CHARACTTERIZATION OF A PLASMID CONFERRING NAD INDEPENDENCE IN *Haemophilus paragallinarum*

by

ANNA ELIZABETH VAN ZYL

Submitted in fulfilment of the requirement for the degree

MAGISTER SCIENTIAE

In the Faculty of Natural and Agricultural sciences
Department of Microbial Biochemical and Food Biotechnology
University of the Free State
Bloemfontein
South Africa

May 2003

Supervisor:	Dr. J. Albertyn
Co-supervisors:	Prof. R.R. Bragg
	Dr. E. van Heerden

Acknowledgements

I acknowledge, with gratitude, the following persons who have made valuable contributions toward this research

Dr. J. Albertyn for his inspiration, availability for consultation, patience and guidance, both with the experimental work, and the writing of this thesis.

My co-study leaders, **Prof. R.R. Bragg and Dr. E. van Heerden**, who's knowledge contributed to the refinement of the study.

All the members of the Department of Microbiological, Biochemical and Food Biotechnology for their kindness, generosity and for lighting up my days in various ways

The National Research Foundation for financial support

Cornelia Casaleggio for her friendship, support, guidance, understanding and advice. Thank you for being like a mother in the lab.

Elsabé Botes for encouragement, sequencing of samples and expert advice on troubleshooting in this regard.

Michel Labuschagne for his friendship, and assistance regarding scanned images for this thesis.

Sarel Marais for his friendship, moral support, advice and for his contribution to my thesis especially with respect to the drawing of chemical reactions.

To all my friends, especially **Gert, Inge-marie, Nelmarie, Nannette and Chaneè** for being sensitive, patient, thoughtful and kind.

René van der Westhuizen for motivation, comradeship and love.

To my parents, thank you Mum and Dad for endless love, inspiration, faith in me, a special thanks for your support through good times and prayers during the bad times through all my years of study.

My sisters **Bertha and Joanie** for inspiration, interest and always being optimistic about my studies, for loving me and caring for me.

Above all to **God the creator of all life** for granting me the opportunity to work with and learn more about DNA, the molecules He designed to encode the very essence of life itself.

This thesis is dedicated with all my love
to my parents and sisters
for they have sacrificed much for the sake of my studies.
May God bless them for this.

Ecclesiastes 9:11

I have seen something else under the sun:
The race is not to the swift
or the battle to the strong
nor does food come to the wise
or wealth to the brilliant
or favour to the learned;
but time and chance happen to them all

Table of Contents

Chapter 3

Materials and Methods 21

Chapter 4

Results 28

Chapter 1

Introduction

Haemophilus paragallinarum infects the upper respiratory tract of poultry causing a disease known as infectious coryza. The disease begins in the nasal passages and sinuses with signs like lacrimation, nasal discharge and swelling around the eyes. Major economic losses are suffered as a result of a reduction in egg size and a reduction of up to 40% in egg production in both laying and breeding flocks, while recovery time can take up to nine weeks.

In many bacteria the essential cofactor NAD^+ is synthesised through the *de novo* pathway, but can also be synthesised by a pyridine salvage pathway (Foster and Moat, 1980) Members of the family *Pasteurellaceae* do not posses either of these pathways for NAD biosynthesis. These bacterial species is growth factor dependent and acquire NAD from their environment. This requirement of NAD^+ has played an important role in classification into the family *Pasteurellacea* (Niven and Lévesque, 1988). For many of these *Pasteurellaceae* species, previously described as NAD^+ dependent, NAD^+ independent variants have been identified. These include strains of *Haemophilus parainfluenzae,* which can cause pneumonia and meningitis in humans (Gromkova and Koornhof, 1990), *Haemophilus ducreyi,* which causes the sexually transmitted disease chancroid in humans (Windsor *et al*, 1991) and *H. paragallinarum,* which causes fowl coryza (Bragg *et al*., 1993b). NAD^+ independence in these strains has been shown to be plasmid mediated (Bragg *et al.*, 1993b; Windsor *et al.,* 1991 and Gromkova and Koornhof, 1990)

A study by Carlisle (1998) revealed the presence of a 6kb plasmid responsible for the NAD^+ independence in *H. paragallinarum* strains. Very little is known about this plasmid except for studies by Taole *et al*. (2002) indicating that plasmid-bearing strains of *H. paragallinarum* are less virulent than the

naturally occurring NAD^+ dependent strains. The genetic information encoded by this plasmid and the means through which independence is conferred is still unknown.

In this study, plasmids99999 from three strains of NAD^+ independent *H. paragallinarum* organisms and a NAD^+ independent *Pasteurella avium* species was isolated. Isolated plasmid DNA was investigated in order to find possible differences between plasmids from different strains.

Structural analysis of the plasmid, through digestion with various restriction enzymes and the construction of restriction map(s) was accomplished followed by sub-cloning and sequencing of the plasmids, in order to identify the genetic information encoded on these plasmid(s).

Chapter 2

Literature review

2.1 Infectious coryza

Infectious coryza (IC) is a bacterial disease of economic importance since it causes an increased culling rate in meat chickens, as well as a reduction in egg production of between 10 to 40 % in laying and breeding hens (Yamamoto, 1981). This is a particular problem on multi age farms, since the severity of the disease increases as the age of the flock increases (Matsumoto, 1988). Recovery time can take up to 9 weeks, which results in a major loss in egg production.

The disease begins in the nasal passage and sinuses and can be recognised when birds show signs that include lacrimation and nasal discharge with swelling around the eyes. The swelling can be unilateral or bilateral. Some chickens develop tracheitis and air sac infection, but the lungs generally stay uninfected. The major problem caused by the disease is degeneration of the ovaries and atrophy of the follicles, which results in the inability of layer hens to produce even after full recovery from the disease (Buys, 1982).

The disease can spread in a number of ways. Hens that recover from infection remain carriers of the disease for up to 9 months. These hens are often sold to uninfected farms, causing the disease to spread. Carrier birds purchased as replacements are regarded as the main source of IC (Yamamoto, 1984). Page (1962) found that the disease could spread for up to 2.4m through the air. In epidemiological studies based on experimental laboratory trials, Page (1962) also established that IC could be transmitted by contaminated drinking water. Although Page (1962) believed that the disease could not be spread through mechanical means Buys (1982) suggested that the cloths used to wash the water bowls also spread the disease.

2.2 Causative agent of Infectious coryza

Different species of bacteria can be isolated from the sinuses of chickens showing signs of IC. These include *Ornithobacterium rhinotracheale, Pasteurella* species such as *P. avium, P. volantium, P. species A* and *Haemophilus paragallinarum.* These organisms are closely related to each other and are difficult to differentiate (Holt *et al.,* 1994). *H. paragallinarum* is the causative agent of IC. *H. paragallinarum* was first isolated by De Blieck (1931) and named *Bacillus heamoglobinophilus coryza gallinarum* (Elliot and Lewis, 1934). Delaplane *et al.* (1938) investigated the growth requirements of these organisms and found their isolates to be dependent on both X factor (haemin) and V factor (nicotinamide adenine dinucleotide: NAD) for growth. The name *Haemophilus gallinarum* was then proposed for these organisms (haem = blood, philus = loving, gallus = chicken). Later, Biberstein and White (1963) found isolates, which were only dependent on NAD for growth. They then proposed the name *Haemophilus paragallinarum* for these isolates, which are similar to *H. gallinarum*, except for their haem independence. However, due to the storage loss of the earlier strains of *H. gallinarum* the haem dependence of these strains cannot be confirmed (Rimler, 1979).

2.3 Identification of *H. paragallinarum*

H. paragallinarum isolates are gram negative coccobacilli or rods, generally less than 1μm in width and of variable length. By occasionally forming threads or filaments these cells can show marked pleomorphism (Sneath and Johnson, 1973). These bacteria are included under the family Pasteurellaceae, being related to the *Actinobacillus* spp. but are biochemically distinguishable (Verschoor *et al*., 1989). *H. paragallinarum* optimally grows at 37°C in the laboratory, although the body temperature of its host is 42ºC.

Yamamoto (1984) developed a biochemical test for the identification of *H. paragallinarum,* which was evaluated with various other methods by Blackall (1983). He concluded that this test based on the ability of *Haemophilus paragallinarum* to form acids from carbohydrates, is the best method for the detection of carbohydrate fermentation. The results of these and other biochemical tests, which are used for the identification of *H. paragallinarum* are listed in table 1.

Table 1: Secondary biochemical tests used in the identification of *H. paragallinarum*.

Reaction	**Results obtained from *Haemophilus paragallinarum***
Gram's stain	-
Catalase	-
Indole	-
Nitrate reduction	+
Oxidase	+
Urease	-
Fermentation of carbohydrates	
Fructose	+
Galactose	-
Glucose	+
Lactose	-
Maltose	+
Mannitol	+
Sorbitol	+
Sucrose	$+^{v}$
Trehalose	-
Xylose	-

v- variable results

Polymerase chain reaction is a powerful technique developed to amplify DNA fragments by using high temperature resistant polymerase. Chen *et al.* (1996) established a polymerase chain reaction technique that can be used to identify *H. paragallinarum* based on sequence specific primers. The three

different primers that were designed by Chen *et al.* (1996) are indicated in Table 2.

Table 2: Sequences of the primers designed by Chen et al (1996) for development of the *H. paragallinarum* specific PCR.

Primer name	Primer sequence
F1	5'-CAA TGT CGA TCC TAC AAT GAG-3'
N1	5'-TGA GGG TAG TCT TGC ACG CGA AT-3'
R1	5'-CAA GGT ATC GAT CGT CTC TCT ACT-3'

Two different combinations of these primers were used [F1/R1 (HPG-1) and N1/R1 (HPG-2)] with HPG-1 yielding a fragment of approximately 1.6 kb, while a 0.5 kb fragment was obtained from the HPG-2 combination. This technique is very specific for the identification of *H. paragallinarum*. Using these primer sets it is possible to differentiate between *H. paragallinarum* and other closely related species isolated from the sinuses of chickens such as *Ornithobacterium rhinotracheale* and *P. avium*, *P. volantium* and *P. species A*.

2.4 Growth requirement

Once isolated from chickens, *H. paragallinarum* can be grown in rich media such as chicken test medium supplemented with chicken blood serum, thymine and nicotine amide dinucleotide (TM/SN). The best-known media for *Haemophilus* species is chocolate agar and blood tryptose agar incubated at 37°C (Sneath and Johnson, 1973). The bacteria are micro-aerophilic and the optimal growth is obtained under enhanced CO_2 concentrations, these conditions can be obtained through the use of a candle jar.

2.5 Virulence factors

Very little is known about the virulence factors of *H. paragallinarum* that cause the telltale symptoms of IC. Virulence factors, which induce a protective immune response, in either vaccinated or infected chickens are likely to be cell surface located antigens (Ogunnariwo and Schryvers, 1992). Possible candidates include outer-membrane proteins (OMP's), polysaccharides and lipopolysaccharides (Blackall, 1989).

2.5.1 Outer Membrane Proteins

It is known that inactivated vaccines confer less protection than natural infections, thereby suggesting that certain protective antigens produced by *H. paragallinarum in vivo,* are either lacking or are produced only in small quantities under *in vitro* conditions (Blackall, 1989).

The iron-regulated OMP's are produced *in vivo* and under iron-restricted conditions *in vitro* in a number of bacterial species (Snipes *et al.*, 1988). Brown and Williams (1985) showed that not only do iron-regulated OMP's from several bacterial species have immunogenic properties, but that the antibodies against them are cross protective.

Blackall *et al*. (1990a) examined the outer membrane protein profiles of four *H. paragallinarum* isolates (0083, 0222, Modesto and HP31). They distinguished major and minor OMP's and designated them OMP A-H. OMP A (87 kDa) is similar in size to the iron regulation proteins found in *Pasteurella multocida* and *Escherichia coli*.

Ogunnariwo and Schryvers (1992) investigated the role of iron-regulated OMP's and the acquisition of iron in *H. paragallinarum*. They found the binding of chicken transferrin by the avian haemophili to be specific, as no binding was detected with conjugates of human or bovine transferrin. It was suggested that *H. paragallinarum* depends on the transferrin receptor for iron

acquisition *in vivo,* which could explain that the loss or drastic alteration of these proteins reduces or eliminates the ability of the bacteria to cause IC.

The conserved molecular sizes and function of the OMP's, among the different avian haemophili species, raises hopes of conserved surface epitopes. The possibility therefore exist of the development of a single cross-protective vaccine against all *H. paragallinarum* (Ogunnariwo and Schryvers, 1992).

2.5.2 Formation of toxins

The ability of microorganisms to attach and colonize specific sites in the host, and the formation of substances such as toxins or enzymes which cause inflammation and damage to the host, can serve as virulence factors. The toxins are responsible for some of the symptoms of infection and the term endotoxin is used to refer to the toxic component of the lipopolysaccharides (Starr and Taggart, 1995). Iritani *et al.* (1981) showed that polysaccharide extracts from *H. paragallinarum* serotype A and C (strain 221 and S1 respectively) are toxic and caused hydropericardium in chickens, but failed to induce hemagglutination-inhibition (HI) antibodies. The role of this component in natural infection has not been established.

2.5.3 The capsule and its role in natural infections

In 1965, Fujiwa and Konno reported on the histopathology in chickens affected by experimentally induced coryza. They found acute catarrhal inflammation of the mucous membranes in the nasal passages, infra-orbital sinuses and trachea. The attachment and subsequent colonization of the mucous membrane, thus important in the infectious process, is characteristic of virulent strains. Sawata *et al.* (1984b) found that non-encapsulated *H. paragallinarum* strains were sensitive to the bactericidal effects of normal freshly prepared chicken serum.

Susumu *et al.* (1982) investigated the relationship between the adhesion of *H. paragallinarum* to chicken embryo fibroblasts (CEF) *in vitro* and virulence in *H. paragallinarum* V- factor dependent strains no. 221, FY-3 and GF. Strains 221 and FY-3 were shown to be pathogenic with strain no. 221 exhibited marked adhesions to the plasma membrane of the CEF, with fuzzy material extending outwards. In some instances the organisms were so closely attached to the plasma membrane that no intervening space could be seen between them and the plasma membrane. The organisms were also seen to occur in the cytoplasm of CEF, being enclosed in a membrane bound vesicle (Susumu *et al*., 1982). Using scanning electron microscopy, strain no. 221 was also seen to adhere to the cilia of the chicken tracheal epithelium, penetrating into the inter-ciliary spaces attaching to the ciliary surface. Generally the two virulent strains, 221 and FY-3, adhered to the CEF *in vitro*, and with increased incubation time the percentage of cells with adherent bacteria was increased (Susumu *et al.,* 1982).

Adhesion of strain no. 221 is thought to enable the cells to colonize the surface of the epithelial cells, by resisting the removal function of bathing secretions and ciliary movement of the respiratory tract. Residence of the organisms between the cilia on the tracheal cells is also thought to protect the cells from the mucociliary clearance mechanism (Susumu *et al.,* 1982).

2.6 Serotyping

The first agglutination for serotyping of the organism was performed by Page (1962). He used a plate method based on the reaction between antibody and particle bound antigen that result in the clumping of the particles, known as agglutination, to try and differentiate between different strains of NAD^+ dependent isolates. The isolates were grouped in three different serovars, namely A, B and C. However, the original type C organisms of Page were lost in storage. Rimler *et al.* (1976) confirmed the serovar A of Page and described the Modesto strain as a new agglutinin serovar C – type isolate.

Kato and Tsubahara (1962) used agglutination tests and reported on the occurrence of three different types I, II and III. Sawata *et al*. (1982) used an agglutination test to serotype isolates of *H. paragallinarum* in Japan, reporting that this serotypes I and II corresponds to the Page serotypes A and C respectively. In additions, Sawata *et al*. (1980) found the Page B organisms to be untypeable. They were considered to be variants of serotype A or C strains that had lost their type specific antigens. However, Yamaguchi and Iritani (1989) reported that the Page B organisms were capable of hemagglutination.

The Page method has been widely used for the serotyping of *H. paragallinarum*, although the occurrence of two major problems such as spontaneous agglutination (Iritani *et al.,* 1978) and a large percentage of untypeable isolates (Blackall and Eves, 1988) resulted in the decrease in popularity of this technique.

Kume and co-workers (1983a) developed a hemagglutination serotyping method for *H. paragallinarum,* which is currently the most widely used method of serotyping this bacterium. Hemagglutination antigens, obtained by potassium thiocyanate extraction and sonification and gluteraldehyde fixed chicken erythrocytes, are used to carry out the hemagglutination test (HA). The application of this serotyping scheme resulted in the identification of three different serogroups consisting of seven different serovars. The serogroups were termed I, II and III the serovars were termed HA-1 to HA-7. Serovars HA-1 to HA-3 was found to belong to serogroup I: serovars HA-4 to HA-6, to serogroup II: and serovar HA-7, to serogroup III. It was established that these serogroups were the same as Page`s (1962) serogroups A, C and B respectively (Sawata *et al*., 1980).

Both Eaves *et al.* (1989) and Blackall *et al.* (1990a) discovered new serovars and it seems most likely that the discovery of new serovars will continue. Kume's scheme was altered by Blackall et al. (1990b) to accommodate newly discovered serovars. In this scheme Blackall *et al.* (1990a) suggested that groups I, III and II of the Kume scheme be changed to A, B and C respectively. This emphasised their relation to the original groups classified by

Page (1962). The existing serovars should therefore be numbered (e.g. A1 to A4) and new serovars can be allocated the next number in the group.

2.6.1 Serotyping of South African strains.

The first serotyping on South African isolates was performed by Buys (1982). He reported the occurrence of type A and B organisms but did not found serotype C organisms among his isolates. Kume *et al*. (1983b) serotyped two South African isolates designated as HA-6. Blackall *et al.* (1990a) later renamed these isolates as C-3. Blackall and Eaves (1988) also serotyped South African isolates and found them to belong to serogroups A-1, B-1, C-2 and C-3. The C-3 organism seemed to be unique to South Africa.

2.7 Vaccination and protection against IC

Chickens can be protected against IC by vaccination. Commercial bacterins prepared from chicken embryos or broth may be autogenous or may contain strains of 2 to 3 different serotypes. The products, inactivated with formalin or merhiolate, must contain at least 10^8 CFU/ml to be effective. They may contain aduvants, stabilizers, or saline diluents (Yamamoto, 1991).

The protective immunity is related to the serotypes of the organism used in the vaccine and the serotype of the organism infecting the bird. Birds vaccinated with a vaccine containing only serovar A, will only be protected against challenge by serovar A isolates. They will be successfully infected with either serovar B or C isolates (Blackall 1991).

Two registered vaccines are manufactured in South Africa, one containing combinations of stereotype A and C-3, the contents of the other is however unknown. A new vaccine, which contains all 4 of the South African serovars, A-1, B-1, C-2 and C-3, has been submitted for registration (Bragg and Greyling 1999 unpublished data). It is anticipated that this vaccine will limit the

amount of infection in South Africa since it will be the only available vaccine that contains all four the serovars occurring in this country. Apart from these, there are no less then 8 registered vaccines, which have been imported into SA. Although most of these vaccines are bivalent, two are trivalent, one containing serovars A-1, B-1 and C-1 and the other containing serotype A-1, a South American A type and C-1.

2.8 Hemagglutinen of *H. paragallinarum*

The first serotyping of *H. paragallinarum* was performed by Page (1962) using a plate agglutination method. The hemagglutination assay used by Kume *et al.* (1983a) for the serotyping of *H. paragallinarum* is currently the most popular method. Both these schemes are dependant on the hemagglutination ability of *H. paragallinarum*.

The hemaglutinins are cell surface antigens which were shown to be involved in attaching *H. paragallinarum* to the mucosal membrane during the first step of infection (Iritani *et al.,* 1977).

Antibodies produced against hemagglutinin-induced antigens play a key role in immunity and serotyping. Kato (1970) reported that vaccinating chickens with hemagglutination-inhibition (HI) antibodies provided protection against *H. paragallinarum* infection in chickens. Based on these results, the HI test has been used to evaluate the immunogenicities in vaccines against infectious coryza in Japan (Sawata *et al.,* 1978).

Sawata and co-workers (1984a) reported on the presence of at least 3 types of hemagglutinen serotype 1 strains. These were subsequently defined as: HA-L, HA-HL and HA-HS. HA-L is a heat-labile, trypsin-sensitive, hylaluronidase-resistant antigen, which is active against glutaraldehyde (GA)-fixed erythrocytes (RBC). The HA-HL was defined as heat-labile and trypsin-resistant, whereas HA-HS was defined as heat-stable and trypsin-resistant.

The HA-HL and HA-HS agglutinate freshly collected chicken RBC but do not agglutinate GA-fixed RBC.

Kume *et al.* (1983a) reported serological and immunologic differences, with HI antibody-producing antigenicity and protective activity, among the different hemagglutinen. They found that protective immunity was only induced when chickens were injected with HA-L but not with HA-HL, although antibodies against HA-HL were present. HA-HS either lacked HI antibody–producing antigenicity against HA-L and HA-HL or protective activity, even though non-specific HI antibody against HA-HS was detected in the injected control rabbit sera. Protective immunity of the HA-L hemagglutinin was completely inactivated by heating at 121^{o}C for 2 hours (Sawata *et al.,* 1979). When the HA-L's were heat-treated at 72^{o}C to 100^{o}C for 30 minutes, protective activity remained to a varying extent, but HI antibody-producing antigenicity of the heat-treated hemagglutinen against HA-L was completely lost by the heating. These results suggested that among the 3 types of hemagglutinen, HA-L is solely responsible for immunogenicity, but antibody formation against HA-L is not always essential for protection.

Recently Hobb *et al.* (2002) succeeded in isolating and purifying a hemagglutinin from strain 0083 belonging to Page serogroup A termed hemagglutinin *hagA*. Following the N-terminal sequencing they identified the gene encoding this protein. The *hagA* gene was cloned and sequenced from the 11 serogroups. Sequence comparisons revealed limited variations that did not correlate with the serological grouping of the different serogroups. Possible explanations suggested by them included:

(1) Other surface proteins, which might be involved in the serotypic differences.
(2) The presence of multiple hemagglutinens, which was previously mentioned by Kume *et al*. (1983b).
(3) Post translational modification of the hemagglutinen.

2.8.1 The role of Hemagglutinen in immunity

The serovar specific hemagglutinin of *H. paragallinarum* was thought to play a role in the protective immunity due to the correlation between the Hemagglutination-inhibition (HI) titre and the protection (Otsuki and Iritani, 1974). This was confirmed when Takagi *et al.* (1991) cloned the total genomic DNA of strain 221 into the vector plasmid, pBR322, which was introduced into *E. coli.* One of the transformed *E. coli* isolates showed hemagglutinating activity. Chickens were vaccinated with killed *E. coli* expressing the hemagglutinin-produced HI antibodies and were protected against challenge by 221.

2.9 Nicotinamide Adenine Dinucleotide

2.9.1 NAD^+ as co-enzyme

NAD^+ is directly involved in energy generating (catabolic) reactions. In the cell it acts as a co-enzyme that functions at the active site of an enzyme. During enzyme action unbound protons (H^+) and electrons are generated, these are picked up by NAD^+ converting it to NADH. Through these oxidation-reduction reactions NAD^+ have a direct impact on virtually every cellular metabolic pathway. (Starr and Taggart 1995.)

2.9.2 ATP Production

NAD^+ function as the first carrier in the electron transport chain, the energy derived from electron transport is used to synthesise ATP. Oxidation of NADH in the presence of ADP results in the formation of 3 ATP molecules. The balanced equation for the reaction is:

$$NADH + H^+ + \tfrac{1}{2} O_2 + 3ADP + 3P_i \rightarrow NAD^+ + 4H_20 + 3ATP$$

The free energy of living cells are captured as ATP`s, hydrolysis of the ATP yields 30.5 kJ/mol energy to the cell (Marks *et al.,* 1996).

2.9.3 DNA Repair

Bacteria use NAD^+ in place of ATP to drive the DNA ligation reaction catalysed by DNA ligase (Ziegler and Oei, 2001). Like the ATP-dependant ligases of other organisms, the NAD^+ dependant enzyme is required for DNA replication and repair. The enzyme catalyses the sealing of nicks between a 3'-hydroxyl and 5'-phosphate group in double-stranded DNA; therefore it is essential for the survival of the organism (Subramang *et al.,* 1996).

2.9.4 Metabolic synthesis of NAD^+

The multiple roles of NAD^+ in cellular functions necessitate that NAD^+ biosynthesis must be actively regulated and proper NAD^+ levels maintained.
NAD^+ can be synthesised from the aromatic amino acid tryptophan. Tryptophan contains a conjuncted indole ring and its metabolism is linked to that of niacin. Tryptophan, like other aromatic amino acids, are synthesised through the shikimate pathway. The pathway is present in bacteria, fungi and plants, but not in animals. Tryptophan degradation yields precursors for NAD^+ synthesis. Tryptophan degradation takes place through a feed forward mechanism, thus tryptophan stimulates its own degradation by allosterically activating tryptophan oxygenase (Fowkes, 1992).

2.9.5 NAD^+ requirements of *Haemophilus*

In prokaryotes NAD^+ is synthesised through the *de novo* pathway. Until 1989 it was believed that the family *Pasteurellaceae*, comprising of the genera *Haemophilus, Pasteurella* and *Actinobacillus*, could not synthesise NAD^+ and

therefore lack the *de novo* pathway for NAD^+ biosynthesis (Foster and Moat, 1980). This requirement of NAD^+ as growth factor has been used as an essential criteria in the classification of the gram-negative bacteria in the genus *Haemophilus* (Hollander and Mannheim, 1975).

In addition to lacking the *de novo* pathway for NAD^+ biosynthesis, the organism possesses only a limited capacity for the uptake of pyridine nucleotides and precursors. Cynamon *et al.* (1988) investigated the NAD^+ requirements of *H. parainfluenzae*. They demonstrated that NAD^+, nicotinamide mononucleotide (NMN) and nicotinamide riboside (NR) can function as V-factors however, nicotinamide (NAm) and nicotinic acid as well as quilinic acid (QA) and the α - anomer of NAD^+, are ineffectual. O' Reilly and Niven (1986) characterised the compounds that can serve as V-factors in porcine haemophili as those possessing an intact pyridine-ribose bond, in the β-configuration, and a pyridine carboxamido group at position 3. $NADP^+$ was found to be the exception to this rule since only one of the 30 strains tested by O' Reilly and Niven (1986) was capable of using purified $NADP^+$ as a pyridine nucleotide source.

2.9.6 NAD^+ synthesis and uptake

Wheat and Pittman (1960) reported that *Haemophilus influenzae* possesses NAD^+ pyrophosphatase, which cleaves NAD^+ to yield NMN and AMP. This enzyme has since been isolated and purified from *H. influenzae,* its activity is required when V factor is supplied in the form of NAD^+. It therefore appears that NAD^+ is not transported intact by *H. influenzae* but is first hydrolysed to yield NMN (Kahn and Adderson, 1996). Similarly Mouahid *et al.* (1992a) found that in *H. paragallinarum* NAD^+ is cleaved to NMN and subsequently NR. Both metabolites enter the cell where they are substrates for the intracellular synthesis of NAD^+. Kasarov and Moat (1973) suggested a biochemical pathway for the biosynthesis of NAD^+ in species of the family *Pasteurellaceae* (Fig. 1).

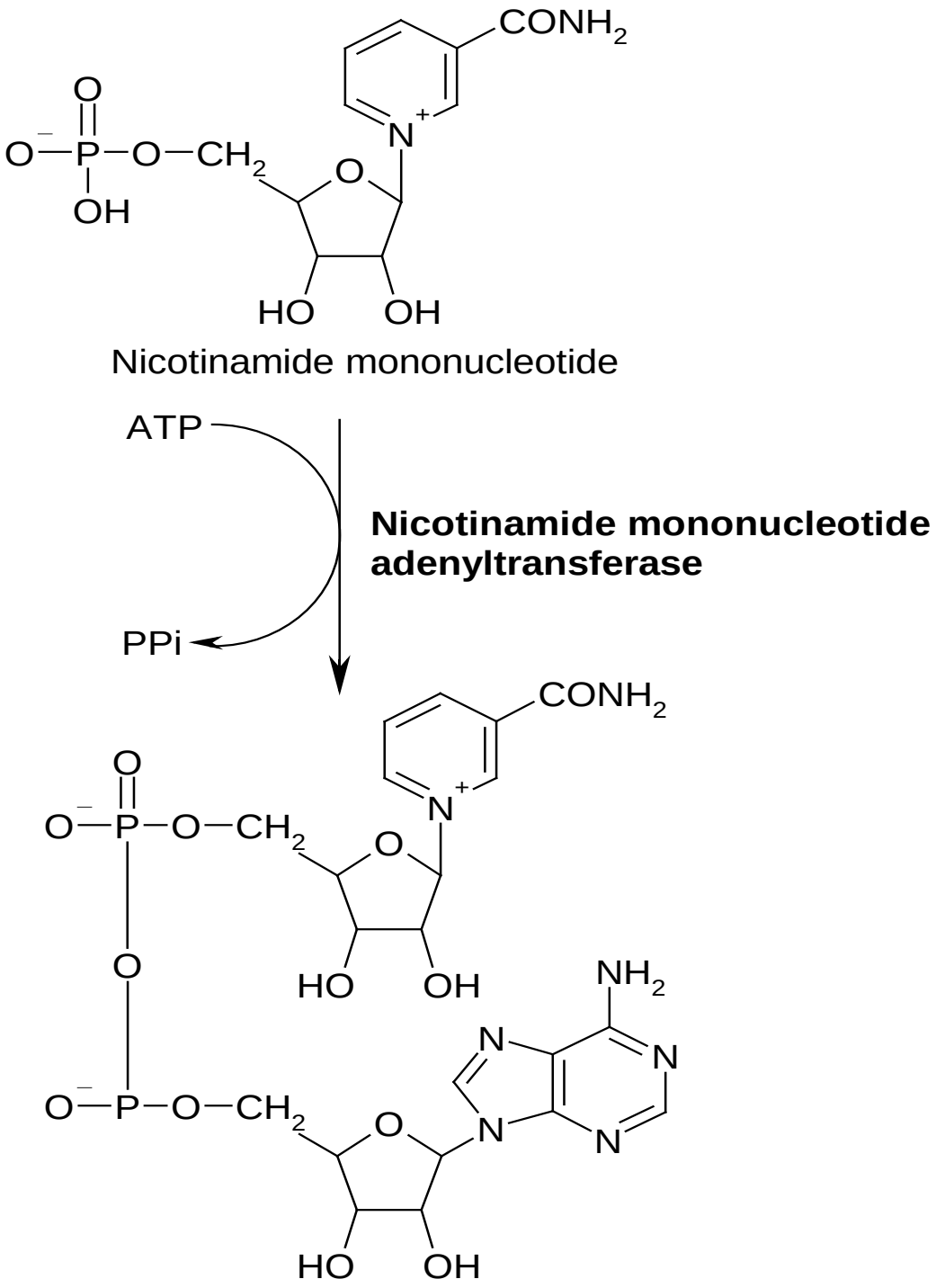

Figure 1. Biochemical pathway for the biosynthesis of NAD^+ as found in the family *Pasteurellaceae*.

2.9.7 NAD^+ independence of *Haemophilus*

The first NAD^+ independent *H. paragallinarum* strain was isolated from chickens showing symptoms of IC in Natal, South Africa (Horner *et al.,* 1992). These organisms differed from NAD^+ dependent *H. paragallinarum* in their ability to grow without V-factor and by forming bigger colonies on blood agar. Horner *et al.* (1992) suggested that these isolates could not be identified as *H. paragallinarum* because of their NAD^+ independence. Although there was a difference in the protein profiles using biochemical tests it was not possible to separate these isolates from *H. paragallinarum* in any other way but their ability to grow in the absence of NAD^+. Mouahid *et al.* (1992a) further characterised the NAD^+ independent isolates, using DNA:DNA hybridisation. They found that there was a high level of genetic relation between the NAD^+ independent and NAD^+ dependent cultures. Mouahid *et al.* (1992b) failed to differentiate between independent and dependant isolates using gaschromatography in combination with mass spectrometric identification of the carbohydrates, fatty acids and phospholipids.

Verschoor *et al.* (1989) developed a panel of monoclonal antibodies (mabs) specific for *H. paragallinarum*. These were raised against strains 0083 (Page's serotype A) and 0222 (Page's serotype B) and two South African field isolates (isolate M 85 obtained in 1985 and isolate SB 86, in 1986) from diseased birds in flocks vaccinated with strain 0083 and 0222. The objective of establishing this panel of mabs was to detect antigenic differences between "vaccine" strains and "field " isolates. Bragg *et al.* (1993a) used these mabs to try and distinguish between the NAD^+ independent and NAD^+ dependent strains. They obtained mab patterns from the NAD^+ independent isolates, which were indistinguishable from those obtained from the NAD^+ dependent isolates. Bragg *et al*. (1993a) therefore suggested that these isolates were in fact NAD^+ independent *H. paragallinarum.* Horner *et al*. (1995) later conceded that their isolates were NAD^+ independent *H. paragallinarum*.

Bragg *et al.* (1993b) made crude DNA extracts from NAD^+ independent organisms. This extract was mixed with competent NAD^+ dependant cells and

incubated for 1 hour. After incubation, liquid casman's medium supplemented with 10% sterile chicken serum was added to the mixture of crude DNA and competent cells. The cultures where incubated for 6 hours, after which they were inoculated on BTA plates without the addition of feeder culture. The NAD^+ independent growth of these previously NAD^+ dependant cells was thus attributed to transferable genetic material. They then suggested that the NAD^+ independence of *H. paragallinarum* is possibly plasmid mediated.

2.10 Plasmids

2.10.1 Plasmid isolated from *H. paragallinarum*

Carlisle (1998) scanned NAD^+ independent *H. paragallinarum* strains for the presence of plasmids and found a small 6kb plasmid. When NAD^+ dependant strains were transformed with this plasmid they became NAD^+ independent.

Plasmids have been found to carry genes that mediate virulence properties of several important enteric bacteria including some *E. coli* strains and *Salmonella dublin* (Chikami *et al.,* 1985). It is therefore possible that the virulence of *H. paragallinarum* can be influenced by its NAD^+ independence. Since evidence suggested that independent strains were less virulent than wild type strains (Taole *et al.*, (2002), it was concluded that transformed isolates could possibly be used as a live vaccine (Bragg personal communication). It was therefore necessary to investigate the impact of transformation on the hemagglutinen of the transformed strains. Bragg *et al.* (1995) showed through hemagglutination and hemagglutination inhibition that the hemagglutinen of 0083, A745/92 and M85 were not affected by transformation. Therefore the use of transformed strains for vaccine production appeared to be a valid approach.

This was further investigated by Taole *et al.* (2002). They isolated a plasmid from a NAD^+ independent strain (1742) and transformed it into the NAD^+ depended strain 46-C3 rendering it NAD^+ independent. Chickens were then

challenged with both the dependent and transformed independent strain of 46-C3. The disease score per bird was obtained by notating the clinical signs of IC on a daily basis as, mild, moderate or severe. It was found that the transformed isolates caused IC, while the wild-type 46-C3 strain was found to be more virulent than the transformed strain.

2.10.2 Plasmids encoding independence in other *Haemophilus* species

Gromkova and Koornhof (1990) isolated the first *Haemophilus* species that did not need NAD^+ *in vitro*. An NAD^+ independent *Haemophilus parainfluenzae* strain was isolated from humans, but the mechanism of this independence was unknown. It was possible to transform the NAD^+ dependant isolates to NAD^+ independence by DNA extraction and transformation. The principle of this method is that characteristics of a bacterium can be transformed to closely related bacteria if cultivated together (Forbisher, 1970). The *H. parainfluenzae* was screened for the presence of a plasmid. Windsor *et al.* (1991) found that the NAD^+ independence of *H. parainfluenzae* was encoded for by a 5.25kb plasmid.

NAD^+ independence in *H. ducreyi* was similarly found to be plasmid mediated by Martin *et al*. (2001) who isolated a 5.25kb plasmid from *H. ducreyi.* Transformation of this plasmid to NAD^+ dependent *Actinobacillus pleuropneumoniae* rendered the organism NAD^+ independent.

Chapter 3

Materials and Methods

3.1 Enzymes and Chemicals

All chemicals used were of molecular biology or analytical grade. The following chemicals were obtained from MERCK Germany: NaCl, starch, glucose, bacteriological agar, NaOH, chloroform, ethanol, isoamyl alcohol, $MgSO_4$, $MgCl_2$, acetic acid, yeast extract, tryptone and EDTA. IDT Technologies Inc.,: Oligonucleotide primers. Roche Biochemicals: Taq DNA polymerase, dNTP`s, Taq polymerase, Tris, RNase A, restriction enzymes (*EcoRI, BamHI, NotI, PstI, MboI, XhoI, SspI, SalI, SacI, SmaI, XbaI, Bsty*I, *Hae*III, *Xcm*I, *Hind*III and *Bgl*II) and the modifying enzyme Klenow polymerase. SIGMA Aldrich, U.S.A.: Thiamine HCl, ethidium bromide, and potassium acetate. Riedel-deHaen, Germany: oleic acid. Fluka Biochemika, Swirtzerland: NAD. Promega, U.S.A.: Phage λ DNA, T4 DNA ligase, pGEM®-T Easy, pGEM-3z and TrueBlue vectors. Biosolve LTD: phenol. Saarchem (PTY) LTD: glycerol. Onderstepoort Vet. Research Institute: chicken serum.

3.2 Bacterial strains

Haemophilus paragallinarum NAD^+ independent strains 1742, 1345, F113-3 and NAD^+ dependent strain 0222, NAD^+ independent *Pasteurella avium* strain 737 (all obtained from The Department of Poultry Diseases, Faculty of Veterinary Science, University of Pretoria) were employed in plasmid comparison. *Escherichia coli* Sure®2, for high efficiency cloning of DNA and white-blue selection (Strategene), was used for propagation of the commercial plasmid vectors. The feeder culture *Staphylococcus aureus* (obtained from The Department of Poultry Diseases, Faculty of Veterinary Science, University of Pretoria).

3.3 Growth medium

Haemophilus paragallinarum 0222 was grown on TM/SN: Test Medium (TM) Agar (Rimler, 1979) supplemented with 5% (v/v) oleic-albumin complex, 1% (v/v) heat inactivated chicken serum (Onderstepoort Veterinary Institute) and 0.0025% (w/v) NAD^+ (Fluka). Media was solidified with 1.5% Agar. All media were incubated under increased CO_2 tension (approximately 5%) at 37ºC in a candle jar. Test medium was autoclaved and the supplements filter sterilised with a 0.22 μm filter (Millipore). TM/S media lacking NAD^+ was used for growth of NAD^+ independent *H. paragallinarum* strains. *E. coli* strains were grown on Luria-Bertani medium (LB) (Sambrook *et al.,* 1989) supplemented with 0.02 $\mu g.ml^{-1}$ ampicillin for plasmid selection.

3.4 Polymerase chain reaction identification of *H. paragallinarum*

The bacterial isolates were confirmed as *H. paragallinarum*, using a modification of the HPG-2 PCR method described by (Chen *et al.,* 1996).
The oligonucleotide primers, HP-1F 5'-TGA GGG TAG TCT TGC ACG CGA ATG-3' and HP-1R 5'-CAA GGT ATC GAT CGT CTC TCT ACT-3' were used. The PCR mix contained a single colony of *H. paragallinarum* suspended in sterilized distilled water, MgCl2 (1.5 mM); dNTP`s (0.2 mM); forward and reverse oligonucleotide primer (100 pmol each) and Taq DNA polymerase (1U). The PCR conditions were 25 amplification cycles preceded by a "hot start" and performed in a PCR Thermal Cycler (Perkin-Elmer, USA). The hot start entailed a 10 min-denaturing step at 94ºC followed by the addition of the Taq DNA polymerase. Each of the following 25 cycles consisted of 30s denaturing step at 94ºC, a 50s annealing period at 55ºC and a 45s elongation period at 72ºC.

3.5 Agarose gel electrophoresis

DNA was analysed on a 1% (w/v) agarose gel containing 2.5 mg/µl ethidium bromide. The agarose gel was prepared and electrophoresed in TAE-buffer [0.1 M Tris, 0.05 M Na_2EDTA (pH 8.00) and 0.1 mM glacial acetic acid]. Gel electrophoreses was conducted for 1 hour at 86 V/cm. DNA bands were visualised under low radiation UV- light.

The relative sizes of the DNA fragments were estimated by comparing their electrophoretic mobility with that of the standards, which were run with the samples on each respective gel. Either 1kb Plus marker (Fig. 2a, Life Technologies) or digested λ DNA (Fig. 2b, Promega) was used as standards.

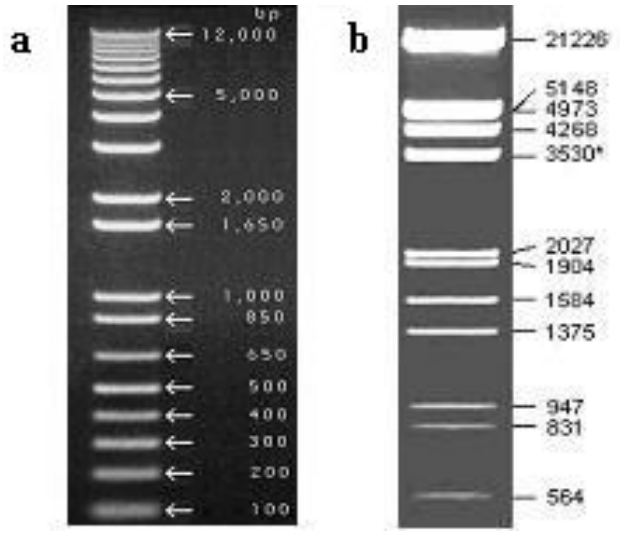

Figure 2. (a) 1kb plus marker (Life Technologies) and (b) λ genomic DNA digested with *Hind*III/*Eco*RI used for determination of relative sizes of DNA fragments during gel electrophoresis.

3.6 Purification of PCR products and restriction enzyme digestions from agarose gels

Purification was achieved using the “NucleoSpin Extract 2 in 1 Kit“ (Amersham) according to the manufacturers instructions.

3.7 Plasmid isolation

H. paragallinarum cells were harvested after 24h incubation in phosphate buffered saline with a hockey stick method. Cells were pelleted by centrifugation at 3000xg. Plasmids were isolated from the cells using an alkaline lysis method. The cells were resuspended in 200μl GTE [50mM glucose, 25mM Tris-CL (pH 8) and 1mM EDTA] and incubated at room temperature for 5 minutes. The cells were then lysed by the addition of freshly prepared 200μl [NaOH/SDS 0.2M NaOH and 1% (w/v) sodium dedocyl sulfate (SDS)] and incubated on ice for 5 minutes. This was followed by the addition of 200μl potassium acetate solution and incubation on ice for 5 minutes. The solution was centrifuged at 10000xg for 15 minutes and the supernatant transferred to a new tube. 600μl Tris-HCL buffered phenol (pH 8.2) was added and the solution vortexed for 1minute, after which it was centrifuged for 10 minutes at maximum speed to extract the DNA. The upper phase was transferred to a new tube, followed by the addition of 600μl chloroform-isoamyl alcohol (24:1 ratio), the mixture was centrifuged for 5 minutes to separate the phases. The upper phase was transferred to a new tube and plasmid DNA was precipitated by addition of 1ml ice-cold absolute ethanol and incubated at –70^{o}C for 20 minutes. The Plasmid was pelleted by centrifugation at 10000xg for 10 minutes and rinsed with 600μl 75% (w/v) ethanol. The pellet was dried under vacuum in a rotary evaporator (Savant, U.S.A.), re-suspended in 40μl TE buffer containing RNase A (pH 8) and incubated at 37^{o}C for 30 minutes. The same method was used for screening recombinant plasmids in *E. coli* cells.

3.8 Plasmid Sonification

DNA was suspended in 500 μl TE – Buffer (10mM Tris, pH 7.8 and 1mM EDTA) followed by sonification performed on ice at 300 W. Sonification was done in cycles of 6 bursts of 10s with a 10s delay between bursts. The sonified DNA was precipitated with 100% EtOH after which it was dissolved in 60 μl TE – buffer.

3.9 DNA ligations

Fragments obtained from genomic DNA digestions, sonifications and PCR products were purified from agarose gels and the desired vector plasmid and DNA combinations were made. Ligations were performed in a total volume of 23 μl containing 10x ligation buffer and 1U DNA ligase (Promega) and incubated at 16^{o} C overnight. PCR products were ligated into pGEM-T Easy vector (Promega) at 4^{o} C according to the manufacturers specifications.

3.10 Preparations of bacterial cells for transformation

Competent *E. coli* cells were prepared according to the method described by Tang *et al*. (1994).

Briefly, the desired *E. coli* strain was cultured overnight at 37^{o}C in 5ml LB broth; 1ml was transferred to 100ml LB broth and allowed to grow to an optimal density (OD_{600}) of 0.9-0.95. The cells were harvested by centrifugation at 4000 g for 5min at 4^{o}C and the pellet re-suspended in 10ml of a solution containing $CaCl_2$ (80mM) and $MgCl_2$ (50mM). The mixture was left on ice for 10min and centrifuged at 4000 g for 5min at 4^{o}C. This was repeated twice with the pellet finally being re-suspended in $CaCl_2$ (0.1M) and glycerol [50% (v/v)] the suspensions were aliquot into 80 μl volumes, snap-frozen using liquid nitrogen and stored at -70^{o}C (Tang *et al.,* 1994).

3.11 Bacterial Transformation

Plasmid DNA (1µg) containing the desired inserts was added to competent *E. coli* cells and left on ice for 30 min then incubated at 42^{o}C for 90 seconds and placed on ice for 4 min. A 40 % (w/v) glucose containing LB medium (800 µl) was added to the suspension and incubated for 1 hour at 37^{o}C. Cells were harvested by centrifugation at 4000 g for 2 min, re-suspended in LB media, plated out on LB agar containing IPTG (20 mg/ml), X-gal (20mg/ml) and ampicillin (30 mg/ml).

3.12 Sequencing

Sequencing was performed using the ABI Prism BigDye™ Terminator v3.0 cycle Sequencing Kit (Applied Biosystems) together with commercial available primers (SP6 or T7) or primers specifically designed according to known sequences obtained from the purified plasmid DNA (Table 3). The sequencing reactions were analyzed on an ABI PRISM™ 377 Automatic DNA sequencer. The DNA sequences were analyzed by using the ABI PRISM™ 377 Automatic DNA Sequencer software.

Table 3.Lists the primers that were designed for the amplification and sequencing of the purified plasmid DNA.

Primer name	Primer Sequence
KSQR	TGT TAT CTC TGC TTA TCA AAA TGC T
KSQF	AAC TAA TTT GGC ATC TTC TAC ATC T
PHP-3R	CTT CGC TTA TGA GCT TAA CCG G
HAEM	CTT CTT TTA ATA CGA CGG GAA ACT
LSF	CTG AAT CAG ACT GGA ATA AAT CTA T
KSF	AGA TGT AGA AGA TGC CAA ATT AGT T
PHP-R1	AAT GAC TTA AAC ATG CTT GTA
PHP F1	CAT ACA CAA CTA TTC TCT GG
KSR	AGC ATT TTG ATA AGC AGA GAT AAC A
PKF 2	CTA GGC GTA AAA AC ACCT TCG CTA A
PHP 4R	CTT GAA AAG TAA AGG GGG ATC G
PHP 4F	TTT ATT AAT TCA AAC CTT GGA ACG G
PKR1	GCC GAA CCA CAA CAC GTA AAA CTA
PKF1	AAT GTT TGG GAG AGA AAA ACG TGC T
PHP R2	AGG GAA CTT ATT CTA TAA GAC

3.13 Enzymatic Activity Assay for NAD^+ synthetase

Cells were grown overnight on BTA plates and harvested by washing the plates with Tris-HCL buffer (pH 8.5). The total volume consisted of 5 ml cells, which was treated with 5 mg lysozyme on ice for 30 min. The cells were then sonified on ice in 6 cycles of 10 s bursts with a 10 s delay between bursts at 200-300 W with a microtip. NAD^+ synthetase activity was determined by measuring the rate of NAD^+ formation at 37^{o}C. The reaction mixture for standard assays, except when otherwise stated, contained 1 mM deamido NAD^+ (N^aAD), 2 mM ATP, 20 mM L-glutamine, 1 mM $(NH_4)_2SO_4$, 5 mM $MgCl_2$, 56 mM KCl, 0.04% BSA, 50 mM Tris-Cl buffer (pH 8), and enzyme in a total volume of 0.25 to 0.5 ml. The reaction mixture was incubated for 30 or 60 min. After incubation, the reaction was terminated by heating in a boiling bath for 30 to 60 s and the reaction mixture was centrifuged at 800 x g for 2 min. NAD^+ was determined on the clear supernatant material by employing alcohol dehydrogenase.

Chapter 4

Results and Discussion

4.1 Growth of bacterial strains

Freeze-dried isolates were reconstituted on blood tryptose agar plates. Growth of the NAD^+ independent *H. paragallinarum* strains 1742, 1345, F113-3 and *Pasteurella avium* strain 737 were obtained on both brain heart infusion (BHI) agar and blood tryptose agar (BTA). NAD^+ dependent strain 0222 was grown at 37^{o}C on test media (TM) plates supplemented with NAD^+ and chicken serum. The isolates were confirmed to be *H. paragallinarum* by using gram staining and PCR test. The *Pasteurella avium* strain was partially identified by gram stain and morphology. *Pasteurella avium* strain 737 was used as a negative control in the PCR test. The PCR reaction using primers HP-1F and HP-1R specific for *H. paragallinarum* yielded one band of approximately 500 bp when separated on a 1% agarose gel (Fig. 3).

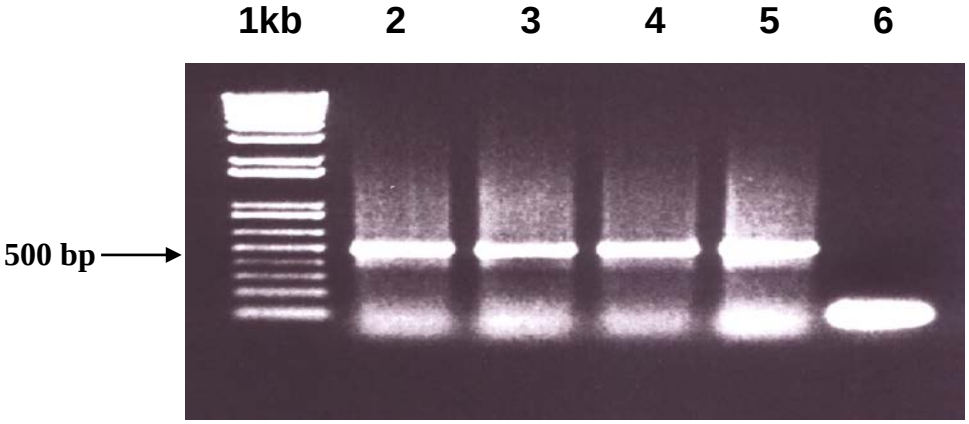

Figure 3. A 1% agarose gel stained with ethidium bromide showing results obtained from the PCR specific for *H. paragallinarum*. Lane 1, 1 kb marker. Lane 2, *H. paragallinarum* strain 1742. Lane 3, *H. paragallinarum* strain 1345. Lane 4 *H. paragallinarum* strain F113-3. Lane 5 *H. paragallinarum* strain 0222, Lane 6 negative control.

4.2 Plasmid Isolation

Haemophilus and *Pasteurella* genera are found in the *Pasteurellaceae* family, it is therefore possible that NAD^+ independence in these organisms is conferred through a similar mechanism contained on a plasmid entity. To enable comparison between the modes of NAD^+ independence active in the different *H. paragallinarum* strains, *P. avium* strain 737 was included in plasmid extraction experiments.

After 24h incubation, cells were harvested and suspended in phosphate buffered saline using a hockey stick method. Plasmid DNA was isolated with an alkaline lyses method. Plasmid DNA was successfully isolated from all three of the independent *H. paragallinarum* strains used in this study (1742, 1345, F113-3). No plasmid could be extracted from *P. avium* strain 737 (Fig. 4, lane2). A possible explanation for this phenomenon could be that the plasmid might have been incorporated into the genomic DNA. The possibility that the mechanism of independence between *H. paragallinarum* and *P. avium* might differ is not excluded. As expected no plasmid could be isolated from the NAD^+ dependent *H. paragallinarum* strain 0222 (Fig. 4, lane1) which acted as a negative control.

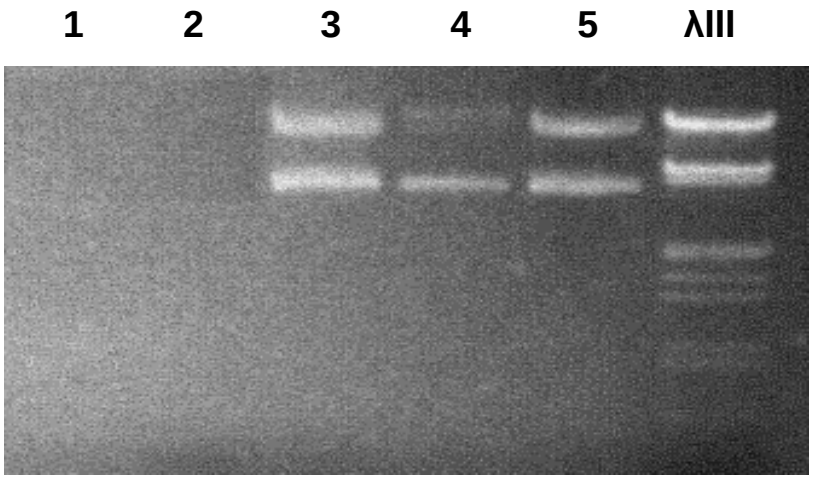

Figure 4. A 1% agarose gel stained with ethidium bromide indicating extracted plasmid DNA. Lane 1) *H. paragallinarum* strain 0222 (NAD dependent), lane 2) *P. avium* strain 737, lane 3) *H. paragallinarum* strain 1742, lane 4) *H. paragallinarum* strain 1345, lane 5) *H. paragallinarum* strain F113-3, lane 6) phage λ DNA digested with *EcoR*1 and *Hind* III.

4.3 Restriction mapping

The plasmid DNA from the three independent *Haemophilus* strains studied was compared by restriction enzyme analyses with a variety of 16 different restriction enzymes. Restriction analysis revealed that the plasmids from the three different strains used in this study were identical (Fig. 5). Plasmid DNA could successfully be digested with six of the 16 enzymes tested. The 16 enzymes which were used are: *EcoRI, BamHI, NotI, PstI, MboI, XhoI, SspI, SalI, SacI, SmaI, XbaI, Bsty*I, *Hae*III, *Xcm*I, *Hind*III and *Bgl*II. The six enzymes, which digested the plasmid as well as the fragment sizes which where obtained on a 1% agarose gel, are shown in Table 4.

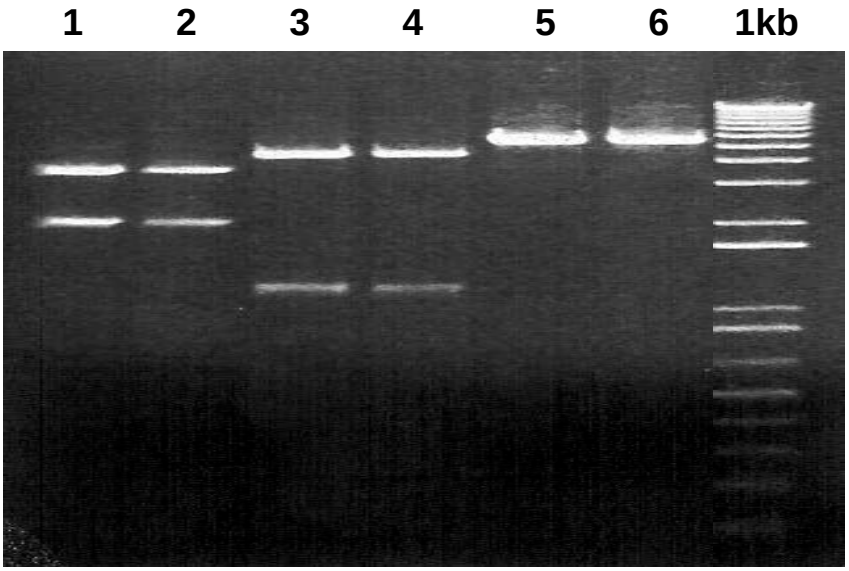

Figure 5. A 1% agarose gel stained with ethidium bromide showing the digestion of the plasmid from *H. paragallinarum* strain 1345 and *H. paragallinarum* strain F113-3 respectivly with three restriction enzymes. Lane 1-2 digestion with *Bsty*I. Lane 3-4 digestion with *Hind*III. Lane 5-6 digestion with *Hae*III. Lane 7 1kb marker.

Table 4. Restriction enzymes with the number of fragments obtained and their approximate sizes.

	*Bgl*II	*Hae*III,	*Xcm*I,	*Bsty*I,	*Hind*III	*XbaI,*
Number of fragments	1.0	1.0	1.0	2.0	2.0	2.0
Fragment sizes (kb)	6.0	6.0	6.0	2.0 & 4.0	1.5 & 4.5	1.7 & 4.3

The six enzymes capable of digesting the plasmid, were used in different combinations with the results obtained listed in Table 5.

Table 5. Restriction enzyme combinations and their approximate fragment sizes.

Restriction enzyme combinations	Fragment sizes (kb)
*Xcm*I – *Hae*III	0.5; 5.5
*Bsty*I – *Hae*III	0.55; 2; 3.45
*Bsty*I - *Hind*III	0.25; 2; 1.5; 2.25
*Hae*III - *Hind*III	1.5; 1.7; 2.8
*Bsty*I – *Xcm*I	1.05; 2; 2.95
*Xcm*I – *Hind*III	1.5; 1.2; 3.3
*Hae*III – *Bgl*II	2.5; 3.5
*Xcm*I – *Bgl*II	3; 3
*XbaI – Bsty*I	2; 0.5; 1.7; 1.8

Using the results obtained from the restriction analysis with single as well as a combination of the different restriction enzymes (Table 4 and 5) an initial restriction map was constructed (Fig. 6).

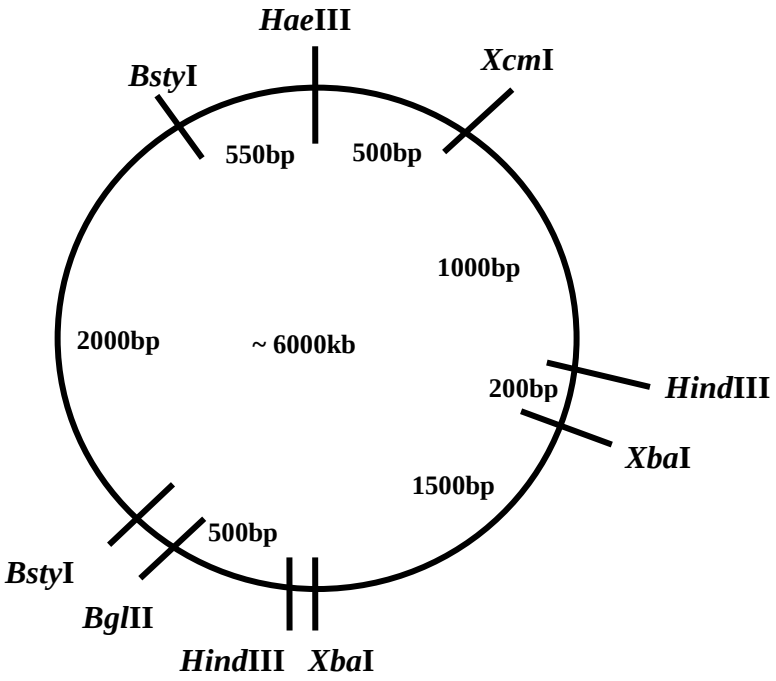

Figure 6. Restriction map of plasmid(s) isolated from NAD^+-independent *H. paragallinarum.*

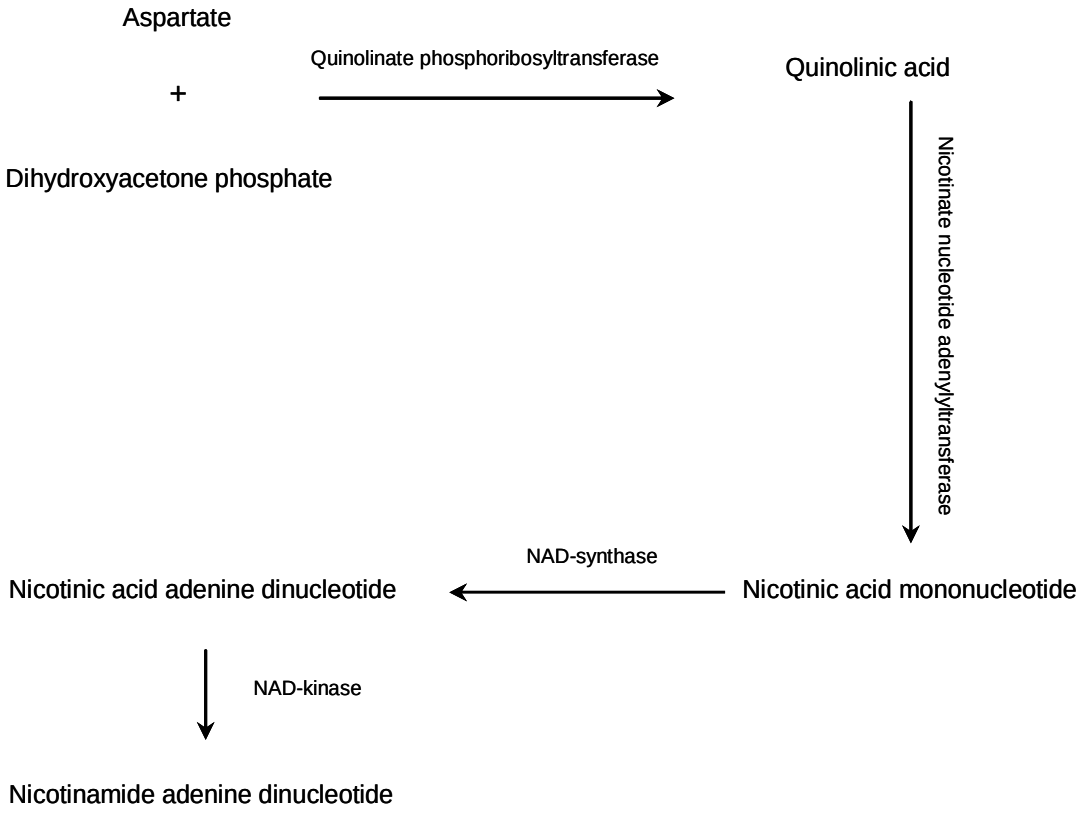

Figure 7. The Preiss-Handler pathway showing NAD metabolism in aerobic bacteria

4.4 NAD^+ Synthetase

The completion of the above mentioned section prompt the question of the precise mode by which the plasmid confers NAD^+ independence and which enzymes may be encoded on the plasmid. Therefore the synthesis of NAD^+ by NAD^+ dependent strain 0222 of *H. paragallinarum* was investigated. This strain was cultured on test media agar plates supplemented with 1% NAD^+, Nicotinic acid, Tryptophane or nicotine amide respectively.

Growth was supported by NAD^+ and Nicotine amide but no growth was observed on plates supplemented with Nicotinic acid and Tryptophane. Closer investigation of the metabolic pathway in question (Fig 7) identified the possible involvement of NAD^+ synthetase, which might aid in conferring NAD independence by aiding the additional synthesis of the co-enzyme. It was thus thought that the plasmid might encode for the enzyme NAD^+ synthetase.

NAD^+ synthetase catalyses the final reaction in both the *de novo* and salvage pathways during the biosynthesis of NAD^+ (Fig. 8). In addition it is considered as an essential factor during germination, overgrowth and survival of bacteria under stressful conditions. NAD^+ synthetase is a pyrophosphatase, the enzyme is a 60 kDa homodimer and the reaction catalysed proceeds in two steps. The reaction involves binding of three substrates to the active site of NAD^+ synthetase. The first is ATP that provides the energy used to drive the reaction, the second is the true substrate diamido-NAD^+ and the third is a combination of ammonium as well as two Mg^{2+} ions (Devedjiev *et al*., 1998).

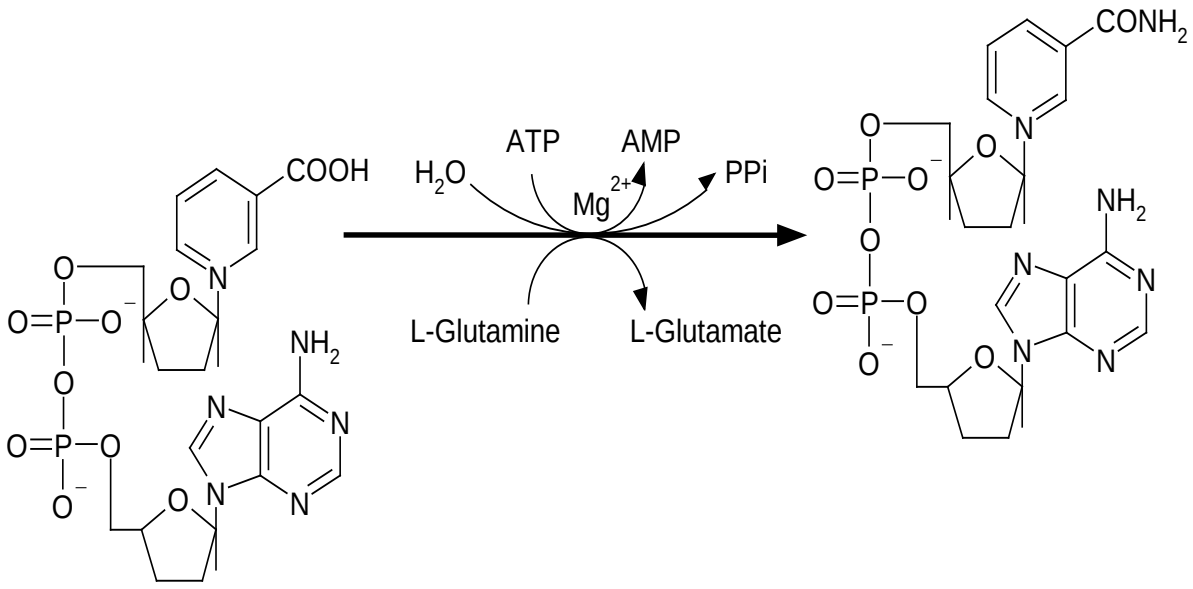

Figure 8. Conversion of deamido-NAD^+ to NAD^+ catalysed by NAD^+ synthetase

Both NAD^+ dependent and independent strains of *H. paragallinarum* were included in the experiment with *Staphyloccus aureus* as positive control, since this bacterium produses NAD^+ in excess. This experiment was conducted to determine the possible presence of NAD^+ synthethase. This determination was done indirectly using a coupled assay, in which NAD^+ acted as co-factor for alcohol dehydrogenase. Activity of commercial NAD^+ synthetase (Sigma-Aldrich) was used as standard. NAD^+ was produced by NAD^+ synthetase in lyophilised cell extracts to which de-amino NAD^+ was added in excess. The reaction mixture was incubated at 37°C for 60 minutes after which enzyme

activity was terminated by boiling the reaction mixture for 1 minute. The NAD^+ formed was converted to NADH by alcohol dehydrogenase (ADH) in the presence of an excess ethanol. The amount of NADH formed was determined spectrophotometrically at 340 nm.

Since the formation of NAD^+ from de-amino NAD^+ is specific to NAD^+ synthetase the presence of this enzyme in NAD^+ dependent and independent strains of *H. paragallinarum* could be deduced. Enzyme activity was calculated as follows:

$$\text{Units.ml}^{-1} = \frac{(\Delta\Delta_{340\text{nm}}\ \text{Test} - \Delta A_{340\text{nm}}\ \text{Blank})(0.65)}{(60)(6.22)(0.2)}$$

Where :

$\Delta A_{340\text{nm}} \equiv$ Absorbance at 340 nm

$0.65 \equiv$ Total volume (in milliters) of assay

$60 \equiv$ Time (in minutes) of assay as per the unit definition

$6.22 \equiv$ millimolar extinction coefficient of NADH at 340 nm

$0.2 \equiv$ Volume (in milliliter) of enzyme used

Unit definition: 1 unit will form 1 µmol of NAD^+ from de-amino NAD^+ per minute at pH 8.5 at 37°C

In terms of the initial hypothesis it was expected that NAD^+ independent strains would display higher activity since the plasmid might contribute to the presence of this protein, whereas the NAD^+ dependent strains would exhibit no syntethase activity.

Activity of commercial NAD^+ synthetase was use as comparative standard in (Fig. 9). The initial time suggested by the manufacturer was 2 minutes per reaction but since we had very low biomass and extraction was necessary we evaluated the time interval and found that increasing the reaction time would increase the absorbance value proportionally, thus making detection of small changes of absorbance more feasible.

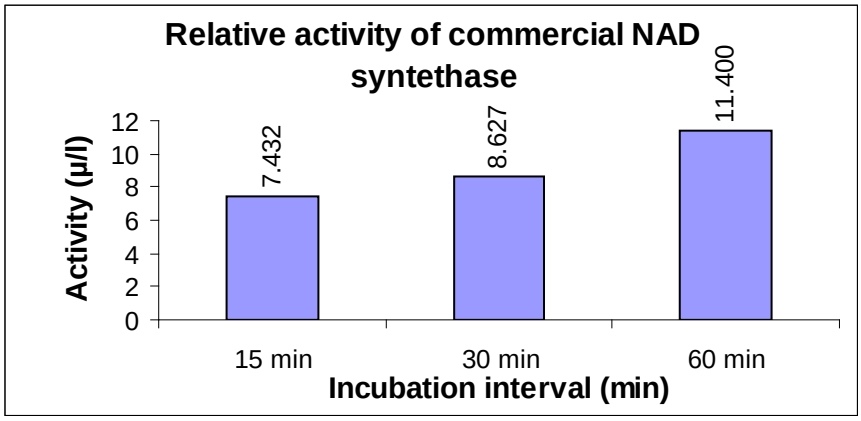

Figure 9. Activity of NAD^+ synthetase

Once the enzyme reaction was standardized the commercial enzyme was substituted by an extracted enzymatic preparation from the *H. paragallinarium* cells. Table 6 displays the activity obtained for the respective NAD^+ synthetase in both NAD^+ dependent and NAD^+ independent strains, although the NAD^+ independent strain exhibited slightly higher production level of NAD^+ it showed no conclusive evidence that the plasmid conferred exclusively for this enzyme and thus did not answer the question of the role of the plasmid in conferring NAD independence.

Table 6. Comparison of NAD^+ synthetase activity (units/µl) in NAD^+ dependent and independent strains of *H. paragallinarum*

Staphylococcus aureus	*H. paragallinarum* NAD^+ dependent	*H. paragallinarum* NAD^+ independent
0.957	0.811	1.003

These results also eliminate the necessity of NAD^+ synthethase encoding on the plasmid as the source of NAD^+ independence. Instead it may be deduced that NAD^+ synthethase is genomically encoded in both NAD^+ dependent and independent forms of *H. paragallinarum.* This was later confirmed through the full sequence of the plasmid.

4.5 Subcloning of plasmid DNA

The initial restriction map of this plasmid was used to direct the subcloning of fragments of the plasmid. Plasmid DNA was digested with *Hind*III for subcloning. Digestion yielded two fragments, a small fragment of approximately 1500bp and a larger fragment of about 4500bp (Fig. 10). These two fragments were cloned into the *Hind*III digested vector pGEM-3Z and subsequently transformed into *E. coli.* Transformation was confirmed through blue white selection on LB agar plates containing ampicilin, IPTG and X-gal. Even though this was numerously repeated, only the small fragment was successfully subcloned (Fig. 9).

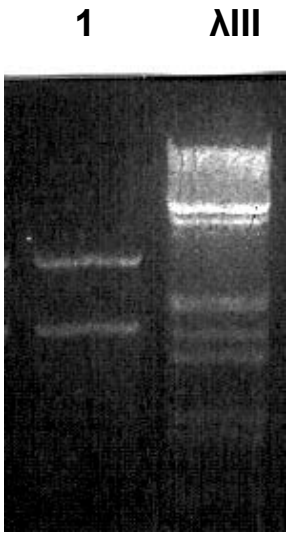

Figure 10. A 1% agarose gel stained with ethidium bromide showing in lane 1, *Hind III* digested pGEM 3z vector containing the 1.5 kb plasmid fraction (lower band), lane 2: phage λ DNA digested with *EcoR*1 and *Hind*III.

Further attempts to sub-clone the remaining part of the plasmid failed, even though various combinations of restriction enzymes and cloning vectors were evaluated. It was therefore decided to revert to shredding of the plasmid DNA through sonification followed by shotgun sequencing of the clones obtained.

4.6 Sonfication of plasmid DNA.

Plasmid DNA was shredded by sonification until only a smear was visible on an agarose gel (Fig. 11). The single stranded DNA overhangs present due to shredding were filled using klenow polymerase enzyme yielding blunt ending DNA fragments. Shredded λ-DNA was used as a positive control in this experiment.

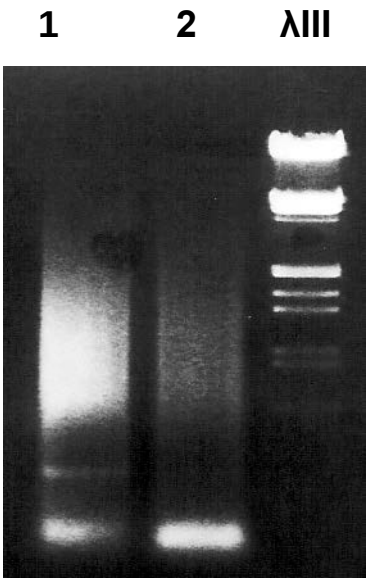

Figure 11. A 1% agarose gel stained with ethidium bromide, lane 1) shredded plasmid DNA visible as a smear, lane 2) shredded phage λ DNA, lane 3) Phage λ DNA digested with *EcoR*1 and *Hind*III.

4.7 Subcloning of shredded DNA.

The klenow treated DNA fragments were ligated into a pGEM-3Z vector, which were pre-digested with *Sma*I (restriction enzyme yielding blunt ends). These recombinant vectors DNA were then transformed into competent *E. coli* Sure®2 cells. Transformation was confirmed by blue white selection on LB agar plates (supplemented with ampicilin, IPTG and X-gal). A very low subcloning success rate was observed with *H. paragallinarum* plasmid fragments. Recombinant plasmid DNA was only found in approximately 1 out

of every 30 white colonies recovered. In comparison, all white colonies sub-cloned with λ DNA, as positive control, contained an insert. The problem therefore seemed to be unrelated to the cloning vector and is likely related to some structural occurrence found in *H. paragallinarum* DNA. The useful transformants yielded clones of various sizes and are shown in Fig. 12.

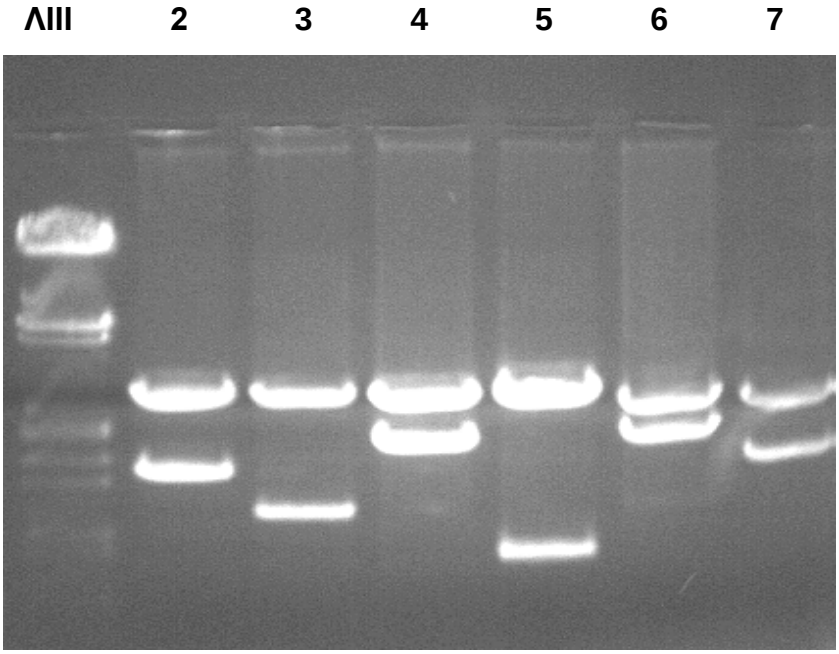

Figure 12. A 1% agarose gel stained with ethidium bromide illustrating the variously sized clones that were obtained from sonification. The upper bands in all lanes represent the pGEM-3Z vector.

4.8 Sequencing of clones.

All clones including the 1.5 kb clone obtained from digestion with *Hind*III were subjected to sequenced using the Sp6 and T7 plasmid primers. The sequences obtained were aligned using Auto Assembler software (Applied Biosystems). Remaining gaps in the sequence were filled through the design of synthetic oligonucleotide primers based on the available sequences. For larger gaps flanking primers were used and the resulting PCR fragment sub-cloned into vector p-GEM-T-easy. Success rate for sub-cloning was still extremely low. Only 1 in 30 white colonies contained inserts, similar to the number of colonies successfully sub-cloned with extracted plasmid. Since

DNA used in this sub-cloning was a PCR product the possibility that the DNA from *H. paragallinarum* is modified in a way hindering sub-cloning was now excluded. Sequences obtained through the use of all available cloned products as well as direct sequencing of PCR products allowed the assembly of the complete sequence of the plasmid DNA.

The complete sequence of the plasmid was found to be 6095 base pairs in length and had a G+C content of 34.4%. This translates into an A+T content of 65.6%. The difficulties experienced in sub-cloning the plasmid may therefore be the result of high AT content. Cloning of AT rich DNA has been found to be extremely problematic. (Razin *et al*., 2001) It is possible that the AT rich cloned sequences behave as transcriptional promoters in *E. coli.* Transcription driven by the insert may then proceed into the vector and interfere with its replication or expression of drug resistance. Furthermore, in an attempt to generate genomic libraries of *Lactobacillus helveticus* assembly of 19000 clones, considered sufficient to provide 4x coverage of the genome, only 70% of the genome was represented. The genome of this organism contains 65% AT. Significant numbers of deleted clones with inserts smaller than 200bp were generated. (Ronald Godiska, Lucigen corp. personal communication)

DNA with a high AT content tends to be more susceptible to sonification fragmentation than DNA with average nucleotide content or GC rich DNA. Sonification of AT rich DNA therefore leads to the formation of large amounts of very small fragments. Such fragments, i.e. smaller than approximately 200 base pairs, are usually not detected when using a standard agarose gel size and concentration (i.e. 1% w/v). This may lead to white colonies with small inserts being designated as false positives, and effectively reduces the observed subcloning success rate. The difficulties in subcloning of AT rich DNA are known and vectors specifically designed to overcome this are commercially available.

4.9 Characterization of the complete sequence of the plasmid encoding NAD^+ independence in *H. paragallinarum*.

The complete sequence of the plasmid encoding NAD^+ independence in *H. paragallinarum* was found to be 6095bp in length. The complete sequence of the plasmid encoding NAD^+ independence in *H. paragallinarum* was used to search sequence databases. Homologues identified in the bacterial genome databases were found from species including *H. influenzae*, *Pasteurella multocida* and *Bacillus anthracis*. The plasmid was found to encode six different proteins (Table 7).

Table 7. Showing the six proteins found on the plasmid

Protein	**Size of open reading frame (amino acids)**
Haemocin immunity	105
hmcD	297
Haemocin structural	90
Haemocin transporter	235
Quinolinic acid phosphoribosyl transferase	372
Replication protein A	321

Fig. 13 represents the relative locations of different restriction sites as well as proteins encoded on the plasmid.

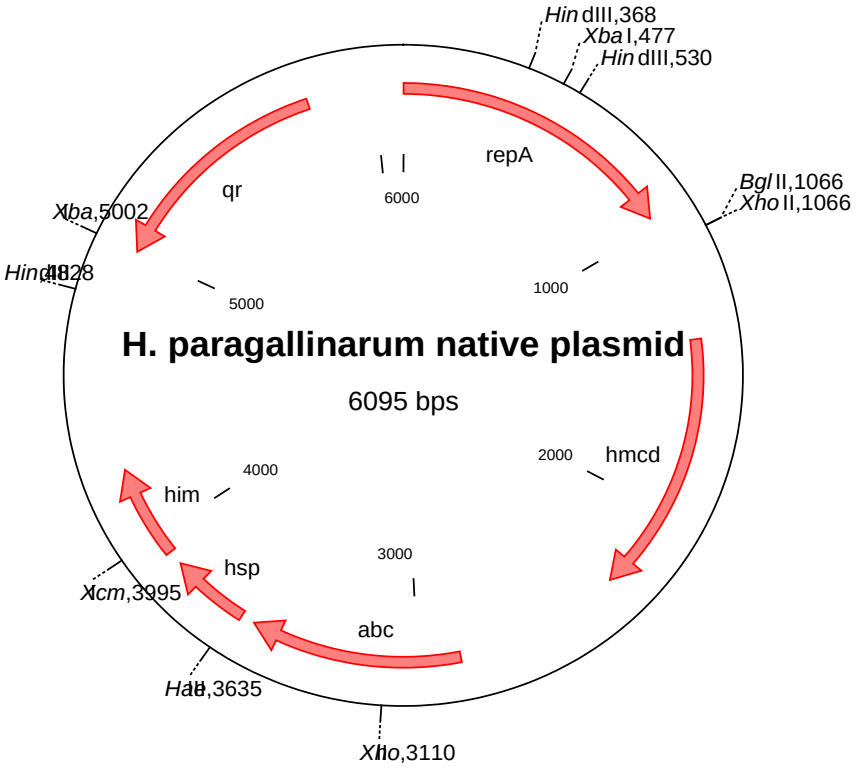

Figure 13. Restriction map of the plasmid from *H. paragallinarum* NAD^+ independent strains. Abbreviations used for the proteins encoded on the plasmid are as follows: repA – replication protein A; hmcD – haemocim putative protein; abc – haemocim transporter; hsp – haemocim structural protein; him – haemocim immunity protein; qr – quinolinate phosphoribosyl transferase.

The complete sequence of the plasmid DNA was translated in all possible reading frames to locate the putative proteins encoded from the plasmid entity (Fig. 14).

```
M  V  D  D    V  C  A    V  K  F    K  T  I  A    K  A  S    Y  R  L
10           20           30           40           50           60
ATGGTTGATG   ACGTTTGTGC   CGTTAAGTTT   AAGACTATCG   CAAAGGCGTC   CTACCGATTA
TACCAACTAC   TGCAAACACG   GCAATTCAAA   TTCTGATAGC   GTTTCCGCAG   GATGGCTAAT

T  L  D  G    A  V  K    K  L  A    S  G  G  T    M  K  P    K  T  H
70           80           90          100          110          120
ACATTGGACG   GTGCTGTCAA   GAAGTTAGCC   TCAGGGGGGA   CGATGAAGCC   TAAAACTCAC
TGTAACCTGC   CACGACAGTT   CTTCAATCGG   AGTCCCCCCT   GCTACTTCGG   ATTTTGAGTG

I  A  V  F    E  L  S    L  C  K    Y  N  R  Q    F  P  G    E  L  I
130          140          150          160          170          180
ATAGCAGTTT   TTGAATTGTC   GTTGTGCAAG   TACAATCGCC   AGTTCCCTGG   TGAATTGATA
TATCGTCAAA   AACTTAACAG   CAACACGTTC   ATGTTAGCGG   TCAAGGGACC   ACTTAACTAT

E  K  A  Y    S  Q  A    R  T  K    R  E  R  L    V  D  R    W  V  H
190          200          210          220          230          240
GAGAAGGCTT   ATAGTCAGGC   GAGAACGAAG   CGTGAAAGGT   TAGTAGACCG   CTGGGTGCAC
CTCTTCCGAA   TATCAGTCCG   CTCTTGCTTC   GCACTTTCCA   ATCATCTGGC   GACCCACGTG

P  E  D  E    K  H  V    T  K  F    A  W  A  S    S  E  S    F  Y  K
250          260          270          280          290          300
CCCGAAGACG   AGAAACATGT   TACTAAGTTC   GCCTGGGCGA   GTAGCGAGAG   CTTTTACAAG
GGGCTTCTGC   TCTTTGTACA   ATGATTCAAG   CGGACCCGCT   CATCGCTCTC   GAAAATGTTC

N  C  G  K    F  Q  S    V  G  C    R  E  I  M    P  H  V    T  Q  L
310          320          330          340          350          360
AATTGTGGCA   AGTTCCAATC   TGTCGGGTGT   CGGGAGATAA   TGCCACACGT   TACCCAGTTA
TTAACACCGT   TCAAGGTTAG   ACAGCCCACA   GCCCTCTATT   ACGGTGTGCA   ATGGGTCAAT

R  A  Q  A    L  Q  Y    Q  L  K    H  V  S  N    L  C  S    V  H  E
370          380          390          400          410          420
CGTGCGCAAG   CTTTGCAATA   CCAGTTAAAG   CACGTATCTA   ACTTGTGCTC   TGTCCATGAG
GCACGCGTTC   GAAACGTTAT   GGTCAATTTC   GTGCATAGAT   TGAACACGAG   ACAGGTACTC

N  R  L  Y    E  L  F    T  Q  Y    R  R  V  K    L  R  V    I  N  L
430          440          450          460          470          480
AATCGTTTAT   ACGAGTTATT   TACACAATAT   AGGCGTGTTA   AGTTGCGTGT   AATCAATCTA
TTAGCAAATA   TGCTCAATAA   ATGTGTTATA   TCCGCACAAT   TCAACGCACA   TTAGTTAGAT

E  D  L  R    D  R  L    Q  V  K    D  K  Y  P    T  F  K    A  F  N
490          500          510          520          530          540
GAAGATTTGA   GAGATCGGTT   ACAAGTTAAA   GATAAATACC   CTACATTCAA   AGCTTTTAAT
CTTCTAAACT   CTCTAGCCAA   TGTTCAATTT   CTATTTATGG   GATGTAAGTT   TCGAAAATTA

Q  W  V  I    K  P  A    I  K  E    I  N  E  K    S  H  L    K  V  E
550          560          570          580          590          600
CAATGGGTAA   TTAAGCCTGC   TATCAAGGAA   ATCAATGAGA   AATCACATCT   GAAAGTGGAA
GTTACCCATT   AATTCGGACG   ATAGTTCCTT   TAGTTACTCT   TTAGTGTAGA   CTTTCACCTT

Y  D  T  I    A  L  R    R  A  V    A  A  L  L    F  T  V    T  A  E
610          620          630          640          650          660
TATGACACAA   TAGCTTTAAG   ACGTGCTGTT   GCCGCATTGC   TTTTCACTGT   AACCGCGGAA
ATACTGTGTT   ATCGAAATTC   TGCACGACAA   CGGCGTAACG   AAAAGTGACA   TTGGCGCCTT
```

```
K  P  V  K    K  V  P    K  F  P    H  K  N  K    Y  G  K    F  V  K
670          680          690          700          710          720
AAACCCGTTA   AAAAAGTCCC   AAAATTCCCA   CACAAAAACA   AGTACGGTAA   GTTTGTGAAG
TTTGGGCAAT   TTTTTCAGGG   TTTTAAGGGT   GTGTTTTTGT   TCATGCCATT   CAAACACTTC

L  D  R  I    D  P  K    M  S  S    A  E  Y  G    S  Y  V    R  D  C
730          740          750          760          770          780
CTAGATCGTA   TCGATCCTAA   AATGAGTTCG   GCTGAATACG   GCAGTTACGT   AAGAGATTGC
GATCTAGCAT   AGCTAGGATT   TTACTCAAGC   CGACTTATGC   CGTCAATGCA   TTCTCTAACG

L  K  I  L    E  D  F    Y  S  N    I  E  D  V    P  N  E    D  L  L
790          800          810          820          830          840
CTGAAAATCC   TTGAAGATTT   CTATTCAAAC   ATTGAAGACG   TACCGAATGA   AGATTTGCTT
GACTTTTAGG   AACTTCTAAA   GATAAGTTTG   TAACTTCTGC   ATGGCTTACT   TCTAAACGAA

Y  Y  W  I    F  L  A    V  N  Q    S  H  K  S    K  L  G    S  K  N
850          860          870          880          890          900
TACTACTGGA   TTTTTCTTGC   GGTAAACCAA   AGCCATAAAT   CAAAACTAGG   CAGTAAAAAC
ATGATGACCT   AAAAAGAACG   CCATTTGGTT   TCGGTATTTA   GTTTTGATCC   GTCATTTTTG

T  F  A  M    S  L  T    W  L  N    E  C  W  L    N  T  W    L  T  N
910          920          930          940          950          960
ACCTTCGCTA   TGAGCTTAAC   GTGGCTAAAT   GAATGTTGGT   TGAATACCTG   GTTAACTAAT
TGGAAGCGAT   ACTCGAATTG   CACCGATTTA   CTTACAACCA   ACTTATGGAC   CAATTGATTA

*
970          980          990          1000         1010         1020
TAAGTTGGTT   GTGAACTGGT   TAACTTGAAA   AGTAAGGTGA   TCGGGCATAC   AATGGAATTA
ATTCAACCAA   CACTTGACCA   ATTGAACTTT   TCATTCCACT   AGCCCGTATG   TTACCTTAAT

1030         1040         1050         1060         1070         1080
GGTCGGGGGC   AGGTTGCAAT   CCTACCCCCT   TCCTAGCCGT   TAGGCAGATC   TATTTAACAA
CCAGCCCCCG   TCCAACGTTA   GGATGGGGGA   AGGATCGGCA   ATCCGTCTAG   ATAAATTGTT

1090         1100         1110         1120         1130         1140
GTAAACTAGC   AAATTAAGCA   TTGCGAGTGC   TTTTAAGCTA   GCCGTAACAG   CTAGTAAAAT
CATTTGATCG   TTTAATTCGT   AACGCTCACG   AAAATTCGAT   CGGCATTGTC   GATCATTTTA

1150         1160         1170         1180         1190         1200
TAGCGTTACT   ATCAGGAATA   ATTCTGATAC   TTTCACATAA   ATCACCTCCT   TAGCACTACA
ATCGCAATGA   TAGTCCTTAT   TAAGACTATG   AAAGTGTATT   TAGTGGAGGA   ATCGTGATGT

1210         1220         1230         1240         1250         1260
AGGCTAAGCG   GCTAGGCTCA   CACTATTAGC   GGTAGTGCGA   GCCGCCGCCA   TAGTGCGAGA
TCCGATTCGC   CGATCCGAGT   GTGATAATCG   CCATCACGCT   CGGCGGCGGT   ATCACGCTCT

1270         1280         1290         1300         1310         1320
ATGATTTTAG   CATAGAACCG   CCTAAAAGGG   CGGTTTTTGT   TTTATGTAAA   CCCTTGTCGG
TACTAAAATC   GTATCTTGGC   GGATTTTCCC   GCCAAAAACA   AAATACATTT   GGGAACAGCC

1330         1340         1350         1360         1370         1380
GCGGGCAACT   TGTTGCACGC   TTGACAAGGG   TTGAATATGT   GACTTGTTTG   TTTTTTGACA
CGCCCGTTGA   ACAACGTGCG   AACTGTTCCC   AACTTATACA   CTGAACAAAC   AAAAAACTGT
```

```
                                M  Y  N  I     N  L  L     N  Y  I     R  L  Y  M
1390          1400          1410          1420          1430          1440
AGGAGGGAGA    AACTATAGTA    ATGTATAATA    TAAATCTATT    AAACTATATC    AGGCTATATA
TCCTCCCTCT    TTGATATCAT    TACATATTAT    ATTTAGATAA    TTTGATATAG    TCCGATATAT

K  T  N     I  L  F     I  F  I  F     N  P  I     L  C  Y     S  N  N  D
1450          1460          1470          1480          1490          1500
TGAAAACCAA    CATATTATTT    ATATTTATAT    TCAACCCCAT    ACTGTGTTAT    TCTAATAATG
ACTTTTGGTT    GTATAATAAA    TATAAATATA    AGTTGGGGTA    TGACACAATA    AGATTATTAC

N  S  I     T  L  N     D  E  F  D     N  F  G     N  F  T     L  V  T  S
1510          1520          1530          1540          1550          1560
ATAATTCTAT    TACATTAAAT    GATGAATTTG    ACAACTTTGG    TAATTTCACA    TTAGTAACAT
TATTAAGATA    ATGTAATTTA    CTACTTAAAC    TGTTGAAACC    ATTAAAGTGT    AATCATTGTA

S  S  I     H  F  K     N  S  Q  E     F  L  R     R  N  V     I  D  K  Q
1570          1580          1590          1600          1610          1620
CATCTTCAAT    ACATTTTAAA    AACAGCCAAG    AATTTTTAAG    AAGAAATGTA    ATTGATAAAC
GTAGAAGTTA    TGTAAAATTT    TTGTCGGTTC    TTAAAAATTC    TTCTTTACAT    TAACTATTTG

N  Y  D     N  S  Y     G  S  E  K     Y  Q  G     L  S  F     E  Q  G  I
1630          1640          1650          1660          1670          1680
AAAATTATGA    CAATTCATAT    GGAAGTGAAA    AATATCAAGG    ACTATCCTTT    GAGCAAGGGA
TTTTAATACT    GTTAAGTATA    CCTTCACTTT    TTATAGTTCC    TGATAGGAAA    CTCGTTCCCT

Q  Y  G     L  T  N     N  I  N  M     S  D  S     I  N  S     F  Y  S  K
1690          1700          1710          1720          1730          1740
TTCAGTATGG    CTTAACAAAT    AACATTAATA    TGTCTGATAG    TATTAATAGT    TTTTACAGTA
AAGTCATACC    GAATTGTTTA    TTGTAATTAT    ACAGACTATC    ATAATTATCA    AAAATGTCAT

Y  S  F     V  S  G     E  E  S  R     K  E  Q     S  Y  D     K  L  R  F
1750          1760          1770          1780          1790          1800
AATACTCATT    TGTATCTGGA    GAAGAAAGTA    GAAAAGAACA    ATCATACGAT    AAATTAAGAT
TTATGAGTAA    ACATAGACCT    CTTCTTTCAT    CTTTTCTTGT    TAGTATGCTA    TTTAATTCTA

D  A  I     N  I  G     L  S  T  K     L  N  N     F  K  S     G  W  T  K
1810          1820          1830          1840          1850          1860
TTGATGCGAT    AAACATTGGA    TTATCTACTA    AATTAAATAA    CTTTAAATCA    GGTTGGACTA
AACTACGCTA    TTTGTAACCT    AATAGATGAT    TTAATTTATT    GAAATTTAGT    CCAACCTGAT

S  I  Y     F  S  S     D  A  I  I     K  N  E     N  T  S     F  F  K  N
1870          1880          1890          1900          1910          1920
AATCAATATA    CTTTTCATCA    GATGCCATAA    TTAAAAATGA    AAATACAAGT    TTTTTTAAAA
TTAGTTATAT    GAAAAGTAGT    CTACGGTATT    AATTTTTACT    TTTATGTTCA    AAAAAATTTT

F  Y  L     D  Y  K     I  D  K  T     I  D  R     I  V  L     S  L  K  T
1930          1940          1950          1960          1970          1980
ATTTCTATTT    AGATTATAAA    ATAGATAAAA    CTATAGATCG    TATAGTTTTA    TCCTTAAAAA
TAAAGATAAA    TCTAATATTT    TATCTATTTT    GATATCTAGC    ATATCAAAAT    AGGAATTTTT

G  L  T     Y  D  S     K  I  K  K     Y  N  P     Y  F  K     P  S  N  V
1990          2000          2010          2020          2030          2040
CAGGATTGAC    ATATGACTCA    AAGATAAAAA    AATATAACCC    ATATTTCAAA    CCATCAAATG
GTCCTAACTG    TATACTGAGT    TTCTATTTTT    TTATATTGGG    TATAAAGTTT    GGTAGTTTAC
```

```
I  T  L     K  P  R     V  D  F  L     V  N  P     Q  I  S     I  S  L  S
2050          2060          2070          2080          2090          2100
TAATAACTCT    AAAACCTAGA    GTAGATTTTT    TAGTTAATCC    ACAGATATCT    ATAAGTTTAT
ATTATTGAGA    TTTTGGATCT    CATCTAAAAA    ATCAATTAGG    TGTCTATAGA    TATTCAAATA

T  E  K     Q  L  K     S  S  E  K     Y  N  E     K  V  T     N  V  S  G
2110          2120          2130          2140          2150          2160
CTACAGAAAA    ACAATTAAAA    AGTAGTGAAA    AGTACAATGA    AAAAGTAACA    AATGTATCTG
GATGTCTTTT    TGTTAATTTT    TCATCACTTT    TCATGTTACT    TTTTCATTGT    TTACATAGAC

M  E  N     F  L  S     I  G  I  S     Y  H  L     N  L  I     N  R  L  Y
2170          2180          2190          2200          2210          2220
GGATGGAAAA    CTTTTTATCT    ATAGGAATAT    CGTATCATTT    AAATTTAATA    AATCGTTTAT
CCTACCTTTT    GAAAAATAGA    TATCCTTATA    GCATAGTAAA    TTTAAATTAT    TTAGCAAATA

F  E  T     T  F  N     K  T  G  N     S  G  A     T  I  N     L  N  Y  E
2230          2240          2250          2260          2270          2280
ATTTTGAAAC    TACTTTTAAT    AAAACAGGTA    ATTCAGGTGC    TACAATCAAC    CTTAACTATG
TAAAACTTTG    ATGAAAATTA    TTTTGTCCAT    TAAGTCCACG    ATGTTAGTTG    GAATTGATAC

K  D  V     *
2290          2300          2310          2320          2330          2340
AGAAAGATGT    ATAATTTAAT    TATCTTATTA    TGTATATCAG    ATTATATATC    TGCAAGAGAA
TCTTTCTACA    TATTAAATTA    ATAGAATAAT    ACATATAGTC    TAATATATAG    ACGTTCTCTT

2350          2360          2370          2380          2390          2400
ATTGTTTCAT    ATAATGAATT    TAAGAATTTC    CATATTATTA    GGCAAACAAA    AAATAATTCT
TAACAAAGTA    TATTACTTAA    ATTCTTAAAG    GTATAATAAT    CCGTTTGTTT    TTTATTAAGA

2410          2420          2430          2440          2450          2460
TGTGGAGCAG    CAGCTCTTGC    CACAATGCTA    AAATATAAAT    TTCATATTAG    TGAAATTAAC
ACACCTCGTC    GTCGAGAACG    GTGTTACGAT    TTTATATTTA    AAGTATAATC    ACTTTAATTG

2470          2480          2490          2500          2510          2520
GAAGATACTA    TTCTATATAA    ATTGAAAAAT    CCAAATGAAG    AAGCATCCTT    TTTTGAACTT
CTTCTATGAT    AAGATATATT    TAACTTTTTA    GGTTTACTTC    TTCGTAGGAA    AAAACTTGAA

2530          2540          2550          2560          2570          2580
GCAAGAATTT    CAAAAGAATT    GAATATAAAT    GCTATTGGAT    TAGCATTAAC    GTTAAAGGAA
CGTTCTTAAA    GTTTTCTTAA    CTTATATTTA    CGATAACCTA    ATCGTAATTG    CAATTTCCTT

2590          2600          2610          2620          2630          2640
TTATTAAACA    TAAATAAACC    AGTAATTGCC    TATGTAAACA    ACAGTTTAAA    CAATGATCAT
AATAATTTGT    ATTTATTTGG    TCATTAACGG    ATACATTTGT    TGTCAAATTT    GTTACTAGTA

2650          2660          2670          2680          2690          2700
TTTATTATCA    TCAATGGTAT    TTTTAATAAA    GAGTTATTAA    TATCAGATGC    AGCAATTGGA
AAATAATAGT    AGTTACCATA    AAAATTATTT    CTCAATAATT    ATAGTCTACG    TCGTTAACCT

2710          2720          2730          2740          2750          2760
AATTACTCAC    TAAAAGTTTC    TGACTTTGAA    AAAATATGGC    TATTGAGAAA    TGATAAAAAA
TTAATGAGTG    ATTTTCAAAG    ACTGAAACTT    TTTTATACCG    ATAACTCTTT    ACTATTTTTT

2770          2780          2790          2800          2810          2820
GGAGATATTT    TATACTTGCA    TAGAGATTCA    AAAGATCACT    TAGAGTTTAT    TGACCATATT
CCTCTATAAA    ATATGAACGT    ATCTCTAAGT    TTTCTAGTGA    ATCTCAAATA    ACTGGTATAA
```

```
                                                  M  Y     T  Y  E     K  Y  R  N
2830          2840          2850          2860          2870          2880
AAAACAAAAC    ATAGATTATT    ATTAAGAAAT    TAAATATGTA    CACATATGAA    AAATATAGAA
TTTTGTTTTG    TATCTAATAA    TAATTCTTTA    ATTTATACAT    GTGTATACTT    TTTATATCTT

N  K  N     H  F  F     H  D  Y  L     F  P  V     K  L  N     K  P  I  S
2890          2900          2910          2920          2930          2940
ATAATAAGAA    CCACTTCTTT    CATGATTATT    TATTCCCAGT    AAAACTTAAT    AAACCAATAT
TATTATTCTT    GGTGAAGAAA    GTACTAATAA    ATAAGGGTCA    TTTTGAATTA    TTTGGTTATA

K  F  I     G  E  N     G  V  G  K     S  P  I     M  E  A     I  A  I  Y
2950          2960          2970          2980          2990          3000
CTAAATTTAT    TGGTGAAAAT    GGCGTAGGTA    AATCACCCAT    AATGGAGGCT    ATAGCTATTT
GATTTAAATA    ACCACTTTTA    CCGCATCCAT    TTAGTGGGTA    TTACCTCCGA    TATCGATAAA

L  G  C     P  A  D     G  G  S  K     N  F  N     F  S  T     E  N  T  H
3010          3020          3030          3040          3050          3060
ATTTAGGTTG    CCCAGCAGAC    GGCGGTTCAA    AAAACTTTAA    TTTTTCAACT    GAAAATACAC
TAAATCCAAC    GGGTCGTCTG    CCGCCAAGTT    TTTTGAAATT    AAAAAGTTGA    CTTTTATGTG

I  Q  I     P  N  V     M  I  K  K     P  T  K     F  P  K     D  P  F  F
3070          3080          3090          3100          3110          3120
ATATCCAAAT    ACCAAATGTG    ATGATAAAAA    AACCAACTAA    ATTTCCTAAA    GATCCATTTT
TATAGGTTTA    TGGTTTACAC    TACTATTTTT    TTGGTTGATT    TAAAGGATTT    CTAGGTAAAA

Y  R  S     K  T  F     Y  T  F  L     S  E  M     K  R  L     D  A  P  E
3130          3140          3150          3160          3170          3180
TTTATAGATC    AAAAACTTTT    TATACCTTTC    TGAGTGAAAT    GAAAAGACTA    GATGCGCCAG
AAATATCTAG    TTTTTGAAAA    ATATGGAAAG    ACTCACTTTA    CTTTTCTGAT    CTACGCGGTC

S  G  G     G  K  K     N  S  Y  Y     G  G  I     E  L  H     K  L  S  H
3190          3200          3210          3220          3230          3240
AATCCGGAGG    GGGTAAAAAA    AATTCTTACT    ATGGAGGGAT    TGAATTACAT    AAATTATCTC
TTAGGCCTCC    CCCATTTTTT    TTAAGAATGA    TACCTCCCTA    ACTTAATGTA    TTTAATAGAG

G  K  S     M  N  A     L  Y  K  N     R  F  N     K  N  G     L  Y  I  W
3250          3260          3270          3280          3290          3300
ATGGAAAATC    AATGAATGCA    CTATATAAGA    ATAGATTTAA    CAAAAATGGG    CTATATATAT
TACCTTTTAG    TTACTTACGT    GATATATTCT    TATCTAAATT    GTTTTTACCC    GATATATATA

D  E  P     E  S  S     L  S  L  S     N  Q  L     K  W  I     E  R  I  V
3310          3320          3330          3340          3350          3360
GGGATGAACC    CGAATCCTCC    TTATCTCTCA    GCAATCAGCT    TAAATGGATT    GAAAGAATCG
CCCTACTTGG    GCTTAGGAGG    AATAGAGAGT    CGTTAGTCGA    ATTTACCTAA    CTTTCTTAGC

N  L  S     R  M  G     A  Q  F  I     I  A  T     H  S  P     I  I  R  Q
3370          3380          3390          3400          3410          3420
TAAATCTAAG    TAGAATGGGA    GCCCAATTTA    TCATTGCAAC    GCATTCACCT    ATTATTAGGC
ATTTAGATTC    ATCTTACCCT    CGGGTTAAAT    AGTAACGTTG    CGTAAGTGGA    TAATAATCCG

T  P  E     S  E  L     L  E  V  T     K  K  G     V  K  T     V  N  F  Q
3430          3440          3450          3460          3470          3480
AAACCCCTGA    ATCAGAGTTA    CTAGAAGTGA    CAAAAAAAGG    TGTAAAAACA    GTTAACTTTC
TTTGGGGACT    TAGTCTCAAT    GATCTTCACT    GTTTTTTTCC    ACATTTTTGT    CAATTGAAAG
```

```
E  T  N     I  Y  Y     M  Y  R  E     F  M  Q     D  N  S     H  T  C  L
3490          3500          3510          3520          3530          3540
AAGAAACAAA    CATATATTAT    ATGTATCGTG    AGTTTATGCA    GGATAATTCT    CACACTTGTC
TTCTTTGTTT    GTATATAATA    TACATAGCAC    TCAAATACGT    CCTATTAAGA    GTGTGAACAG

S  S  R     L  N  V     *
3550          3560          3570          3580          3590          3600
TGAGTAGTAG    ACTAAATGTA    TAAATAATAA    ATTAAGAGGA    GTAACAGTAA    TGAAGAAATT
ACTCATCATC    TGATTTACAT    ATTTATTATT    TAATTCTCCT    CATTGTCATT    ACTTCTTTAA

         M     K  T  Y     F  I  L     L  K  A  I     I  F  F     S  A  S
3610          3620          3630          3640          3650          3660
TTTTATTTTA    TGAAAACATA    TTTTATTTTA    TTAAAGGCCA    TTATTTTTTT    TTCAGCTTCG
AAAATAAAAT    ACTTTTGTAT    AAAATAAAAT    AATTTCCGGT    AATAAAAAAA    AAGTCGAAGC

L  A  S  S     F  D  V     Q  S  S     F  K  S  S     Q  S  E     R  V  T
3670          3680          3690          3700          3710          3720
CTTGCAAGTA    GTTTTGACGT    GCAATCATCC    TTCAAATCCT    CTCAAAGTGA    AAGGGTTACA
GAACGTTCAT    CAAAACTGCA    CGTTAGTAGG    AAGTTTAGGA    GAGTTTCACT    TTCCCAATGT

V  L  G  Q     K  E  L     F  E  I     K  D  G  A     T  S  H     V  Y  T
3730          3740          3750          3760          3770          3780
GTATTAGGGC    AGAAAGAGCT    ATTTGAGATT    AAAGATGGTG    CTACATCTCA    CGTGTATACC
CATAATCCCG    TCTTTCTCGA    TAAACTCTAA    TTTCTACCAC    GATGTAGAGT    GCACATATGG

S  S  I  C     G  E  K     H  G  G     I  Y  S  P     S  G  C     Y  N  C
3790          3800          3810          3820          3830          3840
TCTTCAATAT    GCGGAGAAAA    ACATGGTGGA    ATATATTCTC    CTAGTGGGTG    CTATAACTGC
AGAAGTTATA    CGCCTCTTTT    TGTACCACCT    TATATAAGAG    GATCACCCAC    GATATTGACG

F  N  K  G     L  R  V     N  G  I     R  R  T  *
3850          3860          3870          3880          3890          3900
TTTAATAAAG    GACTTAGAGT    AAACGGTATT    AGACGTACAT    AAAATTATGG    GAATAATAGC
AAATTATTTC    CTGAATCTCA    TTTGCCATAA    TCTGCATGTA    TTTTAATACC    CTTATTATCG

                                  M     L  K  G     Y  K  N     F  A  L  F
3910          3920          3930          3940          3950          3960
GCCCCATGAA    GGGAAAGAGC    TAGATTTAAA    TGTTAAAAGG    TTATAAAAAT    TTTGCTCTTT
CGGGGTACTT    CCCTTTCTCG    ATCTAAATTT    ACAATTTTCC    AATATTTTTA    AAACGAGAAA

Y  T  D     Y  N  I     P  Y  G  C     I  P  Y     L  K  W     G  F  F  K
3970          3980          3990          4000          4010          4020
TTTACACTGA    TTATAATATC    CCATATGGAT    GTATCCCATA    TTTGAAATGG    GGTTTCTTTA
AAATGTGACT    AATATTATAG    GGTATACCTA    CATAGGGTAT    AAACTTTACC    CCAAAGAAAT

I  K  K     V  R  T     M  D  S  N     G  N  T     Y  Y  Y     Y  I  I  Y
4030          4040          4050          4060          4070          4080
AAATAAAAAA    AGTTAGAACA    ATGGATTCTA    ACGGTAATAC    ATATTATTAC    TATATAATAT
TTTATTTTTT    TCAATCTTGT    TACCTAAGAT    TGCCATTATG    TATAATAATG    ATATATTATA

K  K  K     H  T  R     K  A  K  T     L  S  I     L  L  T     K  S  T  N
4090          4100          4110          4120          4130          4140
ATAAAAAAAA    ACATACAAGA    AAAGCAAAAA    CACTATCGAT    TTTACTCACA    AAAAGTACAA
TATTTTTTTT    TGTATGTTCT    TTTCGTTTTT    GTGATAGCTA    AAATGAGTGT    TTTTCATGTT
```

```
C  F  N     L  N  Y     E  R  I  I     G  K  L     L  G  Y     S  K  E  D
4150         4160         4170         4180         4190         4200
ATTGTTTTAA   TTTAAATTAT   GAAAGAATTA   TAGGAAAATT   ATTAGGGTAT   AGTAAAGAAG
TAACAAAATT   AAATTTAATA   CTTTCTTAAT   ATCCTTTTAA   TAATCCCATA   TCATTTCTTC

S  E  F     Y  I  K     N  C  I  S     N  Y  M     N  *
4210         4220         4230         4240         4250         4260
ATAGTGAATT   TTATATAAAA   AATTGTATAT   CTAACTATAT   GAATTAATCT   TCAAGGGCAG
TATCACTTAA   AATATATTTT   TTAACATATA   GATTGATATA   CTTAATTAGA   AGTTCCCGTC

4270         4280         4290         4300         4310         4320
GTTTCCTAAT   AGAGGATGCC   CTTTGATCAC   GGCTTGGTGT   TGTGCATTTT   GATTGAATAG
CAAAGGATTA   TCTCCTACGG   GAAACTAGTG   CCGAACCACA   ACACGTAAAA   CTAACTTATC

4330         4340         4350         4360         4370         4380
TGGAGACATC   GAATCCATAA   AACATTCATA   CAGTGGTGTG   ATAGGAAGCG   TTAATAAAGG
ACCTCTGTAG   CTTAGGTATT   TTGTAAGTAT   GTCACCACAC   TATCCTTCGC   AATTATTTCC

4390         4400         4410         4420         4430         4440
GCATTTGGCA   CTTTCAGCAA   CTAATTTGGC   ATCTTCTACA   TCTTGCGTGC   TGGTTGTGGC
CGTAAACCGT   GAAAGTCGTT   GATTAAACCG   TAGAAGATGT   AGAACGCACG   ACCAACACCG

4450         4460         4470         4480         4490         4500
AGAAGGTAGT   AATAGCGCTT   GCACGGGCGC   GCCTGTACGC   ATTAATAAAT   GGCTGACCAC
TCTTCCATCA   TTATCGCGAA   CGTGCCCGCG   CGGACATGCG   TAATTATTTA   CCGACTGGTG

4510         4520         4530         4540         4550         4560
CGCAGAGTCA   ATTCCGCCGC   TAATGCCGAC   CACATAGCCA   AGGGTATTAT   AATCTTGTCG
GCGTCTCAGT   TAAGGCGGCG   ATTACGGCTG   GTGTATCGGT   TCCCATAATA   TTAGAACAGC

L  L  L  N     P  L  N     *  I  G     *  I  I  F     H  W  T     S  P  Y
4570         4580         4590         4600         4610         4620
TTGTTGCTCA   ACCCATTGAA   TTAAATAGGC   TAAATAATCT   TTCATTGGAC   TTCTCCTTAC
AACAACGAGT   TGGGTAACTT   AATTTATCCG   ATTTATTAGA   AAGTAACCTG   AAGAGGAATG

4630         4640         4650         4660         4670         4680
TATCTATACA   TATTCATAAC   TGTTCCCAAT   GTTTGttaCA   TGAGAAAAAC   GTGGTAAAAA
ATAGATATGT   ATAAGTATTG   ACAAGGGTTA   CAAACaatGT   ACTCTTTTTG   CACCATTTTT
                                      *     M  L  F  V     H  Y  F

4690         4700         4710         4720         4730         4740
TGTACCGTAC   TTAAGCGTGG   ATTAACACGG   TAGCGTCTTC   CCGCTTTTGC   TTGTGCTTTC
ACATGGCATG   AATTCGCACC   TAATTGTGCC   ATCGCAGAAG   GGCGAAAACG   AACACGAAAG
H  V  T     S  L  R  P     N  V  R     Y  R  R     G  A  K  A     Q  A  K

4750         4760         4770         4780         4790         4800
CAGTAAAAAC   AATGGAGATT   TTGAGCAAAG   AGGCACCGAT   GCCATAGTAA   TCCACAGGGT
GTCATTTTTG   TTACCTCTAA   AACTCGTTTC   TCCGTGGCTA   CGGTATCATT   AGGTGTCCCA
W  Y  F     C  H  L  N     Q  A  F     L  C  R     H  W  L  L     G  C  P

4810         4820         4830         4840         4850         4860
ACACCTGTTA   TGTCGAAAGG   CACAGATTGA   GCCTCAAATT   CCCTAATTCT   TTTTTCATCA
TGTGGACAAT   ACAGCTTTCC   GTGTCTAACT   CGGAGTTTAA   GGGATTAAGA   AAAAAGTAGT
V  G  T     I  D  F  P     V  S  Q     A  E  F     E  R  I  R     K  E  D

4870         4880         4890         4900         4910         4920
AAACCCATGG   AAAGCCTGCT   TGATCTAGTT   GAGAATCCCT   CTTCATCTAA   TGCTTTTCTT
TTTGGGTACC   TTTCGGACGA   ACTAGATCAA   CTCTTAGGGA   GAAGTAGATT   ACGAAAAGAA
F  G  M     S  L  R  S     S  R  T     S  F  G     E  E  D  L     A  K  R
```

```
4930        4940        4950        4960        4970        4980
AACGCCACGC  AAATCCTCTT  GTCCAACCCC  CGAGGTTTCT  TGGTGGGTAA  GGAAGTAGCG
TTGCGGTGCG  TTTAGGAGAA  CAGGTTGGGG  GCTCCAAAGA  ACCACCCATT  CCTTCATCGC
L  A  V     C  I  R  K     D  L  G     R  P  K     K  T  P  L     S  T  A

4990        5000        5010        5020        5030        5040
ATGTACCATA  TTGGCGGAGG  TATCAACGCG  CACCGCATAG  AGTTTTTCAC  CGAAGTGTTT
TACATGGTAT  AACCGCCTCC  ATAGTTGCGC  GTGGCGTATC  TCAAAAAGTG  GCTTCACAAA
I  Y  W     I  P  P  P     I  L  A     C  R  M     S  N  K  V     S  T  N

5050        5060        5070        5080        5090        5100
TGGTACTTAT  AAACTATCGG  TAAACACATC  ATTGTTGTAA  TCAACCAAAG  CGGTAAGGGG
ACCATGAATA  TTTGATAGCC  ATTTGTGTAG  TAACAACATT  AGTTGGTTTC  GCCATTCCCC
Q  Y  K     Y  V  I  P     L  C  M     M  T  T     I  L  W  L     P  L  P

5110        5120        5130        5140        5150        5160
AGAATTAGGA  TAGACTTGCA  TATAGGCCTC  GCAAGCCTGA  ATAAGGTTGC  CATTTCTACA
TCTTAATCCT  ATCTGAACGT  ATATCCGGAG  CGTTCGGACT  TATTCCAACG  GTAAAGATGT
L  I  L     I  S  K  C     I  P  R     A  L  R     F  L  T  A     M  E  V
5170        5180        5190        5200        5210        5220
ACATCACCGT  TGAATAAGGG  AATGTGGCAT  TGTGCGTACG  CGCCCATTCC  CTGTTCCGCT
TGTAGTGGCA  ACTTATTCCC  TTACACCGTA  ACACGCATGC  GCGGGTAAGG  GACAAGGCGA
V  D  G     N  F  L  P     I  H  C     Q  A  Y     A  G  M  G     Q  E  A

5230        5240        5250        5260        5270        5280
TGGTATTGGT  GCATTGCATC  GGTGGCTTGC  ATTGTTGCCC  CGCCAATATA  GGCGGCGTAA
ACCATAACCA  CGTAACGTAG  CCACCGAACG  TAACAACGGG  GCGGTTATAT  CCGCCGCATT
Q  Y  Q     H  M  A  D     T  A  Q     M  T  A     G  G  I  Y     A  A  Y

5290        5300        5310        5320        5330        5340
CCATCAGTGG  GTTGCAAGTA  AAATAAACAT  TGCGGTCCAC  CCATAAATAA  TGGGTTTTCC
GGTAGTCACC  CAACGTTCAT  TTTATTTGTA  ACGCCAGGTG  GGTATTTATT  ACCCAAAAGG
G  D  T     P  Q  L  Y     F  L  C     Q  P  G     G  M  F  L     P  N  E

5350        5360        5370        5380        5390        5400
GTTGGCGGGT  CTGAACGTTC  CGCCATGGCG  ATGTTATATA  CATTTGTAAC  CACAGAAGTT
CAACCGCCCA  GACTTGCAAG  GCGGTACCGC  TACAATATAT  GTAAACATTG  GTGTCTTCAA
T  P  P     D  S  R  E     A  M  A     I  N  Y     V  N  T  V     V  S  T

5410        5420        5430        5440        5450        5460
CGACGCACAT  AAATGCCATC  AATAACACCA  TCCTCGGTTG  GCAAAACCCT  CATAAAATCT
GCTGCGTGTA  TTTACGGTAG  TTATTGTGGT  AGGAGCCAAC  CGTTTTGGGA  GTATTTTAGA
R  R  V     Y  I  G  D     I  V  G     D  E  T     P  L  V  R     M  F  D

5470        5480        5490        5500        5510        5520
TCCACAATTA  ATACGTTACA  AAGGGGGAGA  TTACGTCGCC  AACATTTTAT  TGCCCAATTT
AGGTGTTAAT  TATGCAATGT  TTCCCCCTCT  AATGCAGCGG  TTGTAAAATA  ACGGGTTAAA
E  V  I     L  V  N  C     L  P  L     N  R  R     W  C  K  I     A  W  N

5530        5540        5550        5560        5570        5580
TCAATCCTTC  CAGGGTTCCG  TGCAAAACGT  TGTAGTAAGG  CGATGGCTTC  ATCAACACCA
AGTTAGGAAG  GTCCCAAGGC  ACGTTTTGCA  ACATCATTCC  GCTACCGAAG  TAGTTGTGGT
E  I  R     G  P  N  R     A  F  R     Q  L  L     A  I  A  E     D  V  G

5590        5600        5610        5620        5630        5640
CATAAAATGG  CCTGTTTTTT  TTGGAAGAAT  TGCATTGTTA  CCCATTGTTT  TGACCTCTTT
GTATTTTACC  GGACAAAAAA  AACCTTCTTA  ACGTAACAAT  GGGTAACAAA  ACTGGAGAAA
C  L  I     A  Q  K  K     Q  F  F     Q  M  T     V  W  Q  K     S  R  K
```

```
5650        5660        5670        5680        5690        5700
TGTTCCGCAA  TGTGCTTTGC  TTTTAGAAAA  TAGCTTGCGG  AAAAGTACCC  TTCTCCTATC
ACAAGGCGTT  ACACGAAACG  AAAATCTTTT  ATCGAACGCC  TTTTCATGGG  AAGAGGATAG
Q  E  A     I  H  K  A     K  L  F     Y  S  A     S  F  Y  G     E  G  I

5710        5720        5730        5740        5750        5760
TTTTCATCAA  AATTAAAGGT  TTGAGATGGT  CGTCTTGTGA  TATTTACTTT  CTCTTTTTTT
AAAAGTAGTT  TTAATTTCCA  AACTCTACCA  GCAGAACACT  ATAAATGAAA  GAGAAAAAAA
K  E  D     F  N  F  T     Q  S  P     R  R  T     I  N  V  K     E  K  K

5770        5780        5790        5800        5810        5820
TCAATTTCCT  TCATCATAGT  AAAATGGATA  GTAAGTTATT  TATTATGTAT  ATTGTGTGCT
AGTTAAAGGA  AGTAGTATCA  TTTTACCTAT  CATTCAATAA  ATAATACATA  TAACACACGA
E  I  E     K  M

5830        5840        5850        5860        5870        5880
ACATTTTTGC  GATTTACAAC  AAACATATTT  CATAGAAAAG  AGCATTTTGA  TAAGCAGAGA
TGTAAAAACG  CTAAATGTTG  TTTGTATAAA  GTATCTTTTC  TCGTAAAACT  ATTCGTCTCT

5890        5900        5910        5920        5930        5940
TAACACTTGG  GGGCTGGTTG  TGGTCTTGAA  TGATAAAAAA  TGGTTGGATA  GTGACTTTCA
ATTGTGAACC  CCCGACCAAC  ACCAGAACTT  ACTATTTTTT  ACCAACCTAT  CACTGAAAGT

5950        5960        5970        5980        5990        6000
TAGATTTTAT  TTTTTGAGTG  GTTGTATTCT  TTTTCATCTT  TCTCTTCTGT  GTCGGTTAAA
ATCTAAAATA  AAAAACTCAC  CAACATAAGA  AAAAGTAGAA  AGAGAAGACA  CAGCCAATTT

6010        6020        6030        6040        6050        6060
TCGCGATAAG  AGTTTGTTTG  CAAAACAAGC  TATTTTGTCA  TTTCTGAAAT  TAAAATGGTT
AGCGCTATTC  TCAAACAAAC  GTTTTGTTCG  ATAAAACAGT  AAAGACTTTA  ATTTTACCAA

6070        6080        6090   6095
GTCAAGAGGT  TTTTATGAGG  AAATACAGCT  AAGAa
CAGTTCTCCA  AAAATACTCC  TTTATGTCGA  TTCTt
```

Figure 14. The complete sequence of the plasmid encoding NAD^+ independence in *H. paragallinarum.* The six proteins encoded on the plasmid are indicated as follows: repA – dark green; hmcD – turquoise; abc – yellow; hsp – bright green; him – red; qr – purpule. Amino acid sequences are represented either on the upper part of the DNA sequences indicating that the coding direction of the specific open reading frame is present on the Crick strand or on the Watson strand if the amino acid sequence is indicated below the DNA sequence.

4.10 Homologues of the genes encoded on the *H. paragallinarum* plasmid.

4.10.1 Haemocin structural, Haemocim resistance, Haemocin transport and Haemocin immunity gene.

Haemocin is a bacteriocin produced by most type b-encapsulated strains of *Haemophilus influenzae* (Musser *et al*., 1986). This bacteriocin is toxic to virtually all non type b strains of *H. influenzae*, both encapsulated and non-encapsulated. Although little is known about the function of haemocin, evidence to date suggests that DNA synthesis, of susceptible strains, is inhibited by this bacteriocin (Lipuma *et al*., 1990).

Bacterial cells capable of haemocin production must be immune to the toxic effects of the bacteriocin protein. This immunity is mediated via the co-production of a specific immunity protein (Axelsson *et al*., 1995). The immunity gene of *H. influenzae* was cloned and characterised (Murley *et al*., 1997). Immunity was found to be encoded on a 1.5kb chromosomal fragment. The deduced amino acid sequence consists of 105 amino acids with an estimated molecular mass of 12.6 kDa. The immunity gene bears little homology to other proteins in published databases.

Similarly a 3284bp nucleotide fragment of the plasmid from *H. paragallinarum* revealed four open reading frames. Proteins encoded on this fragment of the plasmid all have functions related to production and immunity of the bacteriocin haemocim. The four proteins are: HmcD, haemocin structural, haemocin transporter and haemocin immunity proteins. All four of these show high similarity on both nucleotide and protein level to the proteins found within *H. influenzae*. This is clearly illustrated by the nucleotide and protein alignments shown in Fig's. 15 – 22.

```
H. influenzae HmcD  1   ATGTTTAATAATAATCTATTAAACTATATAAGGATATATATGAAAAACAAAATAT 55
NAD+ plasmid HmcD   1   ATGTATAATATAAATCTATTAAACTATATCAGGCTATATATGAAAACCAAAATAT 55

H. influenzae HmcD  56  TATTTATATTTATATTCAACCCAATACTGTGTTATTCTAATAATTATAATTTTAT 110
NAD+ plasmid HmcD   56  TATTTATATTTATATTCAACCCCATACTGTGTTATTCTAATAATGATAATTCTAT 110

H. influenzae HmcD  111 TACATTAAATGATGAATTTGACAACTCTGGTAAATTCACATTAGTAACATCATCT 165
NAD+ plasmid HmcD   111 TACATTAAATGATGAATTTGACAACTTTGGTAATTTCACATTAGTAACATCATCT 165

H. influenzae HmcD  166 TCATTACATTTAAAAAACAGCCATGAATTTTTAAGAAGAAATGTATTTGATATAC 220
NAD+ plasmid HmcD   166 TCAATACATTTTAAAAACAGCCAAGAATTTTTAAGAAGAAATGTAATTGATAAAC 220

H. influenzae HmcD  221 AAAATTATGACAATTCATATGGAAGTGAAAAATATCAAGTACTATCCTTTGAACA 275
NAD+ plasmid HmcD   221 AAAATTATGACAATTCATATGGAAGTGAAAAATATCAAGGACTATCCTTTGAGCA 275

H. influenzae HmcD  276 AGGAATTCAGTATGGTTTAACAAATAAAATTAATATATCTGGTAGTATTAATAGT 330
NAD+ plasmid HmcD   276 AGGGATTCAGTATGGCTTAACAAATAACATTAATATGTCTGATAGTATTAATAGT 330

H. influenzae HmcD  331 TCATACAGTAAATACTCATTTGTATCTAGAGAAGAAAGTAGAAAGGAACAATCAT 385
NAD+ plasmid HmcD   331 TTTTACAGTAAATACTCATTTGTATCTGGAGAAGAAAGTAGAAAAGAACAATCAT 385

H. influenzae HmcD  386 ACGATAAATTAAGATTTGATACGATAAATATTGGATTATCTACTAAATTAAATAA 440
NAD+ plasmid HmcD   386 ACGATAAATTAAGATTTGATGCGATAAACATTGGATTATCTACTAAATTAAATAA 440

H. influenzae HmcD  441 CTTTAAATCAGATTGGAATAAATCAATATACTTTTCATCAGATGCCATAATTAAA 495
NAD+ plasmid HmcD   441 CTTTAAATCAGGTTGGACTAAATCAATATACTTTTCATCAGATGCCATAATTAAA 495

H. influenzae HmcD  496 AATGAAAATACAAGTTTTTTTAAAAATTTCTATTTAGATTATATAATAGATAAAA 550
NAD+ plasmid HmcD   496 AATGAAAATACAAGTTTTTTTAAAAATTTCTATTTAGATTATAAAATAGATAAAA 550

H. influenzae HmcD  551 CTATAGATCCTATAGTTTTATCCTTAAAAACAGGATTGAAATATTACTCAAAGAT 605
NAD+ plasmid HmcD   551 CTATAGATCGTATAGTTTTATCCTTAAAAACAGGATTGACATATGACTCAAAGAT 605

H. influenzae HmcD  606 AAAAAAATATAACCAATATTTCAAACCATCAAATGTAATAAATATAAAACCTAGA 660
NAD+ plasmid HmcD   606 AAAAAAATATAACCAATATTTCAAACCATCAAATGTAATAACTCTAAAACCTAGA 660

H. influenzae HmcD  661 GTAGATTTTTTAGTTAATCCACAGATATCTATAAGTTTATCTACAGAAATACAAT 715
NAD+ plasmid HmcD   661 GTAGATTTTTTAGTTAATCCACAGATATCTATAAGTTTATCTAGAGAAAAACAAT 715

H. influenzae HmcD  716 TAAAAAGTAGTGAAAAGTACAATGAAAAAGTAACAAATGTATCTGGAATAGAAAA 770
NAD+ plasmid HmcD   716 TAAAAAGTAGTGAAAAGTACAATGAAAAAGTAACAAATGTATCTGGGATGGAAAA 770

H. influenzae HmcD  771 CTTTTTATCTATAGGAATATCGTATAATTTAAATTTAAAAAATCGTTTATATTTT 825
NAD+ plasmid HmcD   771 CTTTTTATCTATAGGAATATCGTATCATTTAAATTTAATAAATCGTTTATATTTT 825

H. influenzae HmcD  826 GAAACTTCTTTTAATACAACAGGTAATTCAGGTGCTACAATCTACCTTAACTTTG 880
NAD+ plasmid HmcD   826 GAAACTACTTTTAATAAAACAGGTAATTCAGGTGCTACAATCAACCTTAACTATG 880

H. influenzae HmcD  881 AGAAAGATTTATAA 894
NAD+ plasmid HmcD   881 AGAAAGATCTATAA 894
```

Figure 15. Nucleotide alignment of the hmcD protein found in *H. influenzae* with that of the hmcD encoded on the *H. paragallinarum* plasmid showing identity of 94.63%.

```
NAD+ plasmid HmcD   1   MYNINLLNYIRLYMKTKILFIFIFNPILCYSNNDNSITLNDEFDNFGNFTLVTSS 55
H. influenzae HmcD  1   MFNNNLLNYIRIYMKNKILFIFIFNPILCYSNNYNFITLNDEFDNSGKFTLVTSS 55

NAD+ plasmid HmcD   56  SIHFKNSQEFLRRNVIDKQNYDNSYGSEKYQGLSFEQGIQYGLTNNINMSDSINS 110
H. influenzae HmcD  56  SLHLKNSHEFLRRNVFDIQNYDNSYGSEKYQVLSFEQGIQYGLTNKINISGSINS 110

NAD+ plasmid HmcD   111 FYSKYSFVSGEESRKEQSYDKLRFDAINIGLSTKLNNFKSGWTKSIYFSSDAIIK 165
H. influenzae HmcD  111 SYSKYSFVSREESRKEQSYDKLRFDTINIGLSTKLNNFKSDWNKSIYFSSDAIIK 165

NAD+ plasmid HmcD   166 NENTSFFKNFYLDYKIDKTIDRIVLSLKTGLTYDSKIKKYNQYFKPSNVITLKPR 220
H. influenzae HmcD  166 NENTSFFKNFYLDYIIDKTIDPIVLSLKTGLKYYSKIKKYNQYFKPSNVINIKPR 220

NAD+ plasmid HmcD   221 VDFLVNPQISISLSREKQLKSSEKYNEKVTNVSGMENFLSIGISYHLNLINRLYF 275
H. influenzae HmcD  221 VDFLVNPQISISLSTEIQLKSSEKYNEKVTNVSGIENFLSIGISYNLNLKNRLYF 275

NAD+ plasmid HmcD   276 ETTFNKTGNSGATINLNYEKDL 297
H. influenzae HmcD  276 ETSFNTTGNSGATIYLNFEKDL 297
```

Figure 16. Alignment of the protein from the plasmid of *H. paragallinarum* with the hmcD protein of *H. influenzae* showing high identity of 87.87% and similarity of 92%.

```
H. influenzae 1   ATGAAAAAATTTTTTATTTTATTAACAGCCA 31
NAD+ plasmid  1   ATGAAAAAATATTTTATTTTATTAAAAGCCA 31

H. influenzae 32  TTATTTTTTCATCAGCTTCGCTAGCAAGTAG 62
NAD+ plasmid  32  TTATTTTTTTTTCAGCTTCGCTTGCAAGTAG 62

H. influenzae 63  TAATGACGTGCAATCATTATTCAAATCATCT 93
NAD+ plasmid  63  TTTTGACGTGCAATCATCCTTCAAATCCTCT 93

H. influenzae 94  CAAAGTGAAAGTGTTACAGTATTATCTCAGA 124
NAD+ plasmid  94  CAAAGTGAAAGGGTTACAGTATTAGGGCAGA 124

H. influenzae 125 AAGAGCTATCTGAGATTAAAGGTGGTGCTAG 155
NAD+ plasmid  125 AAGAGCTATTTGAGATTAAAGATGGTGCTAG 155

H. influenzae 156 ATCTCCCGTGTATACCTGTTCAATATGCGGA 186
NAD+ plasmid  156 ATCTCACGTGTATACCTCTTCAATATGCGGA 186

H. influenzae 187 GCAAAACATGGTGGAGTATATTCTCCTAGTG 217
NAD+ plasmid  187 GAAAAACATGGTGGAATATATTCTCCTAGTG 217

H. influenzae 218 TATGCTATAACTGCTATAATAAAGGACTTAG 248
NAD+ plasmid  218 GGTGCTATAACTGCTTTAATAAAGGACTTAG 248

H. influenzae 249 AATAAACGGTATTAGACGTCCATAA 273
NAD+ plasmid  249 AGTAAACGGTATTAGACGTACATAA 273
```

Figure 17. Nucleotide alignment of the haemocim structural protein found in *H. influenzae* with that of the haemocim structural protein encoded on the *H. paragallinarum* plasmid showing identity of 90.84%.

```
NAD+ plasmid  1   MKKYFILLKAIIFFSASLASSFDVQSSFKSSQS 33
H. influenzae 1   MKKFFILLTAIIFSSASLASSNDVQSLFKSSQS 33

NAD+ plasmid  34 ERVTVLGQKELFEIKDGARSHVYTSSICGEKHG 66
H. influenzae 34 ESVTVLSQKELSEIKGGARSPVYTCSICGAKHG 66

NAD+ plasmid  67 GIYSPSGCYNCFNKGLRVNGIRRT 90
H. influenzae 67 GVYSPSVCYNCYNKGLRINGIRRP 90
```

Figure 18. Alignment of the protein from the plasmid of *H. paragallinarum* with the haemocin structural protein of *H. influenzae* showing high identity of 81.11% and similarity of 85.55%.

```
NAD+ plasmid  1   ATGTACACATATGAAAAATATAGAAATAATAAGAACCACT 40
H. influenzae 1   ATGTACATATATGAAATATATAGAAATAATAAGAACCACT 40

NAD+ plasmid  41  TCTTTCATGATTATTTATTCCCAGTAAAACTTAATAAACC 80
H. influenzae 41  TCTTTCTTGATTATTTATTCCCAGTAAAACTTAATAAACC 80

NAD+ plasmid  81  AATATCTAAATTTATTGGTGAAAATGGCGTAGGTAAATCA 120
H. influenzae 81  AATATCTATATTTATTGGTGAAAATGGCGTAGGTAAATCA 120

NAD plasmid   121 CCCATAATGGAGGCTATAGCTATTTATTTAGGTTGCCCAG 160
H. influenzae 121 ACCATAATGGAGGCTATAGCTATTTATTTAGGTTGCCCAG 160

NAD+ plasmid  161 CAGACGGCGGTTCAAAAAACTTTAATTTTTCAACTGAAAA 200
H. influenzae 161 CAGAAGGCGGTTCAAAAAACTTTAATTTTTCAACTGAAAA 200

NAD+ plasmid  201 TACACATATCCAAATACCAAATGTGATGATAAAAAAACCA 240
H. influenzae 201 TACACATATCCAAATACCAAATATGATGATAAAAAAAGGA 240

NAD+ plasmid  241 ACTAAATTTCCTAAAGATCC-TTTTTTTATAGATCAAAAA 279
H. influenzae 241 ACTAAATTTCCTAAAGATATATTTTTTTATAGATCAGAAA 280

NAD+ plasmid  280 CTTTTTATACCTTTCTGAGTGAAATGAAAAGACTAGATGC 319
H. influenzae 281 CTTTTTATACCTTTATGAGTGAAATGAAAAAACTAGATGC 320

NAD+ plasmid  320 GCCAGAATCCGGAGGGGGTAAAAAAAATTCTTACTATGGA 359
H. influenzae 321 GCCAGAATTTGGAGGGGGTAAAATAAATTCTTACTATGGA 360

NAD+ plasmid  360 GGGATTGAATTACATAAATTATCTCATGGAAAATCAATGA 399
H. influenzae 361 GGGATTGAATTACATAAATTATCTCATGGAGAATCAATGA 400

NAD+ plasmid  400 ATGCACTATATAAGAATAGATTTAACAAAAATGGGCTATA 439
H. influenzae 401 ATGCACTATATATGAATAGATTTAACAAAAATGGGCTATA 440

NAD+ plasmid  440 TATATGGGATGAACCCGAATCCTCCTTATCTCTCAGCAAT 479
H. influenzae 441 TATATTAGATGAACCCGAATCCTCCTTATCTCTCAACAAT 480

NAD+ plasmid  480 CAGCTTAAATGGATTGAAAGAATCGTAAATCTAAGTAGAA 519
H. influenzae 481 CAGCTTAAATTTATTGAAAGAATCGTAACTCTAAGTAGAA 520

NAD+ plasmid  520 TGGGAGCCCAATTTATCATTGCAACGCATTCACCTATTAT 559
H. influenzae 521 ATGGAGCCCAATTTATCATTGCAACGCATTCACCTATTAT 560

NAD+ plasmid  560 TAGGCAAACCCCTGAATCAGAGTTACTAGAAGTGACAAAA 599
H. influenzae 561 TATGCAAACCCCTGAATCAGAGTTACTAGAAGTGACAAAA 600

NAD+ plasmid  600 AAAGGTGTAAAAACAGTTAACTTTCAAGAAACAAACATAT 639
H. influenzae 601 AATGGTGTAAAAACAGTTAACTTTCAAGAAACAAACATAT 640
```

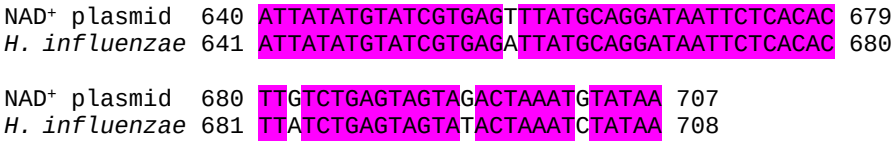

Figure 19. Nucleotide alignment of the haemocim transporter protein found in *H. influenzae* with that of the haemocim transporter protein encoded on the *H. paragallinarum* plasmid showing identity of 95.06%.

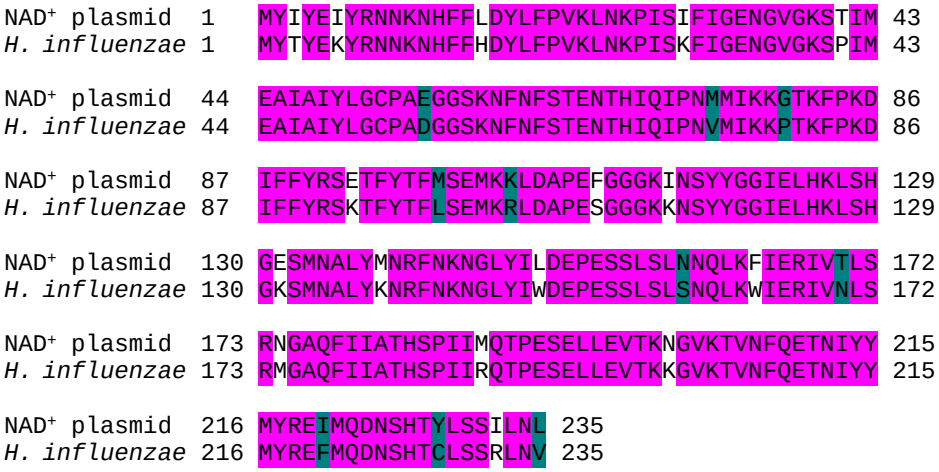

Figure 20. Alignment of the protein from the plasmid of *H. paragallinarum* with the haemocin transporter protein of *H. influenzae* showing high identity of 89.36% and similarity of 93.19%.

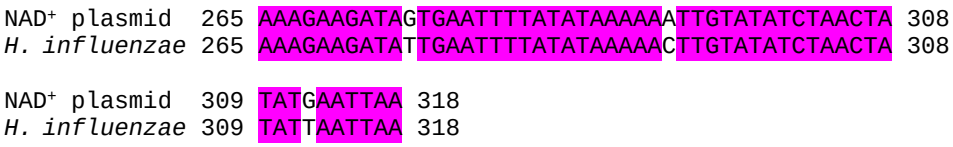

Figure 21. Nucleotide alignment of the haemocim immunity protein found in *H. influenzae* with that of the haemocim immunity protein encoded on the *H. paragallinarum* plasmid showing identity of 93.39%.

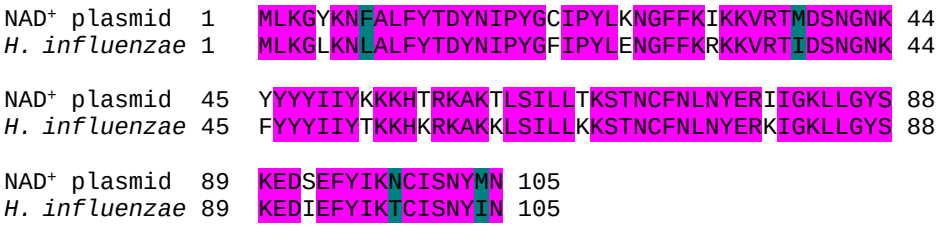

Figure 22. Alignment of the protein from the plasmid of *H. paragallinarum* with the haemocin immunity protein of *H. influenzae* showing high identity of 85.71% and similarity of 89.52%.

Determining the susceptibility of non plasmid bearing strains of *H. paragallinarum* to haemocin toxicity may form the basis of a future study. The presence of a haemocim encoding gene on the plasmid isolated from *H. paragallinarum* suggests a correlation to the plasmid found in *H. parainfluenzae.*

Subsequent secondary pathogenic infection by bacteria and mycoplasmas play an important role in the observable morbidity and mortality resulting from coryza in developing countries. This can be explained by considering the ecological implications of a greater variety of species present in one ecosystem. Such variety usually leads to a greater ratio of available niches being filled, and in pathogenic infection implies greater tissue damage and stress placed on the host (Blackall, 1999). The presence of multiple infective organisms implies competition and a consequent reduction in reproductive rate for all organisms involved. The possibility exits that haemocim may play a role in promoting virulence of plasmid bearing strains of *H. paragallinarum,* by suppressing competitive infection and giving the organism a reproductive

advantage. It may be speculated that the lower severity of symptoms exhibited by chickens infected with plasmid bearing strains of *H. paragallinarum* as observed by Taole *et al*. (2002) is the result of haemocim suppression of secondary infection. Quantification of relative dominance of plasmid bearing strains in infected chickens might offer supporting evidence for the above.

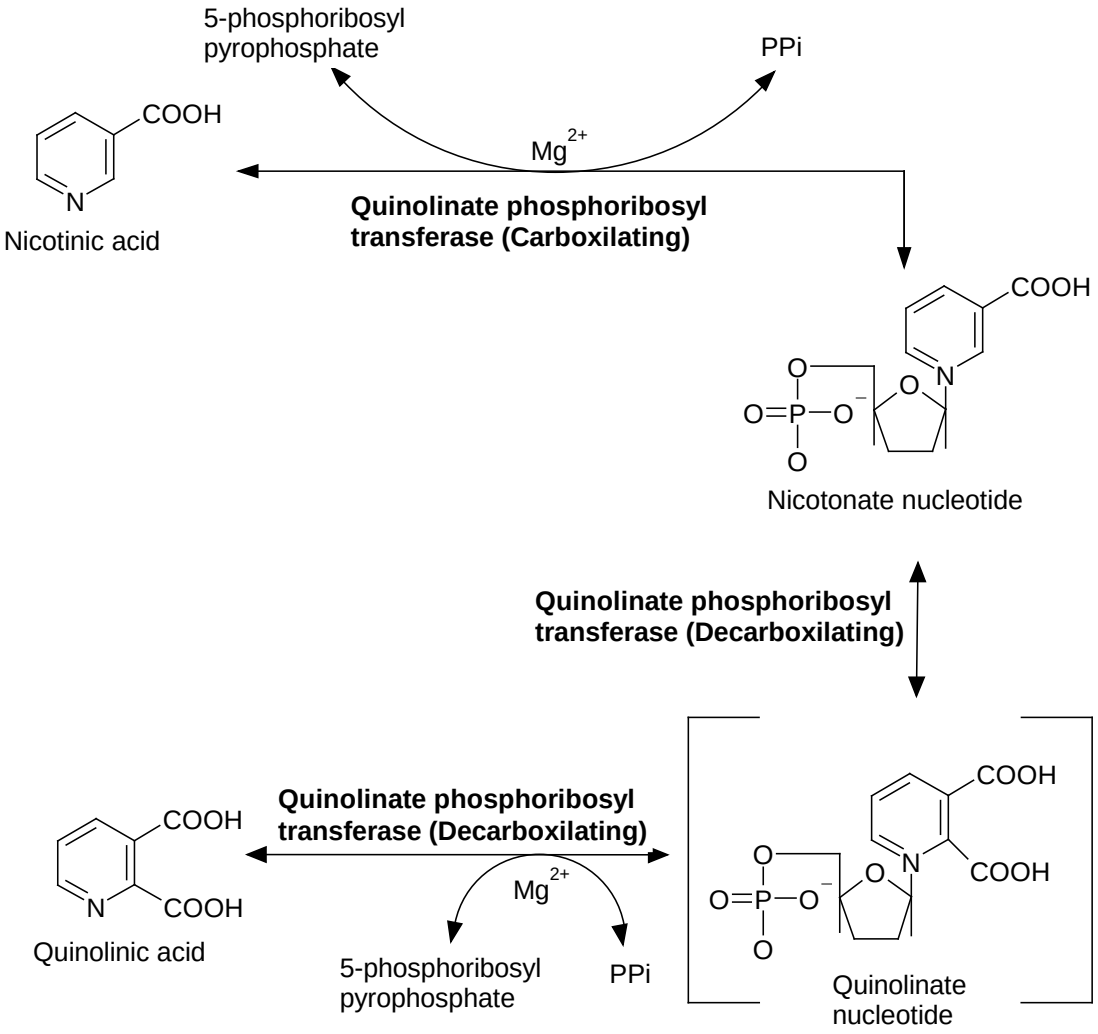

Figure 23. Reaction catalyzed by quinolinic acid phosphoribosyltransferase

4.10.2 Quinolinic acid phosphoribosyltransferase

Previous studies have shown that NAD^+-independent members of the family *Pasteurellaceae* differ from the NAD^+-dependent members solely in their ability to utilize the NAD^+ precursor NAm as V-factor (O`Reilly and Niven, 1986). Quinolinic acid phosphoribosyltransferase (QAPRTase) provides the route for quinolinic acid (QA) metabolism and is an essential step in the biosynthesis of NAD^+. This enzyme catalyzes the production of nicotinic acid mononucleotide (NAMN) from Nicotinamide (Eads *et al*., 1997) (Fig. 23).

NAD^+ independence in *H. ducreyi* was found to be plasmid mediated by Martin *et al*. (2001) who isolated a 5.25kb plasmid from *H. ducreyi.* The gene conferring the independence was subcloned and named NADV. The sequence of the NADV gene did not have any significant homology to proteins of the NAD^+–dependent members of the family *Pasteurellaceae*. Martin *et al*. (2001) used an enzyme assay for QAPRTase and proved that this enzyme is active in the NAD^+ independent strains of *H. ducreyi,* but absent in NAD^+ dependent strains of *H. ducreyi.* This experiment confirmed that the NADV gene encodes QAPRTase.

NAD^+ independence of *H. paragallinarum* is apparently also the result of production of QAPRTase. A deduced amino acid sequence of an open reading frame on the *H. paragallinarum* plasmid shows 65.32% homology with a QAPRTase protein from *Bacillus anthracis*. This protein enables the organism to produce NMN from nicotinamide and 5-phosphoribosyl pyrophosphate. (Fig. 24)

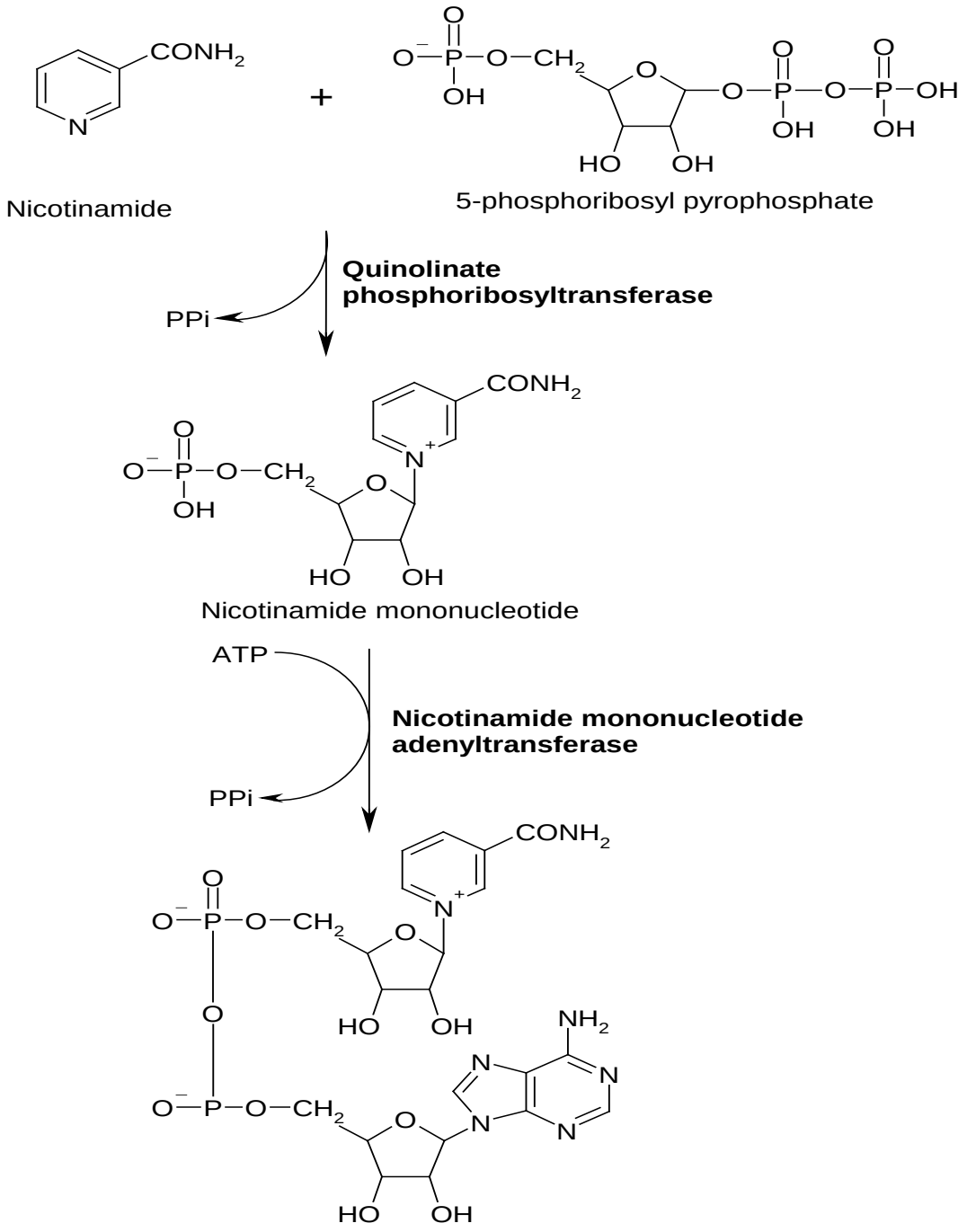

Figure 24. NAD^+ production in NAD^+ independent *H. paragallinarum.*

Although NAD^+ independence in both *H. paragallinarum* and *H. ducreyi* are plasmid mediated and conferred through the same enzyme, very little homology exist between the NADV protein identified by Martin *et al*. (2001) and the amino acid sequence obtained from the plasmid conferring

independence in *H. paragallinarum* as illustrated in Fig. 24. In addition, the NADV gene consist of 495 amino acids compared to the much smaller (123 amino acids less) QAPRtase of *H. paragallinarum*

```
H paragallinarum    1                                   MKE  3
H ducreyi           1   MDNLLNYSSRASAIPSLLCDFYKTSHRIMYPE  32

H paragallinarum    4   IEKKEKVNITRRPSQTFNFDEKIGEGYFSASY  35
H ducreyi           33  CSQIIYSTFTPRSNEQAPYLTQVVSFGFQAFI  64

H paragallinarum    36  FLKAKHIAEQ---KRSKQWVTMQF--FQKKQA  62
H ducreyi           65  IKYLIHYFNDNFFSRDKHDVVTEYSAFIEKTL  96

H paragallinarum    63  ILCGVDEAIALLQRFARNPGRIENWA-IKCWR  93
H ducreyi           97  QLEDTGEHIAKLHELGYLPIRIKAIPEGKTVA  128

H paragallinarum    94  RNLPLCNVLIVE-DFMRVLPTEDGVIDG----  120
H ducreyi           129 IKVPVMTIENTHSDFFWLTNYLETLINVSLWQ  160

H paragallinarum    121 -------IYVRRTSVVTNVYNIAMAERSDPPT  145
H ducreyi           161 PMTSASIAFAYRTALIKFANETCDNQEHVPFQ  192

H paragallinarum    146 ENPLFMGGPQCLFYLQPTDG--YAAYIGGATM  175
H ducreyi           193 SHDFSMRGMSSLESAETSGAGHLTSFLGTDTI  224

H paragallinarum    176 QATDAMHQYQAEQ---GMGAYAQCHIPLFNGD  204
H ducreyi           225 PALSFVEAYYGSSSLIGTSIPASEHSVMSSHG  256

H paragallinarum    205 VVEMATLFR----------------------  213
H ducreyi           257 VDELSTFRYLMAKFPHNMLSIVSDTTDFWHNI  288

H paragallinarum    214 -----------LARPICKSILILPLP------  228
H ducreyi           289 TVNLPLLKQEIIARPENARLVIRPDSGNFFAI  320

H paragallinarum    229 -------------------LWLITTMMCLPIV  241
H ducreyi           321 ICGDPTADTEHERKGLIECLWDIFGGTVNQKG  352

H paragallinarum    242 YKYQNTSVKNSMRCALIPPPIWYIATSLPTK-  272
H ducreyi           353 YKVINPHIGAIYGDGVTYEKMFKILEGLQAKG  384

H paragallinarum    273 ---KPRGLDKRICVALRKALDEEGFSTRSSRL  301
H ducreyi           385 FASSNIVFGVGAQTYQRNTRDTLGFALKATSI  416

H paragallinarum    302 SMGFDEKRIREFEAQSVPFDITGVPCGLLWHR  333
H ducreyi           417 TINGEEKAIFKNPKTDDGFKKSQKGRVKVLSR  448

H paragallinarum    334 CLFAQNLHCFYWKAQAKAGRRYRVNPRLSTVH  365
H ducreyi           449 DTYVDGLTSADDFSDDLLELLFEDGKLLRQTD  480

H paragallinarum    366 FYHVFLM         372
H ducreyi           481 FDEIRQNLLVSRTTL  495
```

Figure 25. Alignment of the QAPRTase protein from the plasmid of *H. paragallinarum* with the NADV protein from the plasmid encoding NAD^+ independence in *H.* ducreyi.

The lack of homology between these two sequences indicates the probability that these two plasmids, although similar in function, might have different

lineages. A higher homology (58%) was observed between the ORF encoding QAPRtase from the genome of *Bacillus anthracis* (Fig. 26). Similarty of 65.32% was observed between the aminio acid sequence deduced from the DNA of *H. paragallinarum* plasmid and that of *Bacillus anthracis* (Fig 27).

```
Bacillus anthracis QAPRTase   1    ATGAAAGAAATTGAATTAAAATTAAAAGGTGAAATAAATAGAC 43
NAD+ plasmid QAPRTase         1    ATGAAGGAAATTGAAAAAAAAGAGAAAGTAAATATCACAAGAC 43

Bacillus anthracis QAPRTase   44   TAACAAATAAAACATTTAAATTTGACGAACGCGTTGGAGAAGG 86
NAD+ plasmid QAPRTase         44   GACCATCTCAAACCTTTAATTTGATGAAAAGATAGGAGAAGG 86

Bacillus anthracis QAPRTase   87   TTGGTTCTCTGCCGTTTACTTTTTAAAGACTAGAGAAATCATC 129
NAD+ plasmid QAPRTase         87   GTACTTTTCCGCAAGCTATTTTCTAAAAGCAAAGCACATTGCG 129

Bacillus anthracis QAPRTase   130  GAAGAATTTCGCCCGAAAAGTGTTGTAACGATGCAATTTTTCC 172
NAD+ plasmid QAPRTase         130  GAACAAAAGAGGTCAAAACAATGGGTAACAATGCAATTCTTCC 172

Bacillus anthracis QAPRTase   173  AAAAGGAAAATGCAGTTCTTTGCGGAACGGATGAAGTCATTGC 215
NAD+ plasmid QAPRTase         173  AAAAAAAACAGGCCATTTTATGTGGTGTTGATGAAGCCATCGC 215

Bacillus anthracis QAPRTase   216  ATTGCTACAAACATTCGCTAAAAATCCAGAGGAACTTGAGATT 258
NAD+ plasmid QAPRTase         216  CTTACTACAACGTTTTGCACGGAACCCTGGAAGGATTGAAAAT 258

Bacillus anthracis QAPRTase   259  AACTCTTTAAAA-GATGGCGATAAAATTAGTCCATTTGAAACA 300
NAD+ plasmid QAPRTase         259  TGGGCAATAAAATGTTGGCGACGTAATCTCCCCCTTTGTAAC- 300

Bacillus anthracis QAPRTase   301  GTGTTAACAATTGATGGTCCTTATGAAAATTTCGGATTTTTAG 343
NAD+ plasmid QAPRTase         301  GTATTAATTGTGGAAGATT-TTATGAGGGTTTTGCCAACCGAG 342

Bacillus anthracis QAPRTase   344  AA-GGTGTTATCGATGGGATTTTAGCTCGTCGTACATCGGTTG 385
NAD+ plasmid QAPRTase         343  GATGGTGTTATTGATGGCATTTATGTGCGTCGAACTTCTGTGG 385

Bacillus anthracis QAPRTase   386  CGACAAACGTATATAACGTTGTCCAAGCTGCGCGT--AGCGTA 426
NAD+ plasmid QAPRTase         386  TTACAAATGTATATAACATCGCCATGGCGGAACGTTCAGACCC 428

Bacillus anthracis QAPRTase   427  GATAAAGAAAAACCAGTTATTTTCATGGGAG-ACCGTGATGAT 468
NAD+ plasmid QAPRTase         429  GCCAACGGAAAACCCATTATTT--ATGGGTGGACCGCAATGTT 469

Bacillus anthracis QAPRTase   469  CATTATACACAACAAGCCGGTGACGGTTATGCAGCATACATCG 511
NAD+ plasmid QAPRTase         470  TATTTTACTTG-CAACCCACTGATGGTTACGCCGCCTATATTG 511

Bacillus anthracis QAPRTase   512  GAGGTATGAGCGCACAAGCGACGCATGCGATGAATGAATGGTG 554
NAD+ plasmid QAPRTase         512  GCGGGGCAACAATGCAAGCCACCGATGCAATGCACCAATACCA 554

Bacillus anthracis QAPRTase   555  GGGCAAAAGTGGTATGGGGACAATGCCTCACGCATTAATTCAA 597
NAD+ plasmid QAPRTase         555  AGCGGAACAGGGAATGGGCGCGTACGCACAATGCCACATTCCC 597

Bacillus anthracis QAPRTase   598  ATGTTTAATGGTGATGTTGTGGAAGCGGCAAA-------AGCA 633
NAD+ plasmid QAPRTase         598  TTATTCAACGGTGATGTTGTAGAAATGGCAACCTTATTCAGGC 640

Bacillus anthracis QAPRTase   634  TATC-------ATAAGAAATTC---CCAGAAGATGAATTAGTT 666
NAD+ plasmid QAPRTase         641  TTGCGAGGCCTATATGCAAGTCTATCCTAATTCTCCCCTTACC 683

Bacillus anthracis QAPRTase   667  GTATTAATAGATTACAACAATGATGTCATTACAGATGCACTTC 709
NAD+ plasmid QAPRTase         684  GCTTTGGTTGATTACAACAATGATGTGTTTACCGATAGTTTAT 726

Bacillus anthracis QAPRTase   710  GCGTAGCACGCGAATTTGGATCAACATTAAAAGGTGTACGTGT 752
NAD+ plasmid QAPRTase         727  AAGTACCAAAACACTTCGGTGAAAAACTCTATGCGGTGCGCGT 769

Bacillus anthracis QAPRTase   753  AGATACGTCTCGCACGATGATTGATCAATACTTCATTCGCCAC 795
NAD+ plasmid QAPRTase         770  TGATACCTCCGCCAATATGGTACATCGCTACTTCCTTACCCAC 812

Bacillus anthracis QAPRTase   796  CCAGAAGTACTTGGAACATTCGATCCACGTGGTGTAAACCCAT 838
NAD+ plasmid QAPRTase         813  CAAGAA--ACCTCGGGGGTTGGA--CAAGAGGATTT------- 844

Bacillus anthracis QAPRTase   839  CGCTTGTATTCGCACTTCGTAAAGCACTTGATGAGGAAGGATT 881
NAD+ plasmid QAPRTase         845  -GCGTG-----GCGTTAAGAAAAGCATTAGATGAAGAGGGATT 881

Bacillus anthracis QAPRTase   882  CCAGCATGTAGACATCGTTGTAACTGGTGGATTTGATGAGAAG 924
NAD+ plasmid QAPRTase         882  CTCAACTAGATCAAGCAGGCTTTCCATGGGTTTTGATGAAAAA 924
```

```
Bacillus anthracis QAPRTase   925  CGTATTCGTGAATTTGAAGCTCAAAACGTACCTGTTGATATAT 967
NAD+ plasmid QAPRTase         925  AGAATTAGGGAATTTGAGGCTCAATCTGTGCCTTTCGACATAA 967

Bacillus anthracis QAPRTase   968  ACGGAGTAGGAAGTAGCTTATTAAAAATGAATATTGGCTTTAC 1010
NAD+ plasmid QAPRTase         968  CAGGTGTACCCTGTGG---ATTACTATGGCATCGGTGCCTCTT 1007

Bacillus anthracis QAPRTase   1011 TGGTGATAA----CGTAGAATTAAATGGAAAACCAGAAGCAAA 1049
NAD+ plasmid QAPRTase         1008 TGCTCAAAATCTCCATTGTTTTTACTGGAAAGC-ACAAGCAAA 1049

Bacillus anthracis QAPRTase   1050 AGCTGGTCGTAAATATCGTCCAAACCCACGTTTAGAGCGTGTT 1092
NAD+ plasmid QAPRTase         1050 AGCGGGAAGACGCTACCGTGTTAATCCACGCTTAAGTACGGTA 1092

Bacillus anthracis QAPRTase   1093 CAATTAGAAAAAAGAGAAGATATGTAA 1119
NAD+ plasmid QAPRTase         1093 CATTTTTACCACGTTTTTCTCATGTAA 1119
```

Figure 26. Nucleotide sequence Alignment of the ORF encoding the QAPRTase protein from the plasmid of *H. paragallinarum* with that from *Bacillus anthracis* showing identity 58%.

```
Bacillus anthracis QAPRTase   1    MKEIELKLKGEINRLTNKTFKFDERVGEGWFS 32
NAD+ plasmid QAPRTase         1    MKEIEKKEKVNITRRPSQTFNFDEKIGEGYFS 32

Bacillus anthracis QAPRTase   33   AVYFLKTREIIEEFRPKSVVTMQFFQKENAVL 64
NAD+ plasmid QAPRTase         33   ASYFLKAKHIAEQKRSKQWVTMQFFQKKQAIL 64

Bacillus anthracis QAPRTase   65   CGTDEVIALLQTFAKNPEELEINSLKDGDKIS 96
NAD+ plasmid QAPRTase         65   CGVDEAIALLQRFARNPGRIENWAIKCWRRNL 96

Bacillus anthracis QAPRTase   97   PFETVLTIDGPYENFGFLEGVIDGILARRTSV 128
NAD+ plasmid QAPRTase         97   PLCNVLIVEDFMRVLPTEDGVIDGIYVRRTSV 128

Bacillus anthracis QAPRTase   129  ATNVYNVVQAARSVDKEKPVIFMGDRDDHYTQ 160
NAD+ plasmid QAPRTase         129  VTNVYNIAMAERSDPPTENPLFMGGPQCLFYL 160

Bacillus anthracis QAPRTase   161  QAGDGYAAYIGGMSAQATHAMNEWWGKSGMGT 192
NAD+ plasmid QAPRTase         161  QPTDGYAAYIGGATMQATDAMHQYQAEQGMGA 192

Bacillus anthracis QAPRTase   193  MPHALIQMFNGDVVEAAKAYHKKFPEDELVVL 224
NAD+ plasmid QAPRTase         193  YAQCHIPLFNGDVVEMATLFRLARPICKSILI 224

Bacillus anthracis QAPRTase   225  IDYNNDVITDALRVAREFGSTLKGVRVDTSRT 256
NAD+ plasmid QAPRTase         225  LPLPLWLITTMMCLPIVYKYQNTSVKNSMRCA 256

Bacillus anthracis QAPRTase   257  MIDQYFIRHPEVLGTFDPRGVNPSLVFALRKA 288
NAD+ plasmid QAPRTase         257  LIPPPIWYIATSLPTKKPRGLDKRICVALRKA 288

Bacillus anthracis QAPRTase   289  LDEEGFQHVDIVVTGGFDEKRIREFEAQNVPV 320
NAD+ plasmid QAPRTase         289  LDEEGFSTRSSRLSMGFDEKRIREFEAQSVPF 320

Bacillus anthracis QAPRTase   321  DIYGVGSSLLKMNIGFTGDNVELNGKPEAKAG 352
NAD+ plasmid QAPRTase         321  DITGVPCGLLWHRCLFAQNLHCFYWKAQAKAG 352

Bacillus anthracis QAPRTase   353  RKYRPNPRLERVQLEKREDM 372
NAD+ plasmid QAPRTase         353  RRYRVNPRLSTVHFYHVFLM 372
```

Figure 27. Alignment of the QAPRTase protein from the plasmid of *H. paragallinarum* with the QAPRTase gene from *Bacillus anthracis* showing identity of 41.13% and similarity of 65.32%.

4.10.3 Characterization of quinolinic acid phosphoribosyltransferase

Quinolinic acid phosphoribosyltransferase (QAPRTase) belongs to a family of glycotransferases and a group of enzymes called phosphoribosyltransferases (PRTases). All enzymes in this group show a common fold (type 1) consisting of a central parallel sheet of 5 strands surrounded by alpha helices with 13 residues, which prove important in building and catalysis. PRTases are responsible for the biosynthesis of pyrimidine, purine and pyridine nucleutides as well as the amino acids histadene and tryptophan.(Eads 1997.) The overall structure of (QAPRTase) can be seen in Fig. 28. QAPRTase consist of 372 amino acids, have a molecular weight of 32428.56 dalton and a theoretical pI value of 4.95.

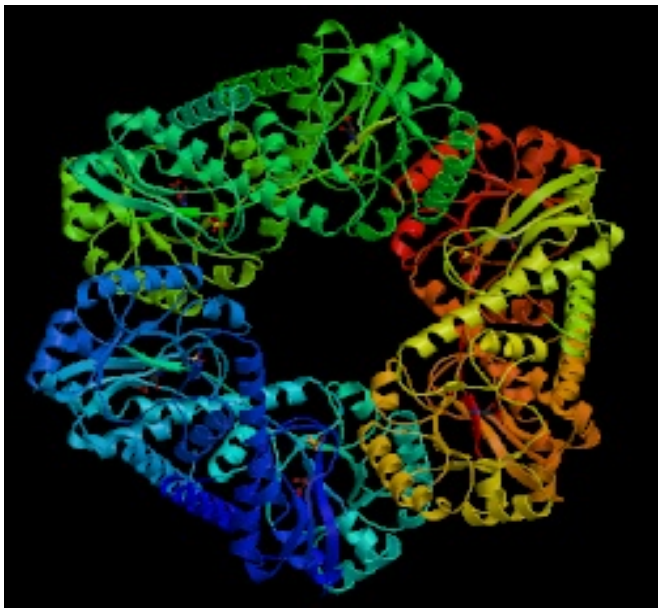

Figure 28. The overall structure of quinolinic acid phosphoribosyltransferase

Species from which QAPRTase has been isolated include *Eschericheria coli*, *Heliobacter pylori*, *Bacillus subtillus*, and *Streptomyces cerevisiae* (http://www.genome.ad.jp/dbget-bin/www-bget?ec:2.4.2.19).

4.10.4 Rep A protein.

Plasmids replicate in an autonomous and self-controlled way through use of the replication machinery of the host. Within plasmids there is an essential region containing genes that are necessary during replication. One such gene, which is common in plasmids, encodes for a protein involved in the initiation of replication, usually termed Rep proteins. Rep proteins recognize specific sequences at the origin of replication; similar to the DnaA initiator protein in bacterial chromosomal replication, where a nucleoprotein initiation complex is generated in which essential macromolecular interactions take place (Rep-DNA, Rep-Rep and Rep with other initiation proteins of the host.) (Bramhill and Kornberg, 1988). This protein exists in a monomer-dimer equilibrium, although it is mainly in the monomeric form. However when the *repA* gene is over-expressed, replication is inhibited. Inhibition under over-expression can assumedly be explained if elevated concentrations of RepA promote its dimerisation and that the RepA dimers would then hinder the interaction of the active RepA monomeric forms with the origin (Ingmer and Cohen, 1993).

The RepA protein encoded on the plasmid conferring NAD^+ independence in *H. paragallinarum* showed similarity to the RepA encoded on the *Pasteurella multocida* PJR2 plasmid on both DNA and protein level (Fig. 29-30)

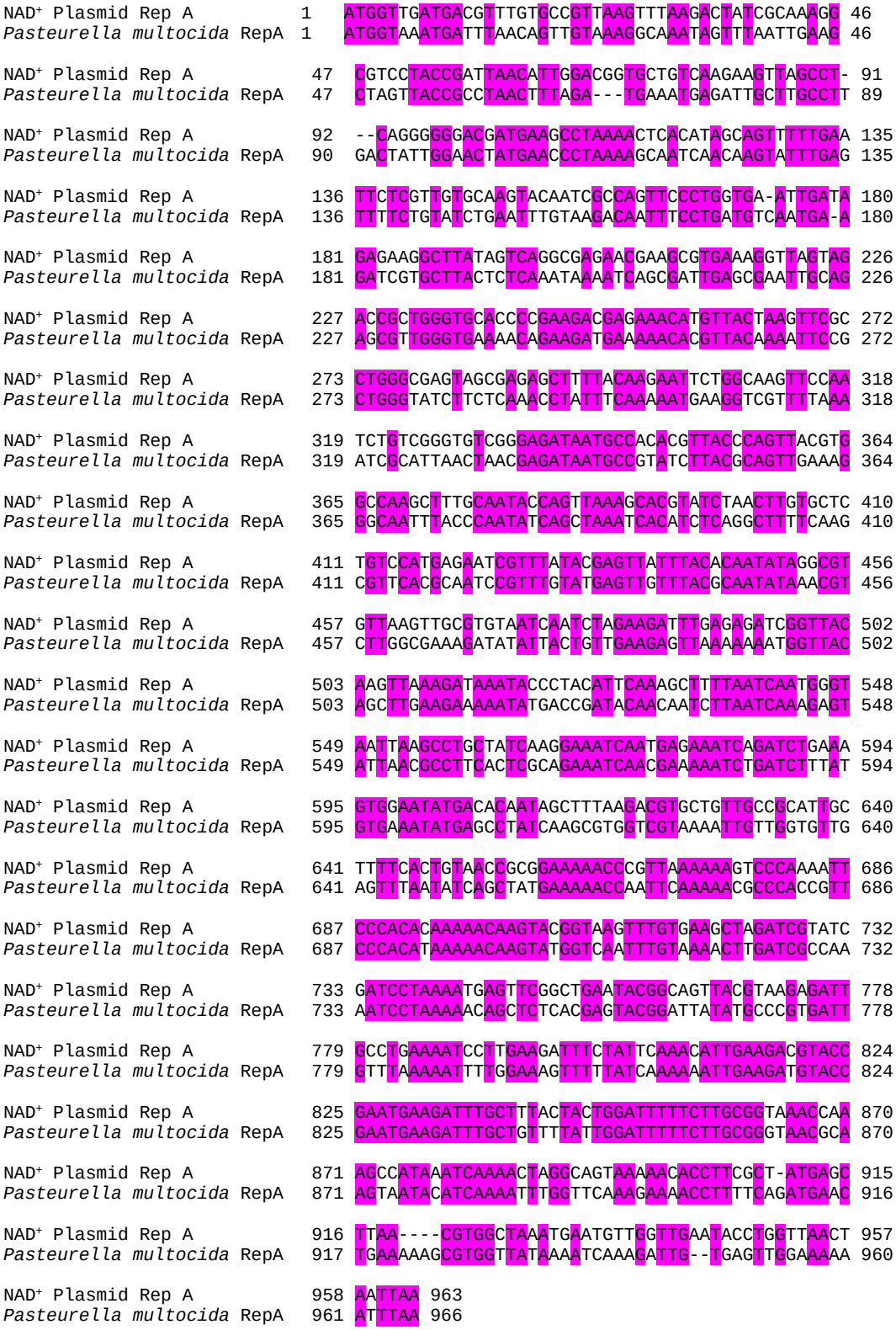

Figure 29. Nucleotide Sequence alignment of the RepA protein found on the *Pasteurella multocida* PJR2 plasmid with RepA protein of *H. paragallinarum* showing an identity of 60,2%.

Figure 30. Amino acid Sequence alignment of the RepA protein found on the *Pasteurella multocida* PJR2 plasmid with RepA protein of *H. paragallinarum*. Showing 55.45% identity 74.76% similarity.

Comparison of the sequence encoding the repA protein of the *Pasteurella multocida* PJR2 plasmid to that encoded on the *H. paragallinarum* plasmid revealed that the similarity between the amino acid sequences (74.76%) is relatively high considering the identity on DNA level is only 60.2%. Both of these sequences have relatively low GC content of 34% (*Pasteurella multocida* PJR2) and 41% (*H. paragallinarum* plasmid).

Chapter 5

Conclusion

NAD^+ independence of *H. paragallinarum* is conferred through a plasmid entity approximately 6Kb in size. The plasmid was successfully isolated from all NAD^+ independent *H. paragallinarum* strains examined during this study. Although different serotypes of the same bacterium can possess different plasmids, all plasmid species isolated from the *H. paragallinarum* strains were approximately 6000 base pairs in length and proved to be identical through restriction analysis.

The plasmid was found to contain open reading frames encoding for six proteins namely: RepA, Haemocin, Haemocin immunity factor, Haemocin transporter, Haemocin structural protein protein and Quinolinic acid phosphoribosyltransferase (QAPRTase). QAPRTase was identified as the enzyme essential to NAD^+ independence of *H. paragallinarum.*

The plasmid encoding NAD^+ independence in *H. paragallinarum* differs significantly in both nucleotide and deduced amino acid sequence from the plasmid encoding for similar metabolic function in *H. ducreyi*. The possibility that plasmid was procured from different sources can therefore not be excluded.

The plasmid sequence was found to consist of a high percentage of Adenine and Thymine. The low success rate in subcloning of plasmid DNA achieved throughout this study may be accounted for by considering the structural characteristics of AT rich DNA. It is therefore possible that these clones could behave like transcriptional promoters in *E. coli.* Transcription driven by these inserts may then proceed into the vector and interfere with its replication or expression of drug resistance

Chapter 6

Summary

Members of the family *Pasteurellaceae* are classified in part by whether or not they require NAD^+ supplement for growth on laboratory media. It is known that this phenotype is determined by a plasmid whose presence allows NAD^+-independent growth of *Haemophilus paragallinarum*.

In this study, this 6-kb plasmid, which was previously shown to be responsible for NAD+ independent growth of *H. paragallinarum* on defined media, was isolated. Isolated plasmid DNA was shredded by sonification and subcloned into vector PGEM-T easy. The recombinant plasmid was transformed into *E.coli* the transformants isolated were sequenced.

Sequence analysis revealed one open reading frame of 1119bp that is predicted to encode a protein with a molecular mass of 43kD. Compared with the sequence databases, this protein was found to have significant sequence homology to Quinolinate phosphoribosyltransferase of *Bacillus anthracis* this enzyme is responsible for the production of nicotinic acid mononucleotide (NAMN) from nicotinate and quinolinate. A 3284bp nucleotide fragment of the plasmid revealed four additional open reading frames. Proteins encoded on this fragment of the plasmid all have significant homology to proteins from *H. influenzae* of which all have functions related to production and immunity of the bacteriocin haemocim. This bacteriocin produced by most type b-encapsulated strains of *H. influenzae*, is toxic to virtually all non-type b strains of *H. influenzae* idependent of encapsulation. This bacteriocin is thought to inhibit DNA synthesis, of susceptible strains. Purification of this bacteriocin and testing its toxicity to other pathogens as a possible antimicrobial drug might form the bases of a future study.

Previous work has indicated that plasmid bearing strains of *H. paragallinarum* are less virulent, thus creating the possibility that more virulent wild type strains can be transformed and used as live vaccines. The influence of transformation with this plasmid on other members of the family *Pasteurellaceae* and the possibility of creating live vaccines should be further investigated

Since species of the genus *Haemophilus* cannot easily be transformed with plasmid, this naturally occurring plasmid could be modified to create a vector, which has specific application in the transformation of *Haemophilus* species.

Chapter 7

Opsomming

Die klassifikasie van bakteriese spesies van die familie *Pasteurallaceae* word gedeeltelik gebaseer op die afhanklikheid van NAD^+ vir groei op laboratorium media. Dit is bekend dat die NAD^+ onafhanklike fenotipe van *Haemophilus paragallinarum* onderskei kan word deur die teenwoordigheid van ’n plasmied.

In hierdie studie is die plasmied van ongeveer 6kbp wat NAD^+ onafhanklike groei op bepaalde media bevorder, geïsoleer. Plasmied DNA is deur middel van sonifikasie gefragmenteer en in vektor pGEM-T Easy gekloneer. Rekombinante plasmiede in *E. coli* getransformeer, geïsoleer en aan basisbaarvolgordebepaling onderwerp.

Ontleding van die basispaarvolgorde toon dat die plasmied ‘n oopleesraam van 1119bp bevat, wat kodeer vir ’n proteïen met ‘n geraamde molekulêre massa van 43kD. Na vergelyking met basispaarvolgorde databasisse is gevind dat hierdie proteïen betekenisvolle homologie met quinoliensuur fosforibosieltransferase, geïsoleer uit *Bacillus anthracis,* toon. kwinolienaat fosforibosieltransferase is verantwoordelik vir die katalitiese produksie van nikotiensuur mononukleotied vanaf nikotienaat en kwinolienaat.

Daar is verder bevind dat die plasmied ’n 3284bp fragment bevat waarop vier oopleesrame aangetref word. Die proteïne waarvoor die leesrame op hierdie fragment kodeer toon betekenisvolle homologie met proteïne geïsoleer vanuit *Haemophilus influenzae,* wat betrokke is by die produksie van en immuniteit teen die bakteriosien haemosien. Die meerderheid tipe b-gekapsuleerde stamme van *H. influenzae* produseer haemosien, terwyl die bakteriosien die groei van bykans alle nie-tipe b, beide gekapsuleerde en nie-gekapsuleerde, stamme belemmer. Daar word vermoed dat die inhibisie effek deur

haemosien bewerkstellig word deurdat die bakteriosien die DNS sintese van vatbare stamme onderdruk. In ‘n verdere studie kan gefokus word op die isolasie en suiwering van haemosien en die bepaling van die mate waarin die bakteriosien die groei van ander patogeniese bakterieë belemmer. Verdere ondersoek kan ook ingestel word na die moontlike aanwending van haemosien as ‘n anti-mikrobiese middel.

Daar is vantevore bevind dat stamme van *H. paragallinarum* waarin plasmiede voorkom minder virulent is na aanleiding van die waarneming dat hoenders wat met hierdie stamme besmet is, minder intense simptome vertoon. Hierdie waarneming kan moontlik verklaar word in terme van die onderdrukking van sekondêre invektiewe mikro-organismes deur haemosien. Verdere ondersoeke kan dus gefokus word op die relatiewe dominansie van plasmied bevattende stamme van *H. paragallinarum*, om bogenoemde hipotese te toets. Dit is ook moontlik dat plasmied bevattende stamme inherent minder virulent is, wat dit moontlik maak om meer virulente wilde stamme te transformeer en sodoende lewende entstowwe te skep. Die uitwerking van transformasie met die geïsoleerde plasmied op ander spesies in die familie *Pasteurellaceae* en die moontlike produksie van lewende entstowwe behoort verder ondersoek te word.

Spesies in die genus *Haemophilus* kan slegs met moeite met plasmiede getransformeer word. Die plasmied geïsoleer vanuit *H. paragallinarum* kan gemodifiseer word en sodoende dien as ‘n vektor vir die transformasie van *Haemophilus* spesies.

Chapter 8

REFERENCES:

Axelsson, L., and Holck, A. 1995. The genes involved in production of and immunity to sakacin A, a bacteriocin from *Lactobacillus sake* *J. Bacteriol* 177:2125-2137.

Bibberstein, E.L. and White, D.C. 1963. A proposal for the establishment of two new *Haemophilus* species. *J. Med. Microbiol*. 2:75-78.

Blackall, P.J. 1983. An evaluation of the methods for the detection of carbohydrate fermentation in avian *Haemophilus* species. *J. Microbiol. Met*. 1: 275-281.

Blackall, P.J. and Eaves, L.E. 1988. Serological classification of Australian and South African isolates of *Haemophilus paragallinarum.* *Aus. J. Vet.* 63:362-363.

Blackall, P.J. 1989. The avian haemophili. *Clin. Microbiol*. 2:270-277.

Blackall, P.J., Rogers, D.G. and Yamamoto, R. 1990a. Outer-membrane proteins of *Haemophilus paragallinarum*. *Avian. Dis*. 34:871-877.

Blackall, P.J., Eaves, L.E. and Rogers,.D.G. 1990b. Proposal of a new serovar and altered nomenclature for *Haemophilus paragallinarum* in the Kume hemagglutinin scheme. *J. Clin. Microbiol*. 28: 1185-1187.

Blackall, P.J. 1991. An evaluation of the cross protection afforded by inactivated infectious coryza vaccines. *Aus. J. Vet.* 68: 266-267.

Blackall, P.J. 1999 Infectious Coryza: Overview of the Disease and New Diagnostic Options *Clinical Microbiol Rev.*12:627-632.

Bragg, R.R., Coetzee, L. and Verschoor, J.A. 1993a. Monoclonal antibody characterization of South African field isolates of *Haemophilus paragallinarum. Onderstepoort. J. Vet. Res.* 60:181-187.

Bragg, R.R., Coetzee, L. and Verschoor, J.A. 1993b. Plasmid encoded NAD independence in some South African isolates of *Haemophilus paragallinarum. Onderstepoort J. Vet. Res.* 60; 147-152.

Bragg, R.R., Purdan, G., Coetzee, L. and Verscoor, J.A. 1995. Effects of transformation on the hemagglutinins of *Haemophilus paragallinarum Onderstepoort J.Vet. Res.* 62:261-270.

Bramhill, D and Kornberg, A. 1988. A model for initiation at origins of DNA replication. *Cell* 54:915-918.

Brown, M.R.W. and Williams. 1985. The influence of environment on envelope properties affecting survival of bacteria in infections. *Ann Rev Microbiol.* 39: 527-556.

Buys, S.B. 1982. Die bereiding van `n bakterien teen *Haemophilus paragallinarum* besmetting, geskik vir Suid-Afrikaanse omstandighede. M.Med.Vet. (Bact.) thesis. Faculty of Veterinary Science, University of Pretoria.

Carlisle, E. 1998. Isolation and characterization of the plasmid Bsc Hons Department of Genetics, University of Pretoria.

Chen, X., Miflin, J.K., Zang, P. and Blackall, P.J. 1996. Development and application of DNA probes and PCR tests for *Haemophilus paragallinarum*. *Avian. Dis.* 40: 398-407.

Chikami, G.K., Fierer, J. and Guiney, G.G. 1985. Plasmid – mediated virulence in *Salmonella dublin* demonstrated by use of a Tn5-*ori*T construct. *Infect. Immun.* 50:420-424.

Cynamon, M.H., Sorg, T.B. and Patapow, A. 1988. Utilization and metabolism of NAD by *Heamophilus parainfluenzae. J. Gen. Microbiol*. 134: 2789-2799.

De Blieck, L. 1931. A haemoglobinophilic bacterium as the cause of contagious catarrh of the fowl. *J. Vet*. 88: 9-13.

Delaplane, J.P., Erwin, L.E. and Stuart, H.O. 1938. The effect of the X-factor, of sodium chloride, and of the composition of the nutrient medium upon the growth of the fowl coryza bacillus, *Haemophilus gallinarum. J. Agric. Res.* 56: 919-926.

Devedjiev, Y., Singh, R., Brouillette, C.H., Brouillette, W., Muccio, D., DeLucas, L. and Jedrzejas, M. 1998. Catalytic cycle of NAD^+ synthetase viewed by X-ray structures of kinetic intermediates. *ACA meeting Arlington Virginia.*

Eads, J.C. 1997. A new function for a common fold: the crystal structure of quinolinic acid phosphoribosyltransferase. *Structure* 5:47-58.

Eaves, L.E., Rogers, D.G. and Blackall P.J. 1989. Comparison of hemagglutinin and agglutinin schemes for the serological classification of *Haemophilus paragallinarum* and proposal of a new hemagglutinin serovar. *J. Clin. Microbiol.* 27:1510-1513.

Elliot, C. and Lewis, M.R. 1934. A hemophilic bacterium as cause of infectious coryza in the fowl. *Am. J. Vet. Med. Ass.* 37: 878-888.

Forbisher, M. 1970. Fundamentals of Microbiology 8th edition. W.B. Saunders Company. Philadelphia. Pp172-173.

Foster, J.W. and Moat, A.G. 1980. Nicotinamide adenine dinucleotide biosynthesis and pyridine nucleotide cycle metabolism in microbial systems. *Microbiol. Rev*. 44: 83-105.

Fowkes, S. 1992. Tryptophan metabolism in chronic disease and AIDS. *Forefront Hlth. Invest.* 36-45

Fujiwa, H. and Kanno, S. 1965. Histopathological studies on infectious coryza of chickens, II. Findings in experimentally infected cases. *Nat. Inst*. *Avian. Hlth.* Quart. 5: 86-96.

Gromkova, R. and Koornhof, H.J. 1990. Naturally occurring NAD independent *Heamphilus parainfluenza. J. Gen. Microbiol.* 136:1031-1035.

Hobb, R.I., Tseng, H., Downes, J.E., Terry, T.D., Blackall, P.J., Takagi, M. and Jennings, M.P. 2002. Molecular analysis of a haemagglutinin of *Haemophilus paragallinarum*. *Microbiol.* 148:2171-2179.

Hollander, R. and Mannheim, G. 1975. Characterization of Hemophilic and related bacteria by their respiratory quinines and cytochromes. *J. Syst. Bacteriol.* 25:102-107.

Holt, J.G. (ed.) 1994. Bergey's manual of determinative bacteriology, 9th Ed. The Williams and Wilkins Co, Baltimore, Md.

Horner, R.F., Bishop G.C. and Haw, C. 1992. An upper respiratory disease of commercial chickens resembling infectious coryza, but caused by a V factor independent bacterium. *Aviam.Path.* 21:421-427.

Horner, R.F., Bishop, G.C., Jarvis, C.J. and Coetzee, L. 1995. T.H.T. NAD (V-factor) Independent and typical, *Haemophilus paragallinarum* infection in commercial chickens: a five-year study. *Avian. Path.* 24:453-463.

Ingmer, H., and Cohen, S. N. 1993. Excess intracellular concentration of the pSC101 RepA protein interferes with both plasmid DNA replication and partitioning. *J. Bacteriol*. 175: 7834 – 7841

Iritani, Y., Hidaka, S. and Katagiri, K. 1977. Production and properties of hemagglutinin of *Haemophilus paragallinarum. Avian. Dis.* 21:39-49.

Iritani, Y., Katagiri, K. and Tsuji, K. 1978. Slide agglutination test of *Haemophilus gallinarum* antigen treated by trypsin to inhibit spontaneous agglutination. *Avian. Dis.* 22:793-797.

Iritani ,Y., Iwaki, S. and Yamaguchi, T. 1981. Biological crude activities of crude polysaccharide extracted from two different immunotype strains of *Haemophilus paragallinarum* in chickens. *Avian. Dis*. 25:29-37.

Kahn, D.W. and B.M. Anderson. 1986. Characterization of *Haemophilus influenzae* nucleotide pyrophosphatase. An enzyme of critical importance for growth of the organism. *J. Biol. Chem*. 261: 6016-6025.

Kato, K. and Tsubahara, H. 1962. Infectious coryza of chickens . Identification of isolates. *Bulletin. Nat. Inst. Animal. Hlth.* 45:21-26.

Kato, K. 1970. Nature of hemagglutination-inhibition antibody response and its relationship to protection in infectious coryza. *J. Vet. Sci.* 32:263.

Kasarov, LB. and Moat, A.G. 1973. Biosynthesis of NAD in *Haemophilus haemoglobinophilus. Biochem. Biophys*. 320:372-378.

Kume, K., Sawata, A., Nakase, Y. and Matsumoto, M. 1983a. Serological classification of *Haemophilus paragallinarum* with a hemagglutinin system. *J. Clin. Microbiol.* 17:958-964.

Kume, K., Sawata, A. and Nakai, T. 1983b. Serologic and immunologic studies on three types of hemagglutinin of *Haemophilus paragallinarum* serotype 1 organisms. *J. Vet. Sci.* 45:783-792.

Lipuma, J.J., Richman, H., and Stull, T.L. 1990. Haemocin, the bacteriocin produced by *Haemphilus influenzae*: species distribution and role in colonization. *Infect. Immun.* 58:1600-1605.

Matsumoto, M. 1988. Prevention of infectious coryza. Vineland Update (25).

Martin, P.R., Shea R.J., and Mulks, M.H. 2001. Identification of a plasmid-encoded gene from *Haemophilus ducreyi* which confers NAD Independence. *J. Bacteriol.* 183:1168-1174.

Marks. D.B., Marks, A. D. and Colleen, D. 1996. Basic Medical Biochemistry A clinical approach. Lippincott Williams and Wilkins. pp305-306.

Mouahid, M., Bisgaard, M ., Morley, A.J., Mutters, R. and Mannheim, W. 1992a. Occurrence of V-factor (NAD) independent strains of *Haemophilus paragallinarum*. *Vet. Microbiol.* 31:363-368.

Mouahid M., Hinz, K.H., Engelhard, E., Mutters, R., Mannheim, W. 1992b. Characterization of *Haemophilus paragallinarum*by analysis of whole cell carbohydrates, fatty acids and phospholipids. *Avian Path.* 21:127-136.

Murley, Y.M., Edlind, T.D., Poszgay, J.M. and LiPuma, J.J. 1997. Cloning and characterization of the immunity gene of *Haemophilus influenzae. J. Bacteriol.* 179: 1684-1689.

Musser, J.M., Barenkamp, S.J., Granoff, D.M., and Selander, R.K. 1986 Genetic relationships of serologically nontypable and serotype b strains of *Haemophilus influenzae. Infect. Immun.*52:183-191

Niven, D.F. and Lèvesque, T. 1988. V-factor–dependent growth of *Actinobacillus pleuroneumoniae* biotype 2 (Bertschinger 2008/76). *Int. J. Syst. Bacteriol.* 38:319-320.

Ogunnariwo, J.A. and Schryvers, A.B. 1992. Correlation between the ability of *Haemophilus paragallinarum* to acquire ovotransferrin-bound iron and the expression of ovotransferrin -specific receptors. *Avian. Dis.* 36:655-663.

O`Rielly, T. and Niven D.F. 1986. Defining the metabolic and growth responses of porcinine Haemophili to exogenous pyridine nucleotides and precursors. *J. Gen. Microbiol.* 132:807-818

Otsulki, K. and Iritani, Y. 1974. Preparation and immunological response to a new mixed vaccine composed of inactivated Newcastle disease virus, inactivated infectious bronchitis virus, and inactivated *Haemophilus paragallinarum. Avian. Dis.* 18:297-304.

Page, L.A. 1962. *Haemophilus* infection in chickens 1. Characteristics of 12 *Haemophilus* isolates recovered from diseased chickens. *Am. J. Vet. Res.* 23:85-95.

Razin, S.V., Ioudinkova, E.S., Trifonov, E.N. and Scherrer, K. 2001. Non-clonability correlates with genomic instability: A case study of a unique DNA Region. *J. Mol. Biol*. 307:481-486.

Rimler R. B., Shotts, E. B., Brown, J. and Davis, R. B. 1976. The effects of atmospheric conditions on the growth of *Haemophilus gallinarum* in a defined medium. *J. Gen. Microbiol*. 92:405-409.

Rimler R.B. 1979. Studies on the patogenic avian *haemophili. Avian. Dis.* 23:1006-1018.

Sambrook, J., Fritsch , E.F. and T. Maniatis. 1989. Molecular cloning : A labratory manual. 2nd Ed Cold spring harbour laboratory press. Cold Spring Harbor.NY. Vol1, pp1.25-1.28

Sawata, A., Kume, K. and Nakase, Y. 1978. *Haemophilus* infections in chickens 2 types of *Haemophilus paragallinarum* isolates from chickens with infectious coryza, in relation to *Haemophilus paragallinarum* strain No. 221. *J. Vet. Sci.* 40:645-652.

Sawata, A., Kume, K. and Nakase, Y. 1979. Antigenic structure and relationship between serotypes 1 and 2 of *Haemophilus paragallinarum*. *Am. J. Vet. Res.* 40: 1450-1453.

Sawata, A., Kume, K. and Nakase, Y. 1980. Biological and serological relationships between Page`s and Sawata`s serotyping of *Haemophilus paragallinarum*. *Am. J. Vet. Res.* 41: 1901-1904.

Sawata, A., Kume, K. and Nakase, Y. 1982. Hemagglutinin of *Haemophilus paragallinarum* serotype 2 organisms: Occurrence and immunologic properties of hemagglutinin. *Am. J. Vet. Res.* 43: 1311-1314.

Sawata, A. and Kume, K. 1984a. Hemagglutinins of *Haemophilus paragallinarum* serotype 1 organisms. *J. Vet. Sci.* 45: 632-635

Sawata, A., Kume, K. and Nakase, Y. 1984b. Susceptibility of *Haemophilus paragallinarum* to bacterial activity of normal and immune chicken serum. *Japn. J. Vet. Sci.* 46:805-813.

Sneath, P.H.A. and Jonson, R. 1973 Numerial taxonomy of *Haemophilus* and related bacteria. *Int. J. Syst. Bacteriol.* 23:405-418.

Snipes, K.P., Hansen, L.M. and Hirsh, M. 1988. Plasma-and-iron-regulated expression of high molecular weight outer membrane proteins by *Pasturella multocida. Am. J. Vet*. 49:1336-1338.

Starr, C. and Teggart, R. 1995. Biology the unity and diversity of life. 7th Ed Wadsworth publishing company, Belmont, California. pp99-100.

Subramang, H.S., Doheerty, AJ., Ashford, and Wigley D.B. 1996. Catalytic domain of an NAD-dependent DNA ligase. *Cell.* 85:607-615

Susumu, U., Nagasawa, Y., Suzuki, T. and Tajima, M. 1982. Adhesion of *Haemophilus paragallinarum* to cultured chicken cells. *Microbiol. Immunol.* 26(11). 1007-1016

Tang, X., Nakata, Y., Li, H.O., Zang, M., Gao, H., Fujuta, A., Sakatsume, O., Ohta, T., Yokoyama, K. (1994). The optimization of preparations of competent cells for transformation of *E. coli*. *Nucleic Acid Research*, 22, 2857-2858.

Taole, M., Albertyn, J., Van Heerden, E. and Bragg, R.R. 2002. Virulence of South African isolates of *Haemophilus paragallinarum.* Part 3: experimentally produced NAD – independent isolate. *Onderstepoort. J. Vet. Res.* 69:189-196

Takagi, M., Ohmae, K., Hirayama, N. and Ohata, S. 1991. Expression of hemagglutinin of *Haemophilus paragallinarum* serotype A in *Escherichia coli. J. Vet. Med. Sci.* 53:917-920.

Verschoor, J.A., Coetzee, L. and Visser, L. 1989. Monoclonal antibody characterization of two field strains of *Haemophilus paragallinarum* isolated from vaccinated layer hens. *Avian. Dis.* 33: 219-225.

Wheat, R. W. and Pittman, J.D. 1960. Degradation of certain nucleotides by *Haemophilus influenzae. J. Bacteriol.* 79:137-141.

Windsor, H.M., Gromkova, R.C. and Koornof, H.J. 1991. Plasmid mediated NAD independence in *Haemophilus parainfluenza. J. Gen. Microbiol.* 137:2415-2421.

Yamaguchi, T. and Iritani, Y. 1989. Hemagglutinating activity and immunological properties of *Haemophilus paragallinarum* field isolates from Japan. *Avian. Dis.* 33:511-515.

Yamamoto, R. 1981. Infectious coryza: diseases of Poultry, 8th Ed. Hofstad, M.S., Barnes, H.J., Calnek, B.W., Reid, W.M., and Yorder, E.R., (Eds). Iowa State University Press. Pp 178-186.

Yamamoto, R. 1984. Infectious coryza: diseases of Poultry, 9th Ed. Hofstad, M.S., Barnes, H.J., Calnek, B.W., Reid, W.M., and Yorder, E.R., (Eds). Iowa State University Press. Pp 180-195.

Yamamoto, R. 1991. Infectious coryza: diseases of Poultry, 10th Ed. Hofstad, M.S., Barnes, H.J., Calnek, B.W., Reid, W.M., and Yorder, E.R., (Eds). Iowa State University Press. Pp 186-195.

Ziegler, M. and Oei, S.L. 2001. Acellular survival switch: poly (ADP – ribosyl) ation stimulates DNA repair and silences transcription. *Bioessays* 23: 543-548

"But in science the credit goes to the man who convinces the world,
not to the man to whom the idea first occurs"

-Francis Darwin