if the expression of *glpA* or *pnpA* are up- or down-regulated by the integration of the different genetic elements this could potentially alter stability of the transcripts that GlpA and PnpA are involved in degrading.

In addition, it has previously been reported that the integration of genetic elements alone can alter gene expression (Dorman, 1995; Ott, 1993). The need to respond rapidly to

changes in the environment requires the presence of global regulators of transcription that as their name suggests, have very widespread effects on gene expression (Dorman, 1994; Dorman, 1995). Many global regulators are proteins, however their effects are confined to those genes which have specific sequences to which they are able to bind. On the other hand, regulatory factors which contribute to the organisation of DNA topology also represent a form of global control which can potentially influence any gene in a given cell (Dorman, 1995). These topological effects include supercoiling, the formation of nucleoprotein complexes and duplex underwinding, which are influenced by numerous environmental factors including growth under anaerobic conditions (Balke & Gralla, 1987; Yamamoto & Droffner, 1985), osmolarity (Dorman & Ni Bhrian, 1993), growth temperature (Goldstein & Drlica, 1984), nutrient limitations (Higgins *et al.*, 1988), and the integration of genetic elements (Zagaglia *et al.*, 1991).

The *Shigella flexneri* virulence plasmid, pINV, integrates site specifically into the *metB* locus of the *S. flexneri* chromosome, and leads to methionine auxotrophy and a decrease in the expression of plasmid-encoded invasin genes. Both of these alterations in gene expression were explained by the differential supercoiling in plasmid and chromosomal genes as a result of the integration event. In addition, the integration of the *rsk* element of *Salmonella typhiumurium* results in a differential state of supercoiling which is also responsible for phenotypic changes (Vandenbosch *et al.*, 1989).

It is therefore possible that thermostable protease expression or secretion may require an element of critical size, such as the *intC* or *vap* elements, to be integrated adjacent to both *askA* and *pnpA*, in order to provide the appropriate topology. The loss of either the *intC* or *vap* elements could alter the local topology and consequently may switch off thermostable protease expression or secretion. This possibility is supported by the observation that benign strains contain a truncated copy of the *intB* element or a truncated *intC* element in one of the two aforementioned positions. Depending on what environmental factors influence

the movement of these genetic elements may globally regulate virulence in *D. nodosus*. In the absence of a transformation system for *D. nodosus*, one could continue to try to select *D. nodosus* colonies that have randomly lost the *intC* element as described in Section 7.2.7.

7.2.11 Northern blot analysis of *glpA* and *pnpA* in *D. nodosus*

In an effort to determine if the genes upstream from the integrated genetic elements may be differentially expressed at the RNA level dependent upon which genetic element was integrated downstream of the element, northern blot analyses were undertaken. DNA fragments from *pnpA* and *glpA* genes were used to probe (Figure 7.16) total RNA from *D. nodosus* strains A198, C305, 1311 and 1311A. These strains were selected primarily because the elements integrated downstream of *askA* and *pnpA* in these strains are different (Figure 7.16), and strains 1311 and 1311A are almost isogenic.

In an initial northern blot experiment, *glpA* was used as a probe (Figure 7.17). In *D. nodosus* strains A198 and C305, the probe hybridised to a transcript of 234 nt, in strain 1311 a second transcript of 434 nt in addition to the 234 nt transcript was detected, whilst no hybridisation to total RNA was observed for strain 1311A. The experimentally-determined mRNA sizes correlate well with the mRNA sizes that were predicted from DNA sequence (Figure 7.18). If transcription were terminated at a near loop 1 or loop 2 the transcript size would be expected to be approximately 260 nt or 420 nt respectively.

However, it is unclear why the expression of transcripts should be different in strains C305 and 1311, since in both the *intC* element is adjacent to *glpA*. Similarly, it is unclear why the expression of the *glpA* transcript should be different for strains A198 and 1311 since both of these contain the same stem-loop structure that may potentially affect the expression of *askA*, *glpA* and $tRNAser_{GCU}$ genes (Figure 7.16). Thus, it was in doubt whether the results accurately reflected genuine differences between the strains of

D. nodosus analysed or not, and consequently, the experiment was repeated using a fresh RNA sample from each strain.

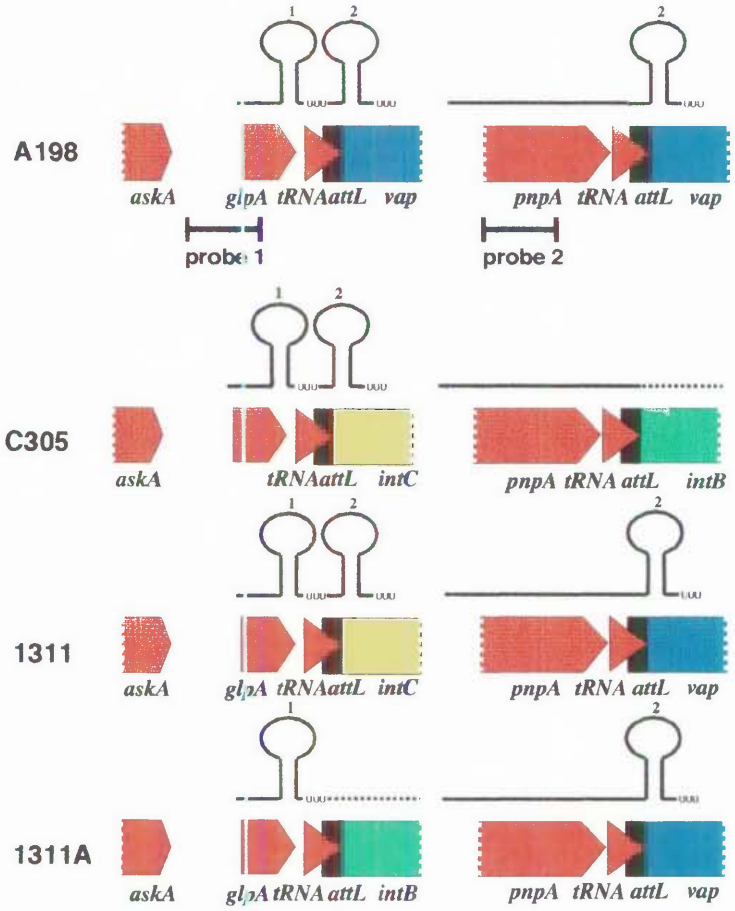

Figure 7.16: Schematic diagram of genetic elements integrated downstream of *askA* and *pnpA* in *D. nodosus* strains A198, C305, 1311 and 1311A. The integrated elements are only shown in part and are distinguished by different colours as follows: the *vap* element (blue), the *intB* element or part thereof (green) and the *intC* element (yellow). Stem-loop structures 1 and 2 are shown (not to scale). The genes are drawn to scale and the position of the *glpA* (probe 1) and *pnpA* (probe 2) probes are indicated.

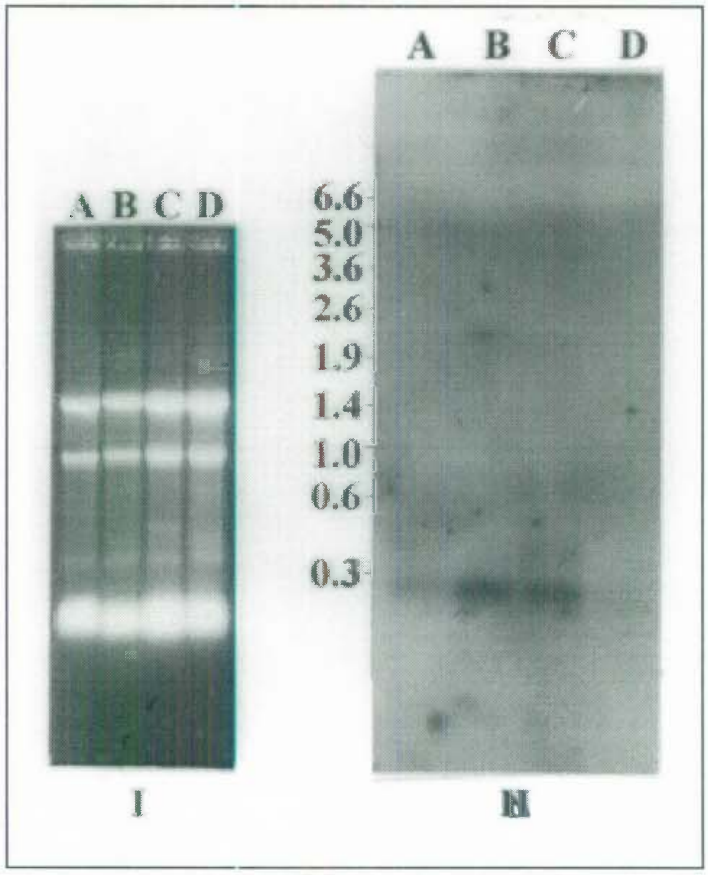

Figure 7.17: Northern blot analysis of *glpA* in *D. nodosus*. Total RNA was prepared from *D. nodosus* strain A198 (lanes A), C305 (lanes B), 1311 (lanes C) and 1311A (lanes D), and separated on a denaturing gel by electrophoresis in duplicate. One set of RNA samples was stained with EtBr as a control (panel I) whilst the other set was used for northern transfer. Results of the northern blot experiment are in kb shown (panel II). RNA marker sizes are indicated at the left of panel II.

Numerous attempts to reproduce this experiment were unsuccessful since they gave variable results. Often no message at all was detected for a strain, whilst at other times hybridisation was visible. The lack of reproducibility might be attributed to numerous factors. At least three of the strains analysed are not isogenic with respect to each other, and the growth characteristics of all four strains analysed are quite different (Section 3.2.9). Consequently, if *glpA* gene is expressed within a very narrow window, the expression of this gene might not be detected. Since the genes that are similar to *glpA*, including *rsmA* (Mukherjee *et al.*, 1996) and *csrA* (Liu *et al.*, 1995) are involved in repression of quorum-sensing molecules and starvation response respectively (Section 7.3) their expression may be

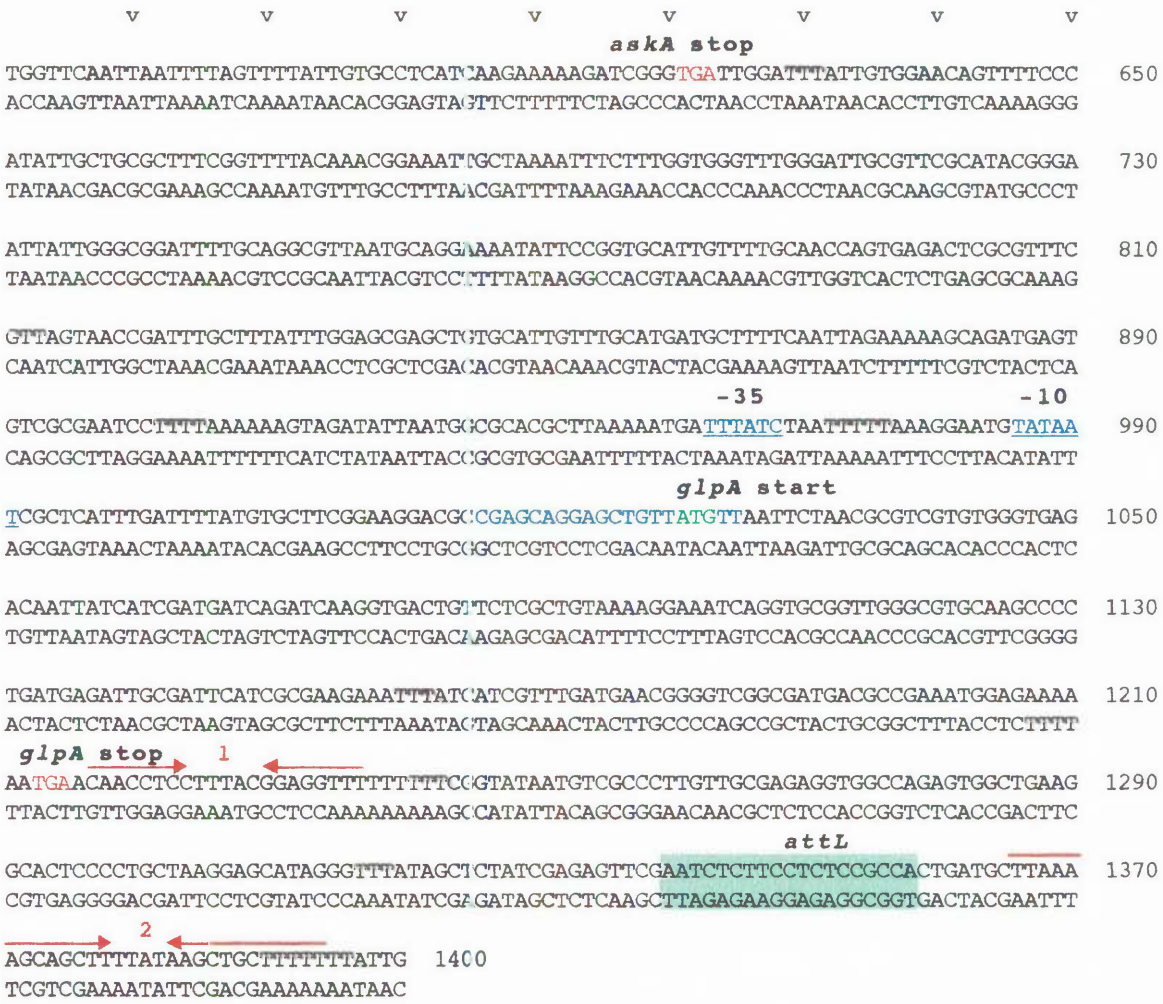

Figure 7.18: Sequence from position 571 to 1400 from λGB321, a C305 lambda clone that includes part of *askA*, and *glpA* and *tRNAser*$_{GCU}$ gene. Features highlighted above include: the *askA* and *glpA* stop codons (red type) are distinguished; the *glpA* start codon (green); the putative -35 and -10 promoter sequences for *glpA*; inverted repeats that form stem-loop structures 1 and 2 are (inverted red arrows); the *attL* in the 3'-end of the *tRNA* gene (green box).

influenced by the stage of growth. Furthermore, it would be expected that genes like *rsmA*, *csrA* and putatively *glpA*, which are involved in modulation of gene expression in response to environmental changes would be tightly regulated.

Although secondary structures inherent in the sequences being analysed in this work may have reduced the efficiency of transfer from gel to membrane, this possibility was addressed by using an alkaline transfer method (Reed, 1987). Unfortunately, this did not clarify results.

It is also possible that the RNA samples analysed may have been partially degraded. However, since bacterial mRNA is known to have a very short half-life, RNA was extracted from *D. nodosus* cells and then used immediately in northern blot experiments in order to minimise the extent of any degradation. In addition, problems with detection of transcripts were not experienced when the *vap* genes were analysed in numerous strains of *D. nodosus* (Chapter 4). It is known, however, that foreign messages from bacteriophages T7, T4, pX174, M13 and R17 have greater chemical stability than host mRNA due to a reduction in the number of sequences sensitive to nuclease action and unique sequences that are present at the 5'-end of certain messages (Alifano *et al.*, 1994). Perhaps since the *vap* regions have been acquired horizontally, the *vap* gene transcripts are more stable than transcripts from the genes of the *D. nodosus* host. *pnpA* sequences were also used as a probe in Northern blot analyses numerous times. No hybridisation to any of the four strains analysed was observed.

In an attempt to address the possibility that northern analysis is not sensitive enough to observe expression of *glpA* and *pnpA* in *D. nodosus* strains A198, C305, 1311 and 1311A, one could design a more sensitive RT-PCR experiment to test the hypothesis that these genes may be differentially expressed dependent upon which element is integrated immediately downstream.

7.3 Discussion

7.3.1 Arrangement and variation of *intC* element sequences in different strains of *D. nodosus*

In this work, seventeen strains of *D. nodosus* were analysed for the presence of the *intC* gene, and genes *orf242* and *orf171*, in an effort to determine whether these genes are clustered together as they are in the strain of origin (Figure 7.1), *D. nodosus* benign strain C3052, and therefore constitute part of a new genetic element. The results indicate that the genes *intC*, *orf242* and *orf171* comprise part of a new genetic element, called the *intC* element. This conclusion is supported by the observation that where *intC* is absent (Group 1 strains), both *orf242* and *orf171* are also absent (Section 7.2.1). In addition, in strain 1311 genes *intC*, *orf242* and *orf171* were lost concomitantly (Sections 7.2.2), which is also consistent with the idea that these sequences are collectively part of a mobile genetic element. Concomitant excision of the *intC* element sequences was also observed in *D. nodosus* strain AC3577 (Section 7.2.3).

Whether *vapG"H"* and the IS*1253* element are also associated with the *intC* element was also investigated. Results suggested that *vapG"H"* and the IS*1253* element are often associated with the *intC* element, though not always (Figure 7.2). It is also notable that in strain 1311, a copy of both *vapG"H"* and IS*1253* sequences were lost together with the *intC*, *orf242* and *orf171* genes (Section 7.2.2 & 7.2.4), and suggest that these genes do comprise part of the *intC* element, even if not immediately adjacent to *orf171* as in some strains.

Similarities between the *vap* element and the *intC* element have been reported previously (Bloomfield, 1997). The *intC* gene has a high degree of amino acid identity to the *intA* gene (Table 7.1) from *vap* region 1 of *D. nodosus* strain A198; *vapG"* has similarity to *vapG'* from *vap* region 3 in strain A198 (Table 7.1); and *vapH"* shares identity with *vapH*

from *vap* region 1 of strain A198. In addition, the *vap* plasmid (Billington *et al.*, 1996) also contains a copy of the IS*1253* element. Despite these similarities between the *vap* element and the *intC* element, genes *orf242* and *orf171* are also part of the *intC* element, and therefore, *intC* element sequences are not remnants of a *vap* element, but instead comprise part a different but related genetic element.

All strains analysed contain a single copy of the *intC* element sequences except strain H1215, which contains two incomplete copies of the *intC* element at an unknown site in the genome. Whether the integration of duplicate copies of the *intC* element is inhibited by *vapG"H"* which, like homologues *vapGH* from the *vap* region 1, may confer superinfection immunity (Section 4.2.3), is as yet unknown.

At this stage, the precise location of the alternative integration site/s for the incomplete copies of *intC* element have not been identified. The identification of the alternative sites for *intC* element integration could be useful because their identification, could potentially lead to the identification of the right-hand ends of other genetic elements, that are integrated in the *D. nodosus* chromosome. For example, the right-hand end of the *intB* element, of the *intC* element and both ends of the putative *intE* element (Section 7.2.9) have not yet been identified. All of these *D. nodosus* elements appear to recognise the same chromosomal target sites, which are found next to *askA* and *pnpA* respectively (Section 7.2.8) and are often found to be integrated in tandem. Therefore, the left-hand end of one element may be adjacent to the right-hand end of another element.

In addition, the integrase genes from native *D. nodosus* bacteriophage DinoHI (*intP*), and from the *intD* element (*intD*) have recently been identified and partially sequenced (B. Shaw & B. Cheetham, unpublished). Both of these integrase genes have a high degree of similarity to the *intA*, *intB* and *intC* family of *D. nodosus* CP4-like integrases, and so may recognise a similar or identical *att* site as these other *D. nodosus* elements. Thus, the

identification of these alternative sites of integration for the *intC* gene might also assist in the identification of other possible integration sites for these other genetic elements.

7.3.2 Loss of the *intC* element was observed in *D. nodosus* strains 1311, AC3577, 819 and 3138

The loss of *intC* element sequences from *D. nodosus* parent strain 1311 was observed to result in the juxtaposition of the *intB* element and *askA* in daughter strain 1311A (Section 7.2.2). This genomic rearrangement was also associated with the loss of the *intD* element sequences (Section 7.2.4), the loss of pDN1 (Section 3.2.6) and the rearrangement of a copy of the IS*1253* element (Section 7.2.4). Furthermore, these rearrangements are occurring in strain 1311 at a frequency high enough to be detected in both Southern blot and PCR experiments.

These rearrangements of *intC* and *intD* element sequences are also observed in *D. nodosus* intermediate strain AC3577. However, unlike strain 1311, in strain AC3577 the excision of the *intC* and element sequences does not seem to result in loss. Instead the *intC* elements sequences appear to be maintained elsewhere in the AC3577 genome, since faint bands corresponding to *intC*, *orf242* and *orf171* are detected in Southern hybridisations (Section 7.2.3). Furthermore, preliminary evidence suggests that the excised *intC* element sequences may be associated either intra- or extrachromosomally with *intD* element sequences in strain AC3577 (Section 7.2.4).

Rearrangement of *intC* element sequences appears to occur in *D. nodosus* strains 819 and 3138. This was unexpected since analysis of the *intC* elements in these strains suggests that the element is interrupted by a truncated copy of the *intB* element, containing the $intB_{VI}$ gene. However, since the *attR* of the *intB* element may be intact in these two strains, the *attL* of the *intC* element and the *attR* of the *intB* element may together result in the

excision of intervening sequences. It is interesting that in strain 819 excision of the *intD* element is also detectable in a small population of cells (Section 7.2.9). Since the putative *intE* element (Section 7.2.9) separates *pnpA* and the *intD* element in strain 3138, it is unknown whether the *intD* element is also excising in this strain. There is evidence to suggest that the *intE* element is also present in benign strains 1169, 2483, 1493 and C305 (Section 7.2.9). No rearrangements of the *intC* or *intD* element sequences were detected in any other strains of *D. nodosus* that were analysed.

7.3.3 Loss of the *intC* and *intD* elements may be concomitant

It is interesting that the strains (1311, AC3577, 819, 3138) in which the *intC* element sequences undergo rearrangement are those strains that also contain the *intD* element sequences (Section 6.2.2). The *intC* and *intD* elements are integrated in different positions of the *D. nodosus* chromosome in strains in which they are present and although both elements contain integrase genes that are related, they are not the same. In addition, the *intD* element can move independently of the *intC* element, although the two elements appear to also move in concert (Sections 7.2.2 and 7.2.4). These results therefore indicate that the *intC* and *intD* elements are different genetic elements. Since the *intC* and *intD* elements can potentially recognise the same integration sites in the *D. nodosus* genome, adjacent to *askA* and *pnpA* respectively, the *trans* action of excisionary proteins or the *trans* action of other regulatory proteins might mediate apparently concomitant excision.

Since the *intC* and the *intD* elements are integrated in different positions in strains 819 and 3138, and are not in tandem in other strains (1311, AC3577, B1006, D1172) in which they are both present, it is most likely that excision from the chromosome is mediated by site-specific recombination, catalysed by element-encoded proteins, such as integrase or excisionase, rather than mediated by general (homologous) recombination. The probability that two separate integrated *intC* and *intD* elements should be lost from different integration

sites, by homologous recombination, in the same *D. nodosus* cell seems remote. It seems more likely that the excision of these two elements is stimulated by the same factors in the host cell. Furthermore, if excision were mediated by homologous recombination, then the loss of the *intC* element and perhaps even the *vap* element should also be apparent in cells that do not contain the *intD* element, and there is no evidence in this regard. The apparent stability of the *vap* and *intB* elements even in strains that contain the *intD* element might be due to the of the divergence in sequences at the attachment sites that flank them, because of the lack of the excisionary machinery, or due to maintenance systems such as the *vapA/toxA* system carried by the *vap* element.

Interactions between genetic elements have been observed previously. For example *Bacteroides* conjugative transposons are known to stimulate the excision of unlinked integrated elements in *trans* due to the *trans*-action of regulatory proteins (Salyers, Shoemaker & Li, 1995a). One group of these unlinked integrated elements are designated nonreplicating *Bacteroides* units (NBUs), which share similarities to the *intD* and *vap* elements. These similarities include the following: *(i)* they are elements of approximately 10-12 kb; *(ii)* they contain *mob* genes and an *oriT* region; *(iii)* they excise and form covalently closed circular intermediates; and *(iv)* they integrate into the 3'-end of *tRNAleu* gene using a mechanism similar to that used by lambdoid phages (Salyers *et al.*, 1995b).

These NBUs are not only mobilised by conjugative transposons but also by conjugative plasmids belonging to the *E. coli* incompatibility group P (Salyers *et al.*, 1995a; Shoemaker *et al.*, 1986; Shoemaker *et al.*, 1993), and have been described as being at least as promiscuous as broad-host-range plasmids belonging to incompatibility groups Q and P. This is interesting because *D. nodosus* strain 1311 contains the small, IncQ-related plasmid pDN1 (Chapter 3). Plasmids belonging to IncQ group are non-self transmissible and may also be mobilised by conjugative plasmids belonging to the IncP group amongst others (Willetts & Crowther, 1981). It is unknown how pDN1 was acquired by the host cell, and it

is likely that a conjugative mechanism may have been involved. Other than the *intD* element, no conjugative element has been identified in *D. nodosus*. Perhaps the *D. nodosus* elements are mobilised by the same conjugative element that was involved in the horizontal transfer of pDN1 into *D. nodosus*.

In addition Tn*916*-type conjugative transposons stimulate the transposition of each other (Flannagan & Clewell, 1991), and evidence also indicates that *Bacteroides* conjugative transposons may also stimulate each others transposition (Li, Shoemaker & Salyers, 1995).

In *Staphlococcus aureus*, a pathogenicity island of 15.2 kb designated the *tst* element was identified. This element shares similarity with the *vap* element and has been described previously (Section 1.6.4.4). This element carries the *tst* gene which encodes toxic shock syndrome toxin 1 (TSST-1), integrates site-specifically into the *S. aureus* genome, and is excised and circularised by the staphylococcal phages ϕ13 and 80α (Lindsay *et al.*, 1998), and thus is induced to excise by the same stimuli. No *xis*-encoded excisionase has been identified in the *tst* element, and it has been proposed that the phage encoded Xis protein catalyses the excision of the *tst* element. Similarly, no putative excisionase has been identified within well characterised *D. nodosus* elements including the *vap*, *intB* and *intC* elements. Excisionase functions may be encoded by the *intD* element of *D. nodosus* since it is only in the presence of this element that the *intC* element excises. Such functions may also be present in the putative *intE* element.

It is unlikely that in *D. nodosus* the *intC* or *intD* elements are induced in a similar way by bacteriophage DinoHI (Section 1.6.9) since the elements excise in strains that do not contain the integrase region of DinoHI. However, since the integrases of these *D. nodosus* elements are most closely related to the integrase genes of integrated prophages rather than those carried by other elements such as integrons (Hall *et al.*, 1995) and conjugative

transposons, it would be interesting to see if these *D. nodosus* elements are also induced by ultraviolet light.

Staphylococcal transducing phages ϕ11 and 80α have been demonstrated to encapsidate plasmids as tandem multimers, comprising of as many copies of the plasmid as would be required to fill a phage headful (Novick, Edelman & Lofdahl, 1986), and thereby transduce them at high frequency. After transfer, the plasmid multimers are resolved into monomers in the recipient cell. The authors also propose that the aforementioned *tst* element may also be packaged as part of a 45 kb *tst* element-phage recombinant (Lindsay *et al.*, 1998).

D. nodosus elements could potentially be transferred in a similar manner, though their transfer could be mediated by conjugation rather than by transduction. If these elements excise cooperatively, and one of these elements carries determinants necessary for conjugative transfer a heterogenous multimer of different elements could be transferred collectively. Since the *vap*, *intB*, *intC*, *intD* and *intE* elements recognise the same target site for integration into the *D. nodosus* chromosome they also have the potential to integrate into each other. In this work evidence suggests that at least the *intC* and *intD* elements are able to excise concomitantly, and in at least one strain may form a multimer. Furthermore, a cooperative mechanism for horizontal transfer might explain why so many related genetic elements appear to integrate in *askA* and *prpA* positions of the *D. nodosus* chromosome.

A cooperative transfer mechanism would enable the recipient cells to acquire numerous traits in one step. Dependent upon the requirements of the host cell, certain traits encoded by one element could be lost, whilst others favourable traits encoded by different elements or cassettes are maintained. Furthermore, the site-specific integration into two sites in the *D. nodosus* genome reduces the probability of any deleterious effects that might occur if integrations were more dispersed in the host chromosome.

7.3.4 Concomitant loss of pDN1 with the *intC* and *intD* elements in *D. nodosus* strain 1311

It is unclear why the loss of pDN1 should be associated with the loss of the *intC* or *intD* elements. Although all three events occurred at the same time in *D. nodosus* strain 1311 to generate *D. nodosus* strain 1311A, results from selection studies (Section 7.2.5) suggest that pDN1 loss can and does occur independently of loss of the *intC* element. Whether pDN1 can be lost without loss of the *intD* element is currently unknown, and may be revealed by Southern blot or PCR analyses of pDN1$^-$ colonies selected in these experiments.

In strain 1311 the loss of pDN1 was associated with the loss of the *intD* and *intC* elements (Sections 7.2.2 and 7.2.4), and may be related to a reduction in the efficiency of multimer resolution of native plasmid pDN1. Multiple-copy plasmids such as pDN1 are distributed randomly to daughter cells upon cell division. In order to maintain a large number of independently segregating units at cell division, a system is required for the efficient resolution of plasmid multimers. These systems usually operate *via* a site-specific recombination mechanism.

Recombination at a *cis*-acting plasmid-encoded sequence is catalysed by site-specific recombinases that all belong to the integrase family of recombinases (Dorman, 1994). In plasmid ColE1, recombination at the *cer* site is mediated by host-encoded XerC, XerD and other accessory proteins (Summers & Sherratt, 1984). If multimer resolution of plasmid pDN1 was similarly dependent upon proteins encoded on the *D. nodosus* chromosome the loss of the *intC* and/or *intD* element encoded integrase genes may have subsequently resulted in a reduced plasmid-multimer resolution efficiency, consequently reducing segregation efficiency, eventually leading to loss of pDN1. It is therefore very interesting that the similarity between XerC and IntB is not just confined to the five key amino acid residues (Arg^{240}, His^{326}, Arg^{329}, Tyr^{363}) that define the lambda family of site-specific recombinases

(Abremski & Hoess, 1992; Argos *et al.*, 1986), but is much more extensive (Table 5.1).

Alternatively, the loss of pDN1 may subsequently destabilise other genetic elements such as the *intC* element. Such an interaction has been observed in the case of RSF1010 derivative, pMMB67. pMMB67 stabilised a derivative of rolling circle plasmid pVT736, in which an open reading frame responsible for stabilisation (*orf2*, which has 80.6% identity in 93aa to *vapD*) of the plasmid was mutated. The mechanism by which pMMB67 stabilised the mutant derivative of pVT736 is unknown and no explanation was offered (Galli & LeBlanc, 1997). However, since plasmid RSF1010 is closely related to native *D. nodosus* plasmid pDN1 (Section 3.2.4), pDN1 may confer a stability on the *intC* element or the *intD* element in a similar manner.

7.3.5 Loss of the *intC* and *intD* elements from strain 1311 generated a putative benign strain, 1311A

Following the loss of the *intC* element and *intD* elements from *D. nodosus* strain 1311A, the colonies appeared smaller and exhibited reduced exoenzyme activity when compared to parent virulent strain 1311 (Sections 7.2.6 and 7.2.7). These phenotypic differences are characteristic of benign strains (Depiazzi & Richards, 1985; Stewart, 1989) of *D. nodosus*. Protease thermostability assays revealed that *D. nodosus* strain 1311A did not produce thermostable proteases and consequently is likely to be avirulent. It is notable that there was no change in the site of integration of the *vrl* regions of 1311 compared to strain 1311A, which appear to be integrated in the same position as virulent reference strain A198, adjacent to the *ssrA* gene (Haring *et al.*, 1995).

The deletion of regions altering or abolishing the expression of virulence-associated factors have been reported in numerous bacterial pathogens (Ott, 1993). It has been proposed previously (Groisman & Ochman, 1996; Mecsas & Strauss, 1996), that the

instability of regions containing virulence determinants might provide an adaptive advantage at certain stages of infection. For example, *E. coli* pathogenicity islands (Pai), PaiI and PaiII, are spontaneously deleted from the host chromosome at a frequency of 10^{-4} (Blum *et al.*, 1994). It has been proposed that the deletion of pathogenicity islands may represent a regulatory mechanism by which the microorganism can either circumvent the host immune response or allow rapid microbial replication by reducing genome size (Groisman & Ochman, 1996; Hacker, 1996; Mecsas & Strauss, 1996) in the absence of factors that favour the expression virulence determinants. However, it has also been reported that both pathogenic and non-pathogenic *E. coli* strains carry immunogenic surface antigens, and, furthermore, that under laboratory conditions *E. coli* growth rates are not related to genome size (Groisman & Ochman, 1996). It is unknown what environmental stimuli might promote the excision of the *intC* and *intD* elements in *D. nodosus*.

Although it has been noted that the proteases secreted by benign strains exhibit different electrophoretic mobilities, are more thermolabile and have decreased elastase activity (Green, 1985; Kortt *et al.*, 1982; Stewart, 1979; Thomas, 1964), the sequencing of the protease genes from both benign and virulent isolates of *D. nodosus* showed that the protease genes were almost identical in benign and virulent isolates (Section 1.2.2.ii). Which of the proteases are thermostable, and which of the proteases are responsible for the virulence-associated thermostability, has not been described in the literature. It would therefore be useful to clarify these points by separating the exoenzymes from virulent strain 1311 and putatively avirulent daughter strain 1311A on a non-denaturing protein gel before and after heat treatment. Protease activity of specific bands could be determined by providing a substrate (such as elastin) as a gel overlay. After incubation at 37°C protease activity could be visualised as clearing in the gel overlay. This would then allow a certain protein band size to be associated with virulence. The molecular weight of these bands could then be compared to those predicted for the protease genes, and consequently the gene/s associated with virulence determined. This may be useful for both diagnosis and further

understanding of the mechanisms by which virulence determinants such as protease are regulated in *D. nodosus*.

Bacterial pathogenesis is typically multifactorial, and consequently one or more virulence factors are required to work together to result in disease. The absence of any one of these virulence factors may or may not render a strain benign. Hence, in bacterial pathogens a small number of strains within a species cause disease (Falkow, 1981). The loss of thermostable protease expression in strain 1311A, suggests that both benign and virulent strains of *D. nodosus* have the potential to produce these virulence-associated proteases, and that regulation of expression altered when the *intC* element was lost from parent strain 1311 generating daughter strain 1311A.

Although many benign strains also contain the *intC* element, in eight of these benign strains (strains C305, 3138, 819, 1469, 2483, 1169, H1204, AC390) the *intC* element is truncated and thus incomplete. The integrity of the *intC* element in the only other benign strain studied in this work (strain 1493) is unknown (Figure 7.2).

The loss of a system required for the export of extracellular thermostable proteases if encoded by either the *intC* or *intD* elements could account for such a phenotypic change. Hence, the similarity between *intD* element encoded *tra* proteins to the Vir proteins which function in the export of T-DNA from *A. tumefaciens* to plant cells is of interest.

It has been suggested that the plant pathogen *A. tumefaciens* has adapted an ancestral conjugative transfer system for the transfer of its oncogenic T-DNA to plant cells. This hypothesis is supported by: *(i)* similarities in DNA processing; *(ii)* the ability of the *vir* encoded T-DNA transfer proteins to mobilise plasmids to plant and bacterial cells; and *(iii)* the presence of genes found in the *vir* operon of *Agrobacterium tumefaciens* Ti plasmids that are homologous and collinear with genes required for the conjugative transfer (*tra*) from the IncN conjugative plasmid pKM101 (Pohlman, Genetti & Winnans, 1994). It has also

been proposed that the animal pathogen, *Bordetella pertussis* has also adapted a homologous conjugal transfer system to mediate the export of pertussis toxin (Weiss, Johnson & Burns, 1993).

Thus, the *intD* element encoded conjugative region may also have the potential to be involved in the export of virulence factors such as proteases in *D. nodosus*. However, the virulent reference strain A198, does not contain a copy of at least the right hand end of the *intD* element, and thus probably not the left end of the *intD* element, but still secretes thermostable protease. Thus, it is unlikely that the loss of the *intD* element is responsible for the observed loss of protease thermostability in strain 1311A. However, it is not impossible that strain A198 and other virulent strains (G1220, H1215) that do not contain the *intD* element, have some alternative export system that is utilised for the secretion of extracellular proteases; and similarly it is also possible that benign strains 819 and 3138 that do contain the *intD* element sequences contain some other mutation that affects the expression of virulence-associated proteases.

7.3.6 The site-specific integration of mobile genetic elements may modulate virulence in *D. nodosus*

It is interesting that based upon the similarity to genes of known function, the *vap* genes *A-D*, *vapG*, *vapH* and *toxA* are likely to be involved with plasmid maintenance and the control of cell division or both (Section 1.6.4). The only gene from within the *vap* element to which no putative function has yet been assigned is *vapE*. The genes from within the *vap* regions do not show similarity to known virulence factors of *D. nodosus* or to virulence factors from other organisms. Though it has been speculated that since the *vap* genes have similarity to genes that interact with the cell division machinery, this interaction may affect virulence (Cheetham, Whittle & Katz, 1998), there is no evidence for a direct role of the *vap* gene products in virulence. Similarly, from genes so far identified in both the

intC (Table 7.1) and *intD* elements (Section 6.2.1, Table 6.1), and from analysis of strains in which these elements are present, there is no evidence that genes encoded by these elements have a direct role in virulence. Although, if the *intD* element it is able to catalyse the excision and mobilisation (Section 6.3 of elements like the *vap* and *intC* elements in *D. nodosus* it may be involved in dissemination of virulence-associated elements in *D. nodosus*.

However, in this work, the loss of protease thermostability (Section 7.2.6), a virulence factor in *D. nodosus*, was clearly associated with the spontaneous loss of the *intC* and *intD* element sequences in strain 1311A. Since the elements themselves do not seem to be involved directly in virulence, it is possible that the excision of these genetic elements may have altered the expression of genes located adjacent to their integration sites, and these alterations affected the expression of virulence determinants.

The comparison of the sequences upstream of the *vap*, *intC* and *intB* elements led to the identification of a putative antisense promoter and stem-loop structure (loop 2, Figure 7.15) that is present when the *vap* and *intC* elements are integrated, but is absent in the corresponding region of the *intB* element (Section 7.2.10). These observations support the hypothesis that virulence in *D. nodosus* might be modulated by the site-specific integration of genetic elements.

The mechanisms by which the integration of genetic elements might reversibly modulate virulence in *D. nodosus* were discussed in Section 7.2.10, involved the posttranscriptional regulation of genes adjacent to the integrase genes of the integrated elements by an antisense binding mechanism, by interfering with the processing of transcripts from upstream genes or by a retroregulation mechanism, or steric hindrance by certain regulatory proteins. The genes that would be posttranscriptionally affected included the *tRNAser* genes, *glpA* and *pnpA*.

7.3.6.1 *tRNAser* genes and virulence

tRNA genes have been demonstrated to play a role in virulence previously. Mutations in genes involved in post-transcriptional modification of *tRNA* genes (Crowley *et al.*, 1997; Durand *et al.*, 1997; Gray *et al.*, 1992; Hromockyj *et al.*, 1992), or mutation of *tRNA* genes (Fernandez-Moreno *et al.*, 1991) and the disruption of *tRNA* genes by the excision of pathogenicity islands (Blum *et al.*, 1994; Ritter *et al.*, 1995; Ritter *et al.*, 1997), have been demonstrated to alter the expression of virulence-associated genes. The later provides a good example of how the integration or excision of genetic elements might inactivate or alter the expression of chromosomal loci.

In uropathogenic *E. coli* 536, the disruption of the *selC* gene encoding selenocysteine-specific tRNAsec impairs anaerobic growth (Blum *et al.*, 1994; Ritter *et al.*, 1995). Similarly, the *leuX* gene encoding tRNA5-leu is necessary for type 1 fimbriae production, flagellation and motility, enterobactin production, and serum resistance, and is necessary for full *in vivo* virulence (Blum *et al.*, 1994; Ritter *et al.*, 1995). The failure to modify position 37 of a subset of *tRNA* genes results in the reduced expression of virulence-related genes *ipaB*, *ipaC*, *ipaD*, *virG* and *virF* in *S. flexneri* 2a (Durand *et al.*, 1997; Hromockyj *et al.*, 1992). Similarly, lack of modification of A-37 in UNN tRNA species reduces the translation efficiency of the *vir* regulon in *A. tumefaciens*, consequently reducing virulence (Gray *et al.*, 1992). tRNAs have also been shown to regulate other prokaryotic genes. In *Streptomyces coelicolor* A3, a mutation in tRNAleuZ encoded by *bldA* results in a mutant which cannot produce aerial hyphae, sporulate or secrete four different antibiotics (Fernandez-Moreno *et al.*, 1991). Similarly, a defective tRNAthr in *Clostridium acetobutylicum* detrimentally affects genes involved in formation of the solvents acetone and butanol (Sauer & Durre, 1992).

In *E. coli*, there are five *tRNA* serine genes designated *serW* ($tRNAser_{GGA}$), *serT* ($tRNAser_{UGA}$), *serX* ($tRNAser_{GGA}$), *serU* ($tRNAser_{CGA}$), and *serV* ($tRNAser_{GCU}$) (Blattner *et*

al., 1997) and six serine codons (Grosjean *et al.*, 1985) (Table 7.6). The $tRNAser_{UCG}$ is not an essential tRNA, since $tRNAser_{UGA}$ can recognise the same codon. The major serine codons in highly expressed genes of *E. coli* are UCU and CCU (Grosjean *et al.*, 1985; Grosjean & Fiers, 1982). If the codon usage in *D. nodosus* were similar, possible reduced efficiency of transcription of $tRNAser_{GGA}$ could effect the levels of both major and minor tRNAs species and thus could significantly affect the translation of certain proteins. The disruption of two of the three key tRNA molecules ($tRNAser_{GGA}$ and $tRNAser_{GCU}$) would be presumed to have an even greater effect on gene expression than one alone.

Table 7.6: Codon and anticodon usage in of serine tRNA in *E. coli*

Relative abundance (%)	Anticodon + 3'-planning base	Expected recognised codons	Freq. High	Freq. Low	Gene
40	**G**GA - A	UC**C**	16	9	*serX*
5	**G**GA - A	UC**U**	18	7	*serW*
19	**V**GA - ms^2i^6A	UC**U** UC**A** UC**G**	1	7	*serT*
5	**C**GA - ms^2i^6A	UC**G**	2	13	*serU*
13	**G**CU -t^6A	AG**C**	9	12	*serV*
19	**G**CU - A	AG**U**	2	11	

Heavy letters correspond to 'wobble' base of the anticodon or the corresponding third position of the codon. The frequency of each codon is expressed per thousand codons for highly expressed genes (high) and lowly expressed genes (low) in *E. coli*. V is uridine-5-oxyacetic acid; ms^2i^6A is *N*-6-(2-isopentenyl)2-methylthioadenosine; t^6A is *N*-[9-(β-D-ribofuranosyl)purine-6-carbamoyl]threonine. Table modified from previous citation (Grosjean *et al.*, 1985) using the sequence of complete genome (Blattner *et al.*, 1997).

Consequently, if tRNAser molecules are involved in virulence gene expression or regulation, the absence of correct processing when the stem-loop structure is absent, or if processing was sterically hindered by DNA binding proteins, could result in the attenuation

of virulence in *D. nodosus*. Similarly, the coupled expression of the *tRNA* gene and the putative antisense promoter would also reduce the amount of functional tRNAser molecules, which could affect the translation efficiency of virulence-associated transcripts, especially if the levels of a minor tRNA species was affected. In the absence of a transformation system for *D. nodosus*, it difficult to test these possibilities directly.

7.3.6.2 *glpA* and virulence

There is the potential that alterations in the levels of *glpA* expression might also affect virulence. GlpA (64 aa) has very high similarity (74% over the first 50 aa) to the RsmA (61 aa) protein from phytopathogen *E. carotovora* (Cui *et al.*, 1995; Mukherjee *et al.*, 1996) which acts as a global repressor of virulence. RsmA has 98% aa identity to CsrA which is a negative regulator of carbon storage and acts a global repressor of the starvation response in *E. coli* B (Liu *et al.*, 1995). When *E. coli* cells enter stationary phase due to nutrient limitation, cell division is inhibited, cell size increases and glycogen biosynthesis is inactivated (Romeo, Gonog & Brun-Zinkernagel, 1993). Since complementation studies have shown that RsmA is able to suppress glycogen synthesis in *E. coli*, RsmA and CsrA are also functionally similar (Cui *et al.*, 1995).

The high degree of similarity between *glpA* from *D. nodosus* and RsmA-like genes, suggests that GlpA and RsmA may have similar functions in both organisms. Thus, the mechanism by which RsmA globally regulates the production of secondary metabolites such as extracellular enzymes, polysaccharides and antibiotics in *Erwinia* sp. is also interesting, since this may give some insight into how virulence might be regulated in *D. nodosus*.

In *E. carotovora*, genes producing secondary metabolites are activated during late exponential phase or early-stationary phase when the bacteria are at high cell density. Production of such secondary metabolites is triggered by nutrient limitation or starvation, or

both, that is experienced by bacteria at high cell density. It has been demonstrated that in *E. carotovora*, production of such secondary metabolites involves a common regulator. The common regulator is the quorum-sensing molecule *N*-(3-Oxohexanoyl)-L-homoserine lactone (HSL), which has been shown to be required for the production of extracellular enzymes (Jones *et al.*, 1993), an antibiotic (Bainton *et al.*, 1992), flagellum formation and motility, and pathogenicity in general (Mukherjee *et al.*, 1996; Pirhonen *et al.*, 1993) in soft rotting *Erwinia* species. HSL production in *Erwinia* sp. is repressed by RsmA, which is thought to bind the *luxI* homologue, *hslI* (encoding homoserine lactone synthase) mRNA accelerating its decay (Mukherjee *et al.*, 1996), thereby repressing virulence gene expression.

It is interesting that an increasing range of target genes, frequently of medical, economic and of scientific importance, appear to be similarly regulated by quorum-sensing systems. These quorum-sensing molecules *N*-acylhomoserine lactones (AHL) accumulate at high cell density (Robson *et al.*, 1997), and interact with a receptor protein belonging to the LuxR family of positive transcriptional activators, and so are able to regulate diverse physiological activities fundamental to the interaction of bacteria with each other and their environment. Sometimes these interactions are symbiotic (Meighen, 1994; Schripsema *et al.*, 1996) but they are often pathogenic (Bainton *et al.*, 1992; Beck von Bodman & Farrand, 1995; Jones *et al.*, 1993; Latifi *et al.*, 1995; Mukherjee *et al.*, 1996; Pearson, Pesci & Iglewski, 1997). Examples of such activities include bioluminescence (Meighen, 1994), swarming (Eberl *et al.*, 1996), plasmid conjugal transfer (Fuqua, Winans & Greenberg, 1996; Stevens, Dolan & Greenberg, 1994; Zhang *et al.*, 1993), and the production of exoenzyme virulence determinants in human and plant pathogens (Cui *et al.*, 1995; Cui *et al.*, 1996; Jones *et al.*, 1993; Passador *et al.*, 1996) (Table 7.7). For pathogens, quorum-sensing systems are thought to provide them with a means of coordinating the expression of virulence factors at high cell density when they are best able to overcome the attack of the

Table 7.7: Bacterial *N*-acyl-L-homoserine lactones and their associated phenotypes

Organism	Major signalling molecule[a]	Regulatory proteins[b]	Phenotypes	Reference
Aeromonas hydrophila	*N*-β-(hydroxybutyryl)- L-homoserine lactone (HBHL)	AhyI/AhyR	unknown	(Swift *et al.*, 1995)
Agrobacterium tumefaciens	*N*-(3-oxyoctanoyl)-L-homoserine lactone (ODHL)	TraI/TraR - TraM	Conjugal transfer of the Ti plasmid	(Zhang & Brooker, 1993)
Enterobacter agglomerans	*N*-(3-oxyhexanoyl)-L-homoserine lactone (OHHL)	EagI/EagR	unknown	(Swift *et al.*, 1993)
Erwinia carotovora	*N*-(3-oxyhexanoyl)-L-homoserine lactone (OHHL)	ExpI/ExpR HslI/?	(5R) Carbapen-2-em-3-carboxlic-acid antibiotic; virulence factors including protease, cellulases, pectinases and exopolysaccharide. Flagellum synthesis and motility.	(Bainton *et al.*, 1992; Jones *et al.*, 1993; Mukherjee *et al.*, 1996)
Erwinia stewartii	*N*-(3-oxyhexanoyl)-L-homoserine lactone (OHHL)	CarI/CarR EsaI/EsaR	virulence factors including protease, cellulases, pectinases and exopolysaccharide	(Beck von Bodman & Farrand, 1995)
Escherichia coli	unknown	?/SdiA[c]	*ftsQAZ*, cell division	(Garcialara, Shang & Rothfield, 1996)
Pseudomonas aeruginosa	*N*-butanoyl-L-homoserine lactone (BHL) &	RhlI/RhlR	Virulence factors: including rhamnolipid, alkaline protease, cyanide, elastase, haemolysin, pyocyanin.	(Latifi *et al.*, 1995)
	N-(3-oxydecanoyl)-L-homoserine lactone (OdDHL)	LasI/LasR	Virulence factors: alkaline protease, elastase, exotoxin A, haemolysin, neuramidase	(Pearson, Pesci & Iglewski, 1997)
Pseudomonas aureofaciens	unknown	PhzI/PhzR	Phenazine antibiotic	(Wood & Pierson, 1996)
Rhizobium leguminosarum	*N*-(3-hydroxy-7cis-tetradecanoyl)-L-homoserine lactone (HtDeHL)	?/RhiR	Expression of rhizosphere genes	(Schripsema *et al.*, 1996)
Serratia liquefaciens	*N*-butanoyl-L-homoserine lactone (BHL)	SwrI/?	Swarming: intitiation of cell differentiation to produce hyperflagellated extremely motile swarmer cells	(Eberl *et al.*, 1996)
Vibrio anguillarum	*N*-(3-oxodecanoyl)-L-homoserine lactone (ODHL)	?	unknown	(Milton *et al.*, 1997)
Vibrio fischeri	*N*-(3-oxyhexanoyl)-L-homoserine lactone (OHHL)	LuxI/LuxR	bioluminescence	(Meighen, 1994)
	N-(3-oxyoctanoyl)-L-homoserine lactone (ODHL)	AinS/AinR	bioluminescence	
Vibrio harveyi	*N*-(3-hydroxybutanoyl)-L-homoserine lactone (HBHL)	LuxM/LuxN –LuxO-LuxR	bioluminescence	(Cao, Wei & Meighen, 1995)
Yersinia enterocolitica	*N*-hexanoyl-L-homoserine lactone (HHL) & *N*-(3-oxyhexanoyl)-L-homoserine lactone (OHHL)	YenI/YenR	unknown	(Throup *et al.*, 1995)

a: only the primary autoinducing molecules synthesised by each organism and predicted to be involved in gene regulation are shown;

b: indicates only *luxI/luxR* type regulators; on average there is no higher than 18-25% identity between LuxR homologues and 28-35% identity for LuxI homologues except for those founds within the same Genus. Although many genes that are homologues of luxI/luxR are genetically linked, several homolgoues are transcribed convergenetly and are not linked to the genes that they regulate (Fuqua, Winans & Greenberg, 1996);

c: non-acylated homoserine lactone has been implicated in *E. coli* starvation response.

host immune system and thereby establish stable infection (Manefield, 1998; Robson *et al.*, 1997).

Recent investigations have identified a group of natural compounds called halogenated furanones (Givskov *et al.*, 1996) that act as autoinducer antagonists by competing with AHLs for the AHL receptor proteins. Consequently, the possibility that these antagonists might be utilised to block intercellular communications and thus pathogenic interactions, is the subject of current quorum-sensing autoinducer research. The alloinhibition of swarming motility in *Serratia liquifaciens* (Eberl *et al.*, 1996) and of bioluminescence in *Vibrio fischeri* and *Vibrio harveyi* (Fuqua, Winnans & Greenberg, 1994) by halogenated furanones has recently been demonstrated.

Even if quorum-sensing molecules are involved in virulence-gene expression in *D. nodosus* it is uncertain whether HSLs antagonists, such as halogenated furanones, would be useful for treatment, since such compounds would not kill or eradicate the pathogen, but simply temporarily block communication. Even if irreversible inhibitors were developed, these inhibitors would need to be applied regularly, in the same way footbaths are used currently (Section 1.4). It is unknown whether these halogenated furanones would be less toxic and thus more user and environmentally friendly than chemicals currently used for footbathing such as formaldehyde, but may be worthy of further investigation.

Although it is currently unknown whether HSLs are involved in virulence gene regulation in *D. nodosus*, it is probable that they are, given the high degree of similarity between RsmA and GlpA, and given that such quorum-sensing systems are being identified increasingly in Eubacteria generally (Fuqua *et al.*, 1996). Recent database searches revealed that numerous other bacteria have RsmA/CsrA homologues including *Haemophilus influenzae* (Fleischmann *et al.*, 1995), *Helicobacter pylori* (Tomb, 1997), *Bacillus subtilis* (Soldo *et al.*, 1996), *Borrelia burgdorferi* (Fraser *et al.*, 1997), *Treponema pallidum* (Fraser

et al., 1998), *Serratia marcescens* (Ang *et al.*, 1998) and *Pseudomonas aeruginosa* (Hindle *et al.*, 1998). The observation that homologues are so widespread in Eubacteria suggests that these global regulatory genes are likely to play a very important role in prokaryotic gene regulation in general, and perhaps also in virulence gene regulation. Thus, investigations to determine whether quorum-sensing molecules are involved in virulence gene expression in *D. nodosus* is an area worthy of further research.

In order to determine whether AHLs are involved in virulence-gene regulation in *D. nodosus*, AHL assays should be performed on stationary phase culture supernatants from benign and virulent strains of *D. nodosus* using the AHL reporter systems described elsewhere (McClean *et al.*, 1997; Milton *et al.*, 1997). Alternatively AHL derivatives or AHL antagonists such as halogenated furanones could be added to benign or virulent *D. nodosus* cultures and the effect on protease thermostability subsequently determined.

The idea that AHLs may be involved in virulence-gene regulation is also supported by studies of opportunistic pathogen, *P. aeruginosa*. *P. aeruginosa* produces elastases that have a role in virulence, and in this organism elastase production is controlled by the quorum sensing molecule, *N*-(3-oxodecanoyl)homoserine lactone (Pearson *et al.*, 1997). The recent identification of an RsmA homologue (Hindle *et al.*, 1998), suggests that in *P. aeruginosa*, as in *E. carotovora*, HSL production and therefore virulence gene expression may also be regulated by an RsmA-like gene.

This is interesting because *D. nodosus* is also known to produce elastases, and this elastase activity is greater in virulent strains than in benign strains of *D. nodosus*(Stewart, 1979), and is associated with virulence. The recent identification of *glpA*, and the high similarity between GlpA and other RsmA- or CsrA-like genes suggests that *glpA* may act as a global regulator of virulence, and the recent loss of an integrated genetic element that may affect *glpA* expression resulted in loss of protease thermostability, a virulence factor in

D. nodosus. These observations support the hypothesis that GlpA and therefore AHLs may be involved in virulence gene expression and regulation in *D. nodosus*.

The parallels between the phenotype observed in *E. carotovora* with RsmA expression and the morphological differences observed in the putatively avirulent strain 1311A is also of interest. In *E. carotovora*, RsmA represses protease production, and reduces motility which consequently reduces colony size (Mukherjee *et al.*, 1996). Similarly, loss of the *intC* element from the position adjacent to *askA* and *glpA* in strain 1311A, resulted in colonies that were significantly smaller in size and exhibited reduced exoenzyme secretion compared to parent strain 1311. *D. nodosus* benign strains do not secrete thermostable proteases and virulent strains have greater twitching motility than benign strains, and so typically have larger colonies than benign strains (Depiazzi *et al.*, 1990; Depiazzi & Richards, 1979; Stewart, 1989; Stewart & Egerton, 1979).

Furthermore, this regulation may involve quorum sensing molecules like HSL and its derivatives, which in turn may regulate the expression of virulence determinants in *D. nodosus*. Consequently, if the integration of a genetic element reduced the production of the putative global repressor of virulence, GlpA, one might expect to observe an increase in the expression of virulence-associated genes.

Also, in *D. nodosus*, *glpA* is located between an aspartokinase gene (*askA*) and a *tRNA*-serine gene with the anticodon GCU, and is similar to the *E. coli* tRNA *serV* gene (Cheetham *et al.*, 1995). It is therefore interesting that in *P. aeruginosa*, *csrA* is similarly located between an aspartokinase gene (*lysC*) and a disrupted copy of a *tRNA* serine gene (*serV*-GCU). In *E. coli*, *csrA* is also immediately adjacent to the *serV* gene (Romeo *et al.*, 1993). In both *E. carotovora* and *S. marcescens* the sequences flanking the *rsmA* genes have not yet been determined. It is unknown whether the tRNA secondary structure is important for the expression or processing of *glpA*.

It is also interesting that the *intD* element contains putative *tra* genes (Shaw and Cheetham, unpublished) that share amino acid identity to those of the *A. tumfaciens* Ti plasmid which are activated by a AHL quorum-sensing system *via* the transcriptional activator TraR and the autoinducer synthase TraI (Winans *et al.*, 1994). Since it is also likely that in *D. nodosus* that the expression of virulence-associated extracellular enzymes in *D. nodosus* is likely to be controlled by acyl homoserine lactone (AHL) quorum-sensing molecules (Cheetham *et al.*, 1998). In addition, results discussed in Chapter 8 suggest that the movement of integrated genetic elements may modulate virulence in *D. nodosus*. Consequently, if AHLs interact with genes present in the *intD* element, as they do with the related *vir* genes in the Ti plasmid of *A. tumefaciens* (Winans *et al.*, 1994), then AHLs could also modulate virulence by influencing the movement of genetic elements in *D. nodosus*.

It is interesting that both GlpA and PnpA contain a KH domain (Section 7.2.10). PnpA is one of the two principal enzymes involved in RNA degradation by and attaching to the 3'-end of an RNA molecule and progressively degrading it to mononucleotides (Gibson, Thompson & Heringa, 1993). GlpA like its homologues may also control gene expression by causing rapid decay of mRNA. The *vap* and *intC* elements are integrated next to two genes whose products may be involved in gene regulation by accelerating the decay of target mRNAs. If the integration of these *D. nodosus* elements up or down regulates the expression of these genes at the transcriptional or translational level, this could up or down regulate the expression of genes that they regulate.

7.3.7 Implications and future investigations of virulence mechanisms in *D. nodosus*

Many of the experiments done to test the pleiotropic effects of RsmA in *E. carotovora* involved the transformation of a soft-rotting strain of *E. carotovora* with an

rsmA$^+$ plasmid. However, it is not possible to perform similar experiments in *D. nodosus* in the absence of a transformation system. However, one could transform a *glpA*$^+$ plasmid into a soft rotting strain of *E. carotovora* and so determine whether GlpA can act as a repressor of virulence like RsmA. Alternatively, a *glpA*$^+$ plasmid could be transformed into a *csrA*$^-$ strain of *E. coli* in an effort to determine whether GlpA can repress glycogen synthesis like CsrA.

In order to test the hypothesis that the integration of certain *D. nodosus* elements downstream from the *askA* and *pnpA* integration sites would affect the expression of genes upstream, northern blot analyses were performed in which both *glpA* and *pnpA* specific DNA fragments were utilised as DNA probes (Section 7.2.11). Although no meaningful data was obtained, the parallels between the expression of *rsmA* in *E. carotovora* and of *glpA* in *D. nodosus* were of interest.

In Soft-rotting strains of *E carotovora* the *rsmA* probe hybridises to two *rsmA*-specific transcripts in northern hybridisation experiments of 1100 bp and 310 bp respectively. In contrast, non soft-rotting strains, the *rsmA* probe hybridises to a 310 bp transcript only (Mukherjee *et al.*, 1996). In *D. nodosus,* the northern blot results do suggest that at least in virulent *D. nodosus* strain 1311, *glpA* also hybridises to two transcripts. It is unclear from these results however, whether that as for *E. carotovora* there is a general trend that virulent strains of *D. nodosus* transcribe two *glpA* messages, whilst benign strains transcribe only one, and this requires further investigation, as discussed previously (Section 7.2.11). Patnogenic isolates might appear to transcribe more of a virulence-repressor if the transcript was controlled by an antisense mechanism.

In *D. nodosus*, we need to determine if expression of *glpA* at the RNA level or the protein level is different in benign compared to virulent strains of *D. nodosus*. Northern blot

analyses of *D. nodosus* proteases should also be done in order to determine whether *glpA* does increase the rate of degradation of transcripts of thermostable proteases.

Further investigation of a large number of strains of *D. nodosus* should be done in order to investigate the hypothesis that in virulent strains the *vap* element is integrated next to *pnpA* and the *vap* element or *intC* element is integrated next to *askA*. Such a study may allow the development of a more efficacious multiplex PCR test to differentiate between benign and virulent strains of *D. nodosus* more rapidly.

This work also has significant implications for footrot eradication programs that are currently in place in Australia (Section 1.4.1). The authorities are in general not concerned about footrot caused by benign strains of *D. nodosus*, since the disease caused by these strains is less severe and self-heals upon the resumption of dry conditions (Stewart, 1989) (Section 1.1.1). In contrast, more serious intermediate and virulent infections result in more severe production costs, are much more contagious and consequently authorities quarantine these properties. However, if *D. nodosus* genetic elements are able to modulate virulence by site-specific integration into the *D. nodosus* chromosome, and have the potential to be transferred horizontally between benign and virulent isolates of *D. nodosus* the authorities cannot afford to be complacent about benign strains of *D. nodosus*. A benign strain could be reversibly converted to a virulent strain and *vice versa* thereby evading eradication measures. There have been reports of sheep that were thought to be infected with a benign footrot strain, later expressing virulent disease.

It is also possible that these *D. nodosus* elements do have an even broader host range if they are transferred *via* conjugation, rather than by a transduction mechanism. Both the *vap* element and the *intD* element are likely to be transferred by a conjugative mechanism, given that they both contain a putative *oriT* region, and the latter contains both mobilisation genes (Section 6.2.1) and putative *tra* genes (B. Shaw & B. F. Cheetham, unpublished). To

assess the host range of the *D. nodosus* elements, other bacteria that are often associated with footrot infections should also be investigated in order to determine whether they contain copies of these *D. nodosus* elements.

There is also the need for epidemiological studies of *D. nodosus* infections, since it is currently unknown whether the recurring infections from one season to the next, on a given property is due to the incomplete eradication of a strain from a previous infection or whether it is due to a new infection from outside the property being introduced to the property. The development of DNA fingerprinting techniques have been proposed for this purpose (Cheetham & Katz, unpublished). Such studies would also indicate whether a benign strain or a virulent strain of *D. nodosus* could be converted to a more or less virulent strain respectively in the field.

A significant inference from the study of the *vap* regions and other mobile genetic elements in *D. nodosus* is that to have an effect on virulence, genes carried by pathogenicity islands or genes affected *via* the integration of pathogenicity islands would not necessarily have to encode classical virulence determinants. It is possible that pathogenicity islands may either encode genes or integrate at a position in the pathogenic host, that ultimately results in the modulation of genes involved in regulation of either housekeeping or virulence functions or both. Though such complex mechanisms would be less easily identified, they may still result in the modulation of virulence. In the future, many other such examples of this may be found.