MOLECULAR AND BIOCHEMICAL CHARACTERISATION OF A *FUGU* NITRIC OXIDE SYNTHASE

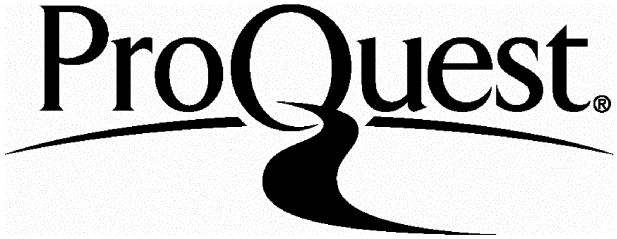
Natalie Jane Wilson

A thesis submitted in accordance with the requirements of the degree of Doctor of Philosophy at the University of London.

The Wolfson Institute for Biomedical Research

University College London

March, 2000.



This thesis is dedicated to my parents with love and thanks.

ABSTRACT.

A neuronal nitric oxide synthase (NOS1) gene was isolated from the the model vertebrate genome of the Japanese pufferfish *Fugu rubripes*. The *Fugu* NOS1 gene was cloned by screening of a genomic cosmid library and PCR of genomic DNA. It was sequenced multiple times on both strands and conserved co-factor/substrate binding sites for haem, L-arginine, Ca^{2+}/Calmodulin, FAD, FMN and NADPH were identifed. In addition, a CaM inhibitory sequence present only in constitutive isoforms was identified along with a PDZ domain specific to nNOS isoforms. At 22,203 bp the *Fugu* NOS1 gene shows a reduction in size of ~7.2 times when compared to the human NOS1 gene. Full-length *Fugu* NOS1 cDNA (*f*NOS) was isolated from the brain of another species of Japanese pufferfish, *Takifugu poecilonotus*, by 5′ and 3′ RACE and PCR. *f*NOS was cloned into the baculovirus-transfer vector pVL1393, sequenced and transfected into *Spodoptera frugiperda* clone 21 cells (*Sf*21) for expression studies. *f*NOS showed the highest amino acid identity to the mammalian and *Xenopus laevis* nNOS isoforms (74-75%). Amino acid identity dropped to 58% when *f*NOS was compared to eNOS isoforms, 54% for iNOS isoforms and 47-52% for the reported invertebrate NOS isoforms. Expression of the recombinant *Fugu* NOS1 protein (*F*NOS) by pVL*f*NOS-infected *Sf*21 cells was detected by western blotting and immunocytochemical staining with a monoclonal anti-nNOS antibody. Expression of functional *F*NOS was greatest after 24 hours as measured by the spectrophometric conversion of oxyhaemoglobin to methaemoglobin by NO on a dual-wavelength spectrophotometer. In the presence of FAD, NADPH, and BH_4, *F*NOS activity was dependent on the addition of L-arginine and was inhibitable by the NOS inhibitors L-NMMA, L-Thiocitrulline and 1400W. Partial purification of *F*NOS was achieved using a 2′5′ ADP sepharose column and 10mM NADPH for elution.

CONTENTS.

LIST OF TABLES.

LIST OF FIGURES.

Chapter 1.

Chapter 3.

Chapter 4.

Chapter 5.

Appendix.

LIST OF ABBREVIATIONS.

1400W	N-(3-(Aminomethyl)benzyl)acetamidine
Abs	Absorbance
ACh	Acetylcholine
ADP	Adenosine diphosphate
APS	Ammonium persulphate
AP	Adaptor primer
ATP	Adenosine triphosphate
BH4	(6R)-5,6,7,8-tetrahydro-L-biopterin (tetrahydrobiopterin)
bp	Base pair(s)
CaM	Calmodulin
cAMP	Cyclic 3′,5′-adenosine monophosphate
cDNA	Complementary DNA
cGMP	Cyclic 3′,5′-guanosine monophosphate
Ci	Curies
cM	Centimorgan
CNS	Central nervous system
Da	Daltons
dATP	2′-deoxyadenosine-5′-triphosphate
dCTP	2′-deoxycytidine-5′-triphosphate
DEPC	Diethyl pyrocarbonate
dGTP	2′-deoxyguanosine-5′-triphosphate
DHR	Discs-large Homologous Region

DHFR	Dihydrofolate reductase
DIG	Digoxygenin
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleoside triphosphates
DOE	United States Department of Energy
DTE	Dithioerythritol
DTT	Dithiothreitol
dTTP	2′-deoxythymidine-5′-triphosphate
dUTP	2′-deoxyuridine-5′-triphosphate
EC	European Community
E. coli	*Escherichia coli*
EDTA	Ethylenediamine tetra-acetic acid
EDRF	Endothelium-derived relaxing factor
EGTA	Ethylene glycol bis (β-aminoethyl ether) N, N, N′, N′-tetra-acetic acid
eNOS	Endothelial NOS cDNA/protein
EtBr	Ethidium bromide
EtOH	Ethanol
FAD	Flavin adenine dinucleotide
FCS	Foetal calf serum
FMN	Flavin mononucleotide
*f*NOS	Full-length *Fugu* NOS1 cDNA
*F*NOS	Recombinant *Sf*21-expressed *Fugu* NOS1
Fugu	*Fugu rubripes*

GCG	Genetics Computer Group
Hb	Haemoglobin
HGMP	Human Genome Mapping Project
HGP	Human Genome Project
HHMI	Howard Hughes Medical Institute
HRPO	Horseradish peroxidase
HUHO	Human Genome Organisation
IC_{50}	Concentration of inhibitor at which enzyme acitivity is reduced by 50%
ICRF	Imperial Cancer Research Fund
IMS	Industrial methylated spirits
iNOS	Inducible NOS cDNA/protein
IPTG	Isopropyl thio-β-D-galactosidase
IS	Immuno Staining
kb	Kilo base pairs
kDa	Kilo daltons
LB	Luria broth
LDL	Low density lipoprotein
L-NIL	L-N^6-(1-Iminoethyl)-L-lysine
L-NIO	L-N^5-(1-Iminoethyl)-L-ornithine
L-NMMA	N^G-monomethyl-L-arginine
LPS	Lipopolysaccharide
MAF	Macrophage activating factor
Mb	Mega base pairs

min	minute(s)
MRC	Medical Research Council
mRNA	Messenger RNA
NADPH	Nicotinamide adenine dinucleotide phosphate
NADPHd	NADPH diaphorase
NIH	National Institutes of Health
NMDA	N-methyl-D-aspartic acid
NNA	N^G-nitro-L-arginine
nNOS	Neuronal NOS cDNA/protein
NO	Nitric Oxide
NO_3^-	Nitrate
NO_2^-	Nitrite
NO_x	Total nitrogen oxides
NOS	Nitric oxide synthase
NOS1	Neuronal NOS gene
NOS3	Endothelial NOS gene
NOS2	Inducible NOS gene
NRC	National research Council
OD	Optical density
OTA	Office of Technology Assessment
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDZ	Acronym of PSD-95, Dlg, Zo1

pfu	Plaque forming units
PMSF	Phenyl methyl sulphonyl fluoride
RACE	Rapid amplification of cDNA ends
RNA	Ribonucleic acid
RNase A	Ribonuclease A
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
*Sf*21	*Spodoptera frugiperda* clone 21 cells
sGC	Soluble guanylate cyclase
SNP	Sodium nitroprusside
STS	Sequence Tagged Site
TB	Terrific broth
TEMED	N, N, N′, N′-tetramethylethylenediamine
UV	Ultra violet
WB	Western blot
X-gal	5-bromo-4-chloro-3-indolyl-β-D galactosidase

ACKNOWLEDGMENTS.

I would like to thank my supervisors Professor Ian Charles and Dr. Mike Lockyer for their encouragement, guidance and advice throughout my studies.

I thank Glaxo Wellcome for funding this project and Professors Ian Charles and Kenneth Powell for providing the additional financial support that allowed me to complete these studies.

Special thanks go to Dr. David Moss for his advice, patience and support during these studies and for taking the time to read this thesis.

I would like to express my gratitude to Neale Foxwell for his technical assistance with the *Fugu* NOS expression and activity work, in particular the NOS assay.

I am indebted to Dr. Shigero Sato, of the Kitasato University School of Fisheries Sciences in Japan, for the gift of pufferfish material and Drs. Saturo Nagata and Stuart McDonald for their assistance in arranging its safe arrival.

Many thanks to Dr. Greg Elgar of the *Fugu* project, MRC HGMP Resource Centre, for his advice, the gift of *Fugu* genomic DNA and the sequence-scanning of twelve *Fugu* cosmid clones.

Thanks are due to all my friends at the Wolfson Institute for Biomedical Research. In particular, I am thankful to my 'support team' from Wakefield Street for their never-ending encouragement and sense of humour and also to Barb who definitely helped to keep me sane.

Finally, biggest thanks go to my parents for their continued support and encouragement and for helping me to get this far.

CHAPTER 1

INTRODUCTION.

CONTENTS.

1.1 The Human Genome Project.

In 1985 Robert Sinsheimer convened a meeting to discuss the first serious proposal to start sequencing the human genome (Sinsheimer, 1989) and planted the idea of a large-scale genome sequencing effort in the mind of Nobel Laureate and Harvard Professor of Biology, Walter Gilbert. Gilbert went on to co-chair a session on the Human Genome Project with Paul Berg at the Cold Spring Harbour (CSH) Symposium on the Molecular Biology of *Homo sapiens* in 1986. Meanwhile, at a meeting in Santa Fe, New Mexico, Charles DeLisi, of the United States Department of Energy (DOE), proposed a research programme concentrating on advanced DNA sequencing technology, computer analysis and methods to order DNA fragments cloned from the human genome (DeLisi, 1988). The interest behind the DOE initiative was based on the use of DNA analysis to detect mutations among atomic bomb survivors. DeLisi's research program, the DOE Human Genome Initiative, and the first US government programme on genome research, began in 1987. Independently of this, and at the same time, Renato Dulbecco, President of the Salk Institute, published a short article in *Science* arguing for sequencing of the human genome as an aid to cancer research (Dulbecco, 1986). In 1986 the Howard Hughes Medical Institute (HHMI) and the National Institutes of Health (NIH) also held meetings on the Human Genome Project and what role they may have to play in its future.

World-wide interest in the Human Genome Project was generated, and in Italy, a pilot project on genome research was funded by the Italian National Research Council (Dulbecco, 1987). In the United Kingdom, Sydney Brenner and Walter Bodmer sent a proposal for an EC genome project to the European community (EC) offices. The result was a joint genome programme between the Medical Research Council (MRC) and the

Imperial Cancer Research Fund (ICRF) which began in 1989 (Alwen, 1990), with the European Commission approving a 2-year human genome effort in 1990.

Back in the United States, a committee appointed by the National Research Council (NRC) of the National Academy of Science published a report that argued for a broader Human Genome Project with funding of $200 million/year for 15 years. A report by the Office of Technology Assessment (OTA) in the same year, commissioned by the House Committee on Energy and Commerce, agreed that genome research would go forward but questioned its management and co-ordination. Congress appointed funds to both the NIH and DOE for genome research in 1988 and the 1988 Memorandum of Understanding led to a co-operative effort between the agencies, culminating in the submission of a 5-year joint research proposal to congress (US Department of Health and Human Services and US Department of Energy, 1990).

Although the United States was the first country to make a financial commitment, the Human Genome Project is a world-wide initiative and in 1988, at the first CSH meeting on genome sequencing, the Human Genome Organisation (HUGO) was formed to mediate international scientific collaboration.

Initial five year goals were set for the Human Genome Project. In the areas of informatics and technology there were to be developments that would support the human genome project such as algorithms and analytical tools for the interpretation of genomic information. Programmes were to be developed to address the ethical, legal and social implications of the Human Genome Project. For the mapping and sequencing of the human genome, a fully connected genetic map was to be completed with markers spaced an average of 2-5 centimorgans (cM) apart, and that each marker was to be identified by a sequence-tagged site (STS). An STS is normally defined by 200-500 nucleotides that

are operationally unique in the human genome i.e. it can be specifically detected via a polymerase chain reaction (PCR) in the presence of all other genomic sequences, allowing the location of individual genes on chromosomes to be established. STS maps and overlapping sets of cloned DNA, with continuity over lengths of 2 million base pairs (bp), were also to be used to create a physical map of all human chromosomes. Markers would be spaced approximately 100,000 bp apart or be closely spaced and unambiguously ordered. Current methods of DNA sequencing were to be improved or new ones invented to bring the cost of large-scale sequencing down to $0.50/bp or less and the sequence of 10 million bp of human DNA determined. Parallel studies on selected model organisms were also included in the initiative as an aid to the understanding of the human genetic information obtained and to add to the development of existing sequencing strategies. Specifically, a genetic map of the mouse genome was to be prepared with the physical mapping of several chromosomes and approximately 20 million bp of sequence were to be aggregated from a variety of model organisms with the focus on stretches of 1 million bp.

The ultimate goal of the Human Genome Project is to produce detailed physical and genetic maps of each of the 24 different chromosomes and to determine the sequence of the 3 billion nucleotides that make up human DNA.

1.2 *Fugu* as a model genome.

The identification of genes in model organisms, and their correlation to the human genome, was one of the original goals of the Human Genome Project because the evolutionary conservation seen between humans and experimental organisms has led to

remarkable insights into biochemistry, regulation, development and disease (Tugendreich *et al.*, 1994). Even the simplest of eukaryotes contains many genes similar enough to those of humans that they can allow comparisons and functional analysis of gene products and those processes fundamental to all eukaryotic cells. The other major advantage of using model organisms, such as *Escherichia coli* (4.7 megabase (Mb) genome) and the yeast *Saccharomyces cerevisiae* (14 Mb genome), is that their high gene density and small genome size makes them suitable for the development of sequencing and mapping technologies and strategies (Jones, 1995). In comparison to these simple model organisms only 1-3% of mammalian genomic DNA contains coding sequence with a gene density in the human genome that averages only one gene for approximately every 50 kb (Jones, 1995). For these reasons most of the large-scale sequencing of the human genome has so far been confined to disease loci alone. This is with the exception of human chromosome 22 for which 33.4 megabase of sequence data containing at least 545 genes and 134 pseudogenes has been reported (Dunham *et al.*, 1999).

However, it was recognised that the usefulness of the more simple model organisms was limited due to the distinct morphology and development of vertebrates. For the study of mammalian gene structure, function and regulation the most relevant model would be a vertebrate such as the mouse, which indeed has served as an excellent model in elucidating developmental and disease processes relevant to humans. For the mouse genome and many of the other vertebrate models in existence, genome mapping and sequencing problems have proven to be as difficult as in humans because their genomes are also large and of comparable complexity to the human genome. Taking all this into consideration, it was Sydney Brenner who recognised that the ideal model for

comparative genome sequencing and analysis would be a vertebrate genome of minimal size and complexity but with maximum homology to the human genome. In 1993 Brenner *et al.* proposed the Japanese pufferfish *Fugu rubripes* (Figure 1.1) as a new compact model vertebrate genome (Brenner *et al.*, 1993).

Figure 1.1. The Japanese pufferfish *Fugu rubripes*.

The tetraodontoid fish *Fugu rubripes* (*Fugu*) is a pufferfish which inhabits the sea surrounding the Japanese islands. As one of the largest pufferfish, and a delicacy amongst the Japanese people, *Fugu* is the most commercially farmed species. As a teleost fish, *Fugu* has a separation time from mammals of around 430 million years making it the most distant extant vertebrate precursor (Powers, 1991). More importantly, it appears to fill the criteria of an ideal 'model' genome.

By 1968 Ralph Hinegardner had assayed the cellular DNA content of 200 different species of teleost fish, and extended the number of species analysed to 275 by 1972 (Hinegardner, 1968; Hinegardner and Rosen, 1972). Both studies confirmed that the amount of cellular DNA present in teleosts ranged from 0.4 to 4.4 picograms (pg) per cell and that specialised fishes had a lower cellular DNA content than the more generalised fishes. It was found that one family of teleost fishes, the tetraodontoid family, had the smallest known vertebrate genomes which ranged from 0.4 to 0.5 pg.

When compared to the 3 pg haploid DNA content of human cells, 0.4 pg equates to approximately 400 Mb. More specifically, the pufferfishes *Tetraodon fluviatilis* and *Tetraodon palembangensis* were found to have haploid DNA contents of 0.39 pg and 0.48 pg per cell respectively. The work of Pizon *et al.* (1984) supported this predicted genome size. This group used reassociation kinetics and density centrifugation to study the genome of another member of the tetraodontoid family, *Arothron diadematus*. With 87% unique (non-repetitive) DNA the *A. diadematus* genome was found to be 80 times more complex than the *E. coli* genome, this figure corresponds to 410 Mb and results in a total genome size of 470 Mb.

Brenner *et al.* used random sequencing and a gene probing approach to present a detailed analysis of the *Fugu* genome (Brenner e*t al.*, 1993). For random sequencing they sonicated *Fugu* genomic DNA to produce 596 clones and 127,831 bp of sequence. 44 clones were sequences almost entirely made up of RNA genes, minisatellite and microsatellite repeat sequences, the most abundant minisatellite repeat unit of 118 bp comprised 2% of the total genome. The remaining 552 clones and their nucleotide sequences were translated into amino acid sequences in all six possible reading frames. The amino acid sequences were searched for homology against the SWISSPROT protein database and fragments of 10 genes were identified. In total, 1,011 bp or 0.8% (1,011/127,831 x 100%) of the analysed nucleotide sequence was found to be coding sequence. To estimate the size of the *Fugu* genome from this figure, a non-redundant coding sequence database for the human genome was created. This database comprised a total of 2,576 genes corresponding to 3,096,858 nucleotides of coding sequence. This included a single example of each mammalian coding sequence in SWISSPROT that could be found, or be reasonably expected to be found, in the human genome. From this

it was concluded that a search of random sequence from the 3,000 Mb human genome against SWISSPROT would identify 0.1% (3,096,858/3,000 Mb X 100) as coding sequence.

As a vertebrate, *Fugu* shares with mammals many anatomical, physiological, immunological and neurological characteristics (Colbert and Morales, 1991), therefore it was assumed that *Fugu* and humans have approximately the same number of genes. By comparing the relative percentages of coding sequence Brenner and colleagues calculated that the *Fugu* genome was 7.68 times smaller than the human genome (0.791/0.103 = 7.68) and was therefore 380 Mb (3000 Mb/7.68). To support this figure they hybridised 3 independent non-amplified λ2001 phage libraries with single-copy-gene probes. In a total of 1.379×10^6 phage 56 copies of the single-copy genes were found, which equated to 1 complete genome per 24,625 phage. Assuming that all sequences were equally represented in the libraries, and with an average insert size of 16.4 kb, a *Fugu* genome size of 404 Mb (16.4 kb x 24,625) was predicted. With both estimates in accordance with the 380 Mb previously found for the *Tetradon* species, and the close correspondence of both of the estimates, this group concluded that the assumption that *Fugu* has a similar gene repertoire to human was correct. Thus, the predicted genome size of *Fugu* was confirmed as approximately 400 Mb making it 7.5 times smaller than the human genome. In addition, the genome was found to have a G+C content of 44.2%, comparable to mammalian genomic DNA, with more than 90% of the genome being non-repetitive, unique sequence (Brenner *et al.*, 1993).

Brenner and colleagues proposed that the compactness of the *Fugu* genome was due to the apparent reduction of intron size with three quarters of introns being small, (60 - 120 bp) giving a modal value of ~80 bp. They also found a relative paucity of repetitive

DNA and absence of dispersed repeats in the *Fugu* genome. In addition, they reported that the expected high gene density had been confirmed by initial sequencing of contiguous stretches of DNA (Brenner *et al.*, 1993).

As a general rule it has been found that *Fugu* genes are smaller than their human counterparts although the *Fugu* mRNA transcripts are of approximately the same size. Intron/exon boundaries are generally conserved between species and the reduction in gene size has been shown to be the result of reduced intron size. Smaller intergenic distances also help to account for the overall reduction in the size of the *Fugu* genome. Based on the prediction of 60 - 70,000 genes in the human genome, and the assumption that *Fugu* has the same number of genes in only 400 Mb, the *Fugu* genome would average one gene every 6-7 kb. In common with other tetraodontiformes *Fugu* is predicted to have 22 to 24 chromosomes (Elgar *et al.*, 1996), and as seen in mammalian genomes, there is a codon bias for G or C in the third position.

One of the first genes to be isolated from *Fugu* was the housekeeping gene for glucose-6-phosphate dehydrogenase (G6PD) which catalyses the reaction for NADPH production in erythrocytes (Mason *et al.*, 1995). It was found that compared to its human counterpart, the *Fugu* G6PD gene was 3.75 times smaller due to a reduction in intron size but that the intron/exon structure was identical to the human gene. Over the coding region the identity seen between the two species was 71% for the nucleotide sequence and 76% for the amino acid sequence.

Many different genes have since been isolated from *Fugu* including, for example, those coding for a variety of transcription factors (Schüddekopf *et al.*, 1996; Verma-Kurvari & Johnson, 1997; Göttgens *et al.*, 1998; Wentworth *et al.*, 1999; Tassone *et al.*, 1999), G protein-coupled receptors (Yamaguchi *et al.*, 1996; Yamaguchi & Brenner, 1997; Naito

et al., 1998), a homologue of complement component C9 gene (Yeo *et al.*, 1997) and T-cell receptor gene homologues (Rast *et al.*, 1995). These findings support the argument that *Fugu* has not only a similar number of genes as mammals but a comparable gene repertoire too.

Many studies utilising *Fugu* as a model genome have made use of the 430 million years of evolutionary distance between *Fugu* and humans. There is maximum stringency for genome sequence comparisons as all unconstrained sequences have had the maximum time to randomise by mutation. Only those essential sequences, coding or non-coding, that are required for functions common to all vertebrates will be conserved. For example, in 1995 the *Fugu* Huntington's disease gene (*Fr*HD) was cloned (Baxendale *et al.*, 1995) and showed an exact reduction in size of 7.5 times, from 170 kb to 23 kb, with conservation of all 67 of the exons found in the human HD gene. Again there was conservation of intron/exon boundaries and good nucleotide and amino acid identity between the two species (69% and 73% respectively) which led to the identification of potential regulatory regions. In addition to identifying conserved regions within the coding sequence of genes, *Fugu* has also been used to identify conserved regulatory elements. This approach was used to identify conserved enhancer elements present between mouse and *Fugu* for both the *Hoxb-1* and *Hoxb-4* genes (Marshall *et al.*, 1994; Aparicio *et al.*, 1995). In these studies murine *Hoxb-1* and *Hoxb-4* genes were found to be functionally regulated by *Fugu* enhancer elements. However, further work on *Fr*HD called into question the use of *Fugu* sequences in transgeneic studies when it was found that *Fr*HD mRNA was not processed correctly in mouse cells or in transgenic mice (Sathasivam *et al.*, 1997).

The area of conserved gene order is the most contentious issue surrounding the use of *Fugu* as a model genome. Positional cloning, the use of conserved synteny and conserved linkage, is a particularly attractive technique for isolating new genes flanking either side of candidate regions. Also, regions of conserved synteny would accelerate the mapping and ordering of genes.

Conserved synteny, where two or more genes are located on homologous chromosomes in different species regardless of their order on either chromosome (Koop and Nadeau, 1996), has now been demonstrated across many regions of the *Fugu* genome. This makes *Fugu* an attractive model for the positional cloning of new candidate disease genes and is particularly aided by the smaller intronic and intergenic regions present in the *Fugu* genome. The sequence information obtained could then be used to isolate the corresponding human gene of interest. This approach has been taken by several groups who initiated the mapping of genes on the *Fugu* genome where mapping of human chromosomes has proven difficult and the identity of genes and their order remains unknown.

Human chromosome 4q35 is a very repetitive region of the human genome and contains many pseudogenes and low copy repeats making transcriptional mapping of this area very problematic. This region contains the locus for the neuromuscular disorder Facioscapulohumeral Muscular Dystrophy (FSHD) and it was suggested that the development of a synteny map of the FSHD region could be used to accelerate the identification of FSHD candidate genes (Grewal *et al.*, 1998). Likewise, Ali *et al.* (1997) have initiated mapping of a region of the *Fugu* genome, homologous to human chromosome 9, as an aid to the physical mapping of human chromosome 9q34.

However, in the *Fugu* genome the case for conserved linkage and gene order, where two or more genes are located in the same order on homologous chromosomes, is not so clear cut. Initial work by several groups demonstrated linkage of groups of two or three *Fugu* genes (How *et al.*, 1996; Sandford *et al.*, 1996; Schofield *et al.*, 1997; Trower *et al.*, 1996). For example, the work of Trower and colleagues has demonstrated that the order of the FOS/S31iii125/S20i15 genes in the *Fugu* genome was identical to the order of the homologous genes in the human genome. In humans the cFOS, S3iii125 and S20i25 genes are located in the AD3 locus of chromosome 14q24.3, a region spanning over 600 kb and associated with familial Alzheimer's disease. In *Fugu*, the three genes are found in a region of only 12.4 kb. However, these studies have covered relatively small distances in terms of human DNA and do not provide complete sequence data for the regions. In addition, they compared genes whose order in the human genome was already known.

More recently, extensive work over larger regions of the *Fugu* genome has been performed with conflicting results. A *Fugu* genomic region of less than 100 kb, corresponding to a 1.5 Mb region of human chromosome 11, specifically 11p13, was isolated in 3 overlapping cosmids (Miles *et al.*, 1998). In humans, this region contains two well studied disease genes, PAX6 and WT1, which are essential for normal mammalian development, and also the reticulocalbin (RCN1) gene which encodes a calcium binding protein of the endoplasmic reticulum. Heterozygous deletion of this region gives rise to Wilm's tumour, aniridia, genitourinary abnormalities and mental retardation, collectively known as WAGR syndrome. In humans, the order of these three genes is WT1/RCN1/PAX6 with the RCN1 gene being transcribed in the opposite direction to the WT1 and PAX6 genes, this is also the case for *Fugu*, and so although

these three genes are not members of the same gene family their order and orientation is identical between *Fugu* and human.

Additionally, Sequencing of 148 kb of *Fugu* genomic DNA around the wnt1 locus identified fifteen genes (Gellner and Brenner, 1999). Nine of the genes identified were already known genes with eight mapped to human chromosome 12q13. Two of the genes, wnt1 and wnt10b, are known to be tightly linked within 2.3 cM on murine chromosome 15.

Neurofibromatosis type 1 (NF1) patients with deletions of over 500 kb generally suffer from a more severe manifestation of neurofibromatosis, accompanied by atypical anomalies including mental retardation. This is thought to be a contiguous gene syndrome. In *Fugu*, three genes were identified 3′ of the NF1 gene in the following order, A-kinase anchor protein (AKAP84), an unidentified gene designated as BAW (between AKAP84 and WSB1) and WD-40-repeat protein (WSB1) (Kehrer-Sawatzki *et al.*, 1999). However, in humans, whilst these genes have all been localised to the same region of chromosome 17 they have been found to map in a different order (WSB1, AKAP84, BAW). In addition, further information available on the location of the retinoic acid receptor α (RARA) gene distal to the CC chemokine gene cluster in 17q11.2-q12, has ordered the genes as follows: NF1- CC chemokines- WSB1- RARA- AKAP84/BAW or alternatively WSB1- NF1- CC chemokines- RARA- AKAP84/BAW. Whichever order is correct the difference to the order in *Fugu* is apparent. Thus, whilst synteny and linkage of the genes in this region is conserved gene order is not.

Work performed on the Surfeit gene cluster also questions the conservation of gene order in *Fugu*. Initial work by Armes *et al.* (1997) showed that the structure and predicted products of the individual *Fugu* Surfeit genes were highly homologous to their

human counterparts. However, when compared to the single tightly clustered locus found in mammals and birds, *Fugu* surfeit gene homologues were present at 3 separate loci in the *Fugu* genome. Further work by Gilley and Fried (1999) questioned the utility of *Fugu* to facilitate the identification of human disease genes by comparative positional cloning. This group expanded their study to seven *Fugu* genes closely linked to the Surfeit genes in two regions of the *Fugu* genome. They mapped and ordered the human homologues of these genes and found that all seven genes mapped to a 3 Mb region of human chromosome 9q34.1, approximately 2-4 Mb proximal to the Surfeit genes. However, although this demonstrates conserved synteny, the gene order differs greatly suggesting intrachromosomal rearrangements have occurred during the 900 million years of divergent evolution that separates *Fugu* and humans.

Additional evidence questioning the use of *Fugu* as a good model for positional cloning analysis came from the laboratory of Reboul *et al.* in 1999. This group analysed the interferon/interleukin-10 receptor gene cluster which is present on human chromosome 21 and linked to the GART gene. They showed that whilst the large-scale structure of the GART-cytokine receptor gene cluster locus was conserved in human and chicken it was not in *Fugu*. They identified a *Fugu* GART homologue but not the cytokine receptor gene cluster. Instead, they identified a novel predicted gene 5′ of the GART gene, which had no homology to the SON3 gene present 5′ of GART in the human genome. Neither did this gene have homology to the *Fugu* homologue of the yeast YDR140w gene 3' of the GART gene demonstrating that the gene had not been inverted. Also present on the *Fugu* cosmid containing the GART gene were two copies of the putative neurotransmitter receptor gene (PNR) and the VEGF gene both of which

have been mapped to human chromosomes 6q23 and 4q34 respectively. Thus, in contrast to *Fugu* none of these genes are physically linked in the human genome.

Nevertheless, *Fugu* has become established as a model vertebrate genome with few amendments to its initial characterisation. Contrary to previous reports, repetitious dispersed elements have been found in the *Fugu* genome, both long terminal repeat (LTR) and non-LTR retrotransposons named sushi and Maui respectively (Poulter and Butler, 1998; Poutler *et al.*, 1999), and although they are not evenly dispersed throughout the genome there may be as many as 3000 copies of Maui present (Poulter *et al.*, 1999). Additionally, a comprehensive analysis of 11.338 Mb of *Fugu* genomic DNA detailed the characterisation of 6042 microsatellite repeat sequences (Edwards *et al.*, 1998). This study demonstrated that whilst the sparseness of the *Fugu* genome may be attributed in part to a lack of minisatellite repeat sequences and transposable elements, as compared to mammalian genomes, it is not due to a lack of microsatellite repeat sequences.

Another pufferfish, *Tetraodon fluviatilis* (*T. fluviatilis*), has been introduced as a model for genome studies to complement the use of *Fugu rubripes* (Crnogorac-Jurcevic *et al.*, 1997). Unlike *Fugu rubripes*, *T. fluviatilis* is a freshwater pufferfish that can easily be maintained and bred in an aquarium. Comparison of mitochondrial cytochrome b sequences from the two species indicated a separation time of not more than 18-30 million years. *T. fluviatilis* has since been used to characterise the genomic structure and promoter region of a *c-fos* homologue (Chang *et al.*, 1997), to characterise and orient *snf5* and *set* gene homologues (Yao *et al.*, 1998), and to analyse the spermine synthase gene homologue in comparison to the same region in *Fugu rubripes* and the zebrafish *Danio rerio* (Boeddrich *et al.*, 1999).

To summarise, at ~400 Mb the *Fugu* genome is compact and between seven and eight times smaller than the human genome. As a vertebrate, it has a gene repertoire similar to that of humans. *Fugu* genes are reduced in size compared to their human counterparts due to a lack of intronic DNA. Generally, there is good conservation of gene sequences and organisation between the *Fugu* and human genes compared to date. *Fugu* transcripts are of a comparable, if not identical, length to their human counterparts. For these reasons, it was decided that the model vertebrate genome of *Fugu rubripes* should be searched for homologues of the nitric oxide synthase (NOS) gene family.

1.3 The discovery of nitric oxide.

In 1980 Furchgott and Zawadski reported that the acetylcholine-dependent relaxation of arterial smooth muscle only occurred in the presence of intact endothelium. They demonstrated that the relaxation was mediated by a labile factor which diffused from the endothelium to the adjacent smooth muscle. This factor became known as endothelium-derived relaxing factor (EDRF) (Furchgott and Zawadski, 1980). An increase in the levels of cyclic 3′,5′-guanosine monophosphate (cGMP) was found to accompany the endothelium-dependent relaxation of blood vessels (Rapoport, 1983). It was already known that cGMP biosynthesis was catalysed by the guanylate cyclase enzymes (Hardman and Sutherland, 1969; Ishikawa *et al.*, 1969; Schultz *et al.*, 1969; White and Aurbach, 1969) and it had previously been shown that the simple gas nitric oxide (NO) caused the activation of guanylate cyclase and therefore an increase in cGMP levels (Arnold *et al.*, 1977). However, the connection between the two observations was not made at this time.

Whilst Furchgott and Zawadski were investigating EDRF, investigations into the metabolism of inorganic nitrogen oxides in humans showed that there was endogenous synthesis of nitrate (NO_3^-) (Green *et al.*, 1981b). Experiments in germ-free animals had shown that the NO_3^- produced was mammalian in origin and not the result of intestinal microbial metabolism as first thought (Green *et al.*, 1981a). Four years later it was shown that exposure to bacterial lipopolysaccharide (LPS), increased NO_3^- levels in the blood and urine of LPS-sensitive mice (Stuehr and Marletta, 1985). In the same study it was shown that *in vitro* immunological stimulation increased the production of nitrite (NO_2^-) and NO_3^- by activated mouse peritoneal macrophages. It was concluded that the murine macrophages were therefore the most likely source of LPS-induced NO_2^- and NO_3^- synthesis.

The vasodilatory effects of EDRF were subsequently shown to be mediated through the activation of soluble guanylate cyclase (sGC) (Ignarro *et al.*, 1986; Förstermann *et al.*, 1986; Mulsch *et al.*, 1987). The chemical identity of EDRF as NO was established in 1987 when the two moieties were shown to have identical biological properties. In addition, NO was demonstrated to be generated by vascular endothelial cells in an amount which could account for the biological activity of EDRF (Palmer *et al.*, 1987; Khan and Furchgott, 1987; Ignarro *et al.*, 1987). It was then demonstrated, by three independent groups, that NO was the originator of the NO_2^- and NO_3^- produced by activated macrophages (Marletta *et al.*, 1988; Hibbs *et al.*, 1988; Stuehr *et al.*, 1989).

Other studies revealed the presence of NO in the central nervous system (CNS). For example, activation of *N*-methyl-*D*-aspartic acid (NMDA) receptors by glutamate was shown to increase cGMP synthesis in neighbouring cells (Garthwaite *et al.*, 1988). This rise in cGMP led to the identification of a factor also able to relax vascular smooth

muscle. Further work demonstrated that this factor had properties that were indistinguishable from those attributed to NO.

1.4 The generation of NO.

L-arginine was identified as the substrate for NO synthesis in the vascular endothelium (Palmer *et al.*, 1988a; Sakuma *et al.*, 1988; Schmidt *et al.*, 1988) and the L-arginine analogue, N^G-monomethyl-L-arginine (L-NMMA) as an inhibitor of NO biosynthesis in this tissue (Palmer *et al.*, 1988b; Rees *et al.*, 1989b). L-NMMA had previously been shown to inhibit the L-arginine dependent generation of nitrite and nitrate as well as the production of citrulline from macrophages (Hibbs *et al.*, 1987). Further experiments revealed that citrulline and NO were co-products of the same enzymatic reaction (Palmer and Moncada, 1989; Moncada and Palmer, 1990). It is now known that NO is formed by an NADPH-dependent five-electron oxidation of L-arginine to form L-citrulline and NO (Bredt and Snyder, 1994).

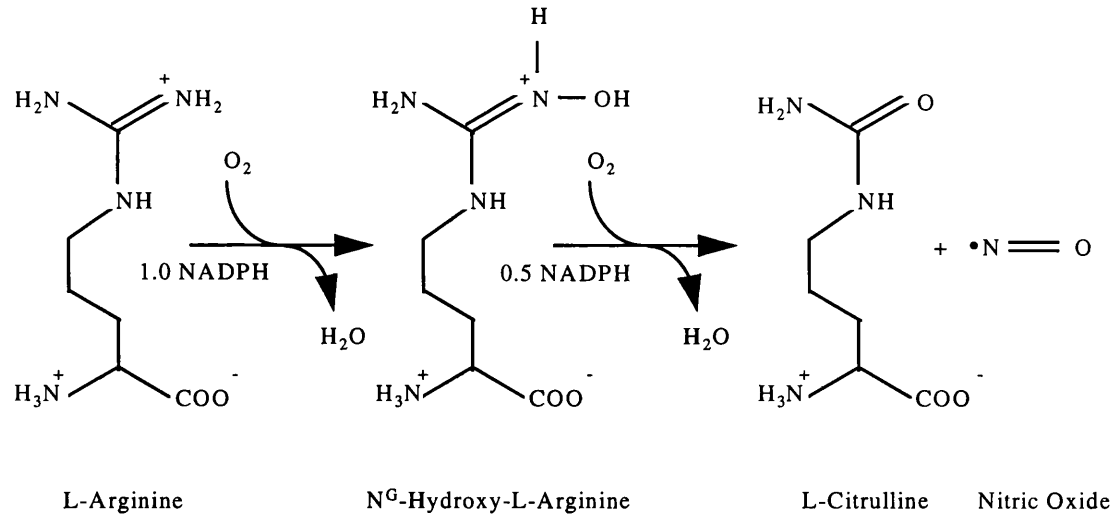


Figure 1.1. The generation of nitric oxide.

The synthesis of NO from L-arginine is catalysed by the family of NO synthase (NOS) enzymes (EC 1.14.13.39). In mammals, there are at least three distinct NOS isoforms, neuronal (nNOS; NOS1), inducible (iNOS; NOS2) and endothelial (eNOS; NOS3), named after the tissues in which they were first identified (neurons; cytokine-induced immune cells; vascular endothelium). All three have been cloned from a variety of tissues and species as summarised in table 1.1 and the overall structure of each isoform is illustrated in figure 1.1a.

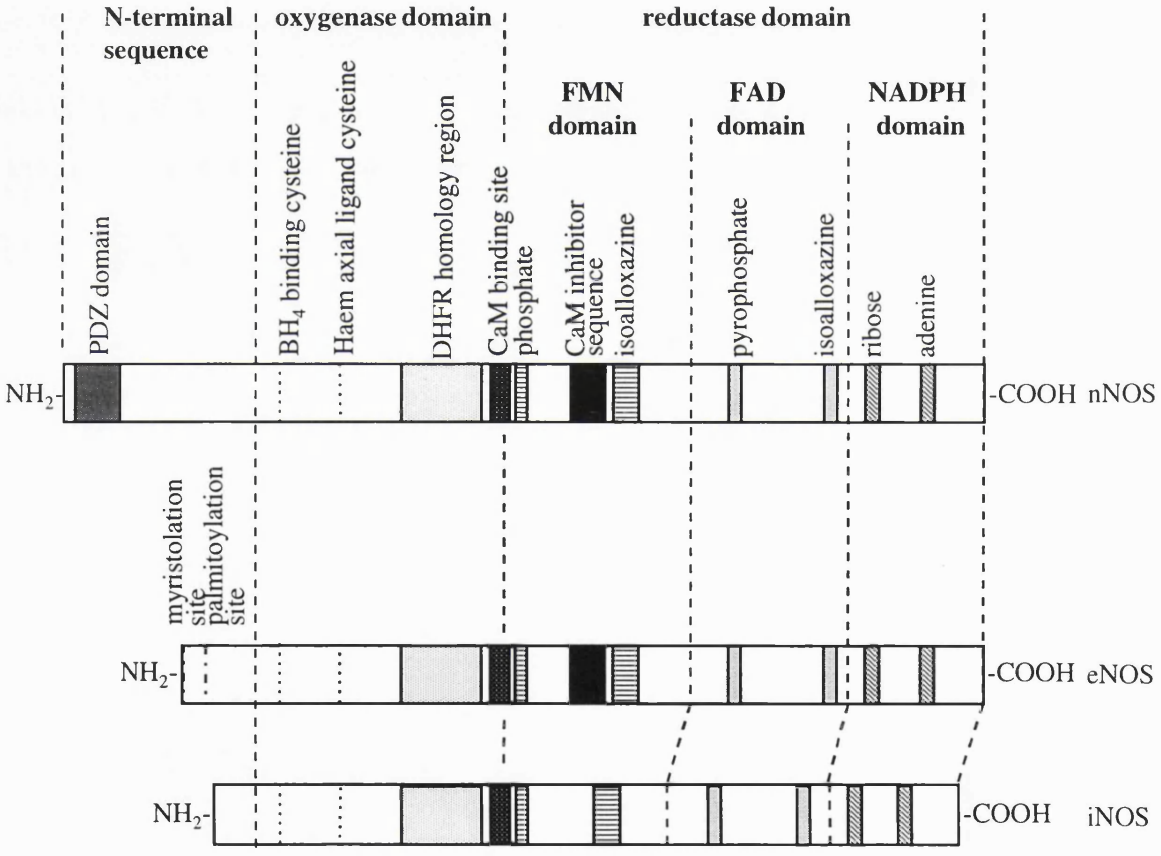


Figure 1.1a. Domain structure of NOS isoenzymes. (adapted from Hemmens & Mayer, 1999. From: *Methods in Molecular Biology, Vol. 100. Nitric Oxide Protocols.* Edited by Michael A. Titheradge).

facing P38

Table 1.1. Cloned NOS isoforms.

Isoform	Cell/Tissue	Species	Reference
NOS1/ nNOS	Brain	Rat	Bredt *et al.*, 1991; Charles *et al.*, 1993; Black and de Montellano, 1995
	Brain	Mouse	Ogura *et al.*, 1993
	Brain	Human	Nakane *et al.*, 1993;
	Placenta		Hall *et al.*, 1994
	Brain	Rabbit	Jeong and Kim, 1997
	genomic	*Drosophila melangoster*	Regulski and Tully, 1995
	Salivary gland	*Rhodnius prolixus*	Yuda *et al.*, 1996
		Lymnaea stagnalis	Korneev *et al.*, 1998
	genomic	*Anopheles stephensi*	Luckhart *et al.*, 1998; Luckhart and Rosenberg, 1999
	genomic	*Xenopus laevis*	Scheinker *et al.*, 1999
NOS2/ iNOS	Macrophage	Mouse	Lowenstein *et al.*, 1992; Lyons *et al.*, 1992; Xie *et al.*, 1992
	MTAL cell line	Mouse	Kone *et al.*, 1995
	Liver	Rat	Adachi *et al.*, 1993
	Astroglia	Rat	Galea *et al.*, 1994
	Vascular smooth muscle	Rat	Geng *et al.*, 1994
	Neutrophils	Rat	Kosuga *et al.*, 1994
	Islets of Langerhans	Rat	Karlsen *et al.*, 1995
	Hepatocytes	Human	Geller *et al.*, 1993
	DLD1 colorectal carcinoma	Human	Sherman *et al.*, 1993
	Articular chondrocyte	Human	Charles *et al.*, 1993; Maier *et al.*, 1994
	Foreskin fibroblast HSF42	Human	Chartrain *et al.*, 1994
	Macrophage	Bovine	Adler *et al.*, 1995
		Guinea pig	Uchida *et al.*, 1999
	Macrophage	Chicken	Lin *et al.*, 1996
		Cyprinus carpio	Saeij *et al.*, 1999
NOS3/ eNOS	Aortic endothelium	Bovine	Sessa *et al.*, 1992; Nishida *et al.*, 1992; Lamas *et al.*, 1992
		Mouse	Gnanapandithen *et al.*, 1996
		Pig	Zhang *et al.*, 1996
	Endothelium	Human	Janssens *et al.*, 1992; Marsden *et al.*, 1992

For the structural and functional characterisation of the NOS family of enzymes high levels of protein are required. Expression of recombinant NOS isoforms has been achieved in a number of systems as summarised in table 1.2.

Table 1.2. Expression of Recombinant NOS.

Isoform	Species	Expression system	Reference
nNOS	Rat	293 kidney cells	Bredt *et al.*, 1991; McMillan *et al.*, 1992
	Rat	Baculovirus/insect cell	Richards and Marletta, 1994; Charles *et al.*, 1993; Harteneck *et al.*, 1994; Riveros-Moreno *et al.*, 1995
	Rat	*E.coli*	McMillan and Masters, 1995 Roman *et al.*, 1995; Gerber and de Montellano, 1995
	Rat	Yeast	Black and de Montellano, 1995
	Human	COS cells	Nakane *et al.*, 1993
	Human	Baculovirus/insect cell	Nakane *et al.*, 1995
	Drosophila	Baculovirus/insect cell	Regulski and Tully, 1995
	Rhodnius prolixus	Baculovirus/insect cell	Yuda *et al.*, 1996
iNOS	Rat	293 kidney cells	Karlsen *et al.*, 1995
	Rat	COS cells	Adachi *et al.*, 1993
	Mouse	Baculovirus/insect cell	Moss *et al.*, 1995
	Mouse	E.coli	Wu *et al.*, 1996
	Mouse	Yeast	Sari *et al.*, 1996
	Human	CHO cells	Charles *et al.*, 1993;
	Human	293 kidney cells	Geller *et al.*, 1993
	Human	Baculovirus/insect cell	Nakane *et al.*, 1995
eNOS	Bovine	COS cells	Sessa *et al.*, 1992; Nishida *et al.*, 1992; Lamas *et al.*, 1992
	Bovine	Baculovirus/insect cell	List *et al.*, 1996; Busconi and Michel, 1995
	Bovine	*E.coli*	Marteseck *et al.*, 1996
	Human	3T3 cells	Janssens *et al.*, 1992;
	Human	Baculovirus/insect cell	Chen *et al.*, 1995; Nakane *et al.*, 1995; Seo *et al.*, 1995; Chen *et al.*, 1996
	Human	*E.coli*	Rodríguez-Crespo *et al.*, 1996

In 1991, the first complete primary structure of a NOS enzyme was deduced by Snyder and colleagues when they cloned, sequenced and expressed a rat brain NOS cDNA (Bredt *et al.*, 1991). In the C-terminal half of the enzyme they identified a number of recognition sites needed for enzyme function that were already familiar from other enzymes. The structural units they identified included a basic amphipathic α helix calmodulin-binding consensus sequence, which was consistent with the requirement of the enzyme for calmodulin as a cofactor (Bredt and Snyder, 1990), and a cAMP-dependent protein kinase phosphorylation consensus sequence. In addition, they found a NADPH binding domain, with well-defined points of contact for the ribose and adenine rings. Furthermore, consensus binding domains for the co-factors flavin mononucleotide (FMN) and flavin-adenine dinucleotide (FAD), including distinct binding sites for the pyrophosphate and isoalloxazine moieties of FAD, were identified. Haem and (6R)-5,6,7,8-tetrahydro-L-biopterin (tetrahydrobiopterin; BH_4) were also identified as co-factors required for the synthesis of NO (Schmidt and Murad, 1991).

Analysis of the predicted amino acid sequence with the FASTA programme (Pearson and Lipman, 1988) revealed very little identity of the N-terminal half of the enzyme to known proteins. Some homology of the C-terminal half of NOS was found to rat cytochrome P-450 reductase, with 36% identity and 58% similarity over 641 amino acids. The normal function of cytochrome P-450 reductase is to provide cytochrome P-450s with reducing equivalents required for their activity. With a P-450-type haem moiety present in NOS, the same reaction was proposed for this domain (Marletta, 1993). In addition, the N-terminal fragment of both inducible and constitutive NOS isoforms have been shown to have a maximum absorption at ~450nm (λ_{max}) upon reduction and treatment with carbon monoxide. This spectral characteristic is relatively rare and limited to the

cytochrome P-450 enzyme family (Marletta, 1994). The cytochrome P-450 supergene family codes for enzymes which carry out the oxidative metabolism of endogenous and xenobiotic compounds (Nelson *et al.*, 1993). All except one member of the family require a separate flavoprotein reductase (normally P-450 reductase) and sometimes an iron-sulphur protein to transfer electrons into the haem prosthetic group responsible for the oxidative catalysis. The exception is the fatty acid monooxygenase $P\text{-}450_{BM3}$ isolated from *Bacillus megaterium*, whereby a single polypeptide encompasses both the flavoprotein reductase and the haem moiety (Narhi and Fulco, 1986). However, although superficially NOS has the appearance of a self-sufficient mammalian P-450 (i.e. an enzyme containing both reductase and haem domains) there is no significant homology of the NOS N-terminal sequence with the cytochrome P-450s (Marletta, 1994).

Proteolysis of rat nNOS further supported the idea that the NOS enzymes had a bi-domain structure (Lowe *et al.*, 1996). This study demonstrated the presence of a central calmodulin-binding sequence flanked by an arginine/haem-binding N-terminal fragment and a C-terminal fragment that catalysed the NADPH-dependent reduction of cytochrome c.

A more detailed knowledge of the NOS family of enzymes has since been elucidated by site directed mutagenesis, expression and crystallisation studies, in conjunction with the nucleotide sequence information available. Each NOS isoform consists of a reductase domain at the COOH terminus and an oxidative domain at the NH_2 terminus. Both within these domains, and the common consensus binding sites present for haem, calmodulin, FMN, FAD and NADPH (Bredt *et al.*, 1991; Vorherr *et al.*, 1993; Zhang

and Vogel, 1994; Venema *et al.*, 1997), residues crucial for enzyme function have been identified.

A cysteine residue (nNOS C415; iNOS C194; eNOS C184) conserved in all known NOS sequences was found to be crucial for the binding of haem as mutation of this residue resulted in loss of haem binding (Richards and Marletta, 1994; Chen *et al.*, 1994; Sari *et al.*, 1996). A single cysteine residue (C99) was identified in human eNOS as being largely responsible for BH_4 binding since mutation of this residue to alanine resulted in a large loss of BH_4 binding as measured by enzyme activity (Chen *et al.*, 1995). This was confirmed by mutation of the equivalent cysteine residue in rat nNOS (C331) to alanine (Martásek *et al.*, 1998).

More recent work on the crystal structure of bovine eNOS has shown that C101, the equivalent of human eNOS C99, is part of a zinc binding motif (C - $(X)_4$ - C) which spans residues 96-101 (Raman *et al.* 1998). The zinc binding motif has conserved cysteine residues that are involved in the co-ordination of a zinc atom which functions both to stabilise the BH_4-binding site and facilitate recognition of the pterin. Miller and colleagues (1999) have since shown that the zinc present in NOS does indeed have a structural, rather than catalytic role, which is important for the maintenance of an optimally functional and enzymatically active constitutuve NOS. The more detailed anatomy of substrate binding in eNOS has assigned guanidinium binding to residues E363 and W358, amino acid carboxylate binding to N368, Y359 and Q249 and α-amino (haem propionate (pyrole A)) binding to E363 with the hydrophobic β-sheet roof of the haem pocket formed by V338 and F355. Arginine forms hydrogen bonds with W358, E363, Q249 and N368, it also forms weaker bonds with V338, Y359 and F355.

In rat nNOS residues 564-715 were found to have homology to the dihydrofolate reductase (DHFR) homology fragment (Nishimura *et al.*, 1995). This region has been identified as the L-arginine binding domain as it was found to bind the L-arginine analogue N^G-nitro-L-arginine (NNA).

For the constitutive nNOS and eNOS isoforms, but not the iNOS isoform, a calmodulin autoinhibitory sequence has also been identified (Salerno *et al.*, 1997). This 45 amino acid insert is present between the regions responsible for FMN binding and using the x-ray crystal structure available for flavodoxins, the insert was positioned in three dimensional terms relative to the calmodulin binding site. It is thought that this element docks with a site on the constitutive NOS isoforms in such a way as to impede calmodulin binding and therefore enzymatic activation.

The NOS enzymes function as dimeric proteins (Abu-Soud and Stuehr, 1993; Baek *et al.*, 1993) and mutations affecting the ability of dimers to form have also been identified. In iNOS mutation of E411, P452 and P461 has been found to affect enzyme activity and also dimerisation (Gachhui *et al.*, 1997), whereas mutation of G450 and A453 affect BH_4 binding (Cho *et al.*, 1995). Two studies have localised the region required for dimerisation in murine iNOS to between residues 65 and 114 (Crane *et al.*, 1997; Crane *et al.*, 1998). It was demonstrated that in contrast to the crystal structure of a murine iNOS truncated by 114 N-terminal amino acids (Δ114 murine NOS2) which was a monomer, crystals of a murine iNOS truncated by only 65 N-terminal amino acids (Δ65 murine iNOS) were dimers. Similarly, for bovine eNOS N-terminal deletion mutations of 91 and 105 amino acids (Δ91 and Δ105) and showed increased monomerisation over wild type eNOS whilst Δ52 deletion mutants were normal (Rodriguez-Crespo *et al.*,

1997). Monomerisation has also been demonstrated for rat nNOS where residue H431 had been mutated (Perry *et al.*, 1998).

Crane *et al.* (1997) also made available the first 3-dimensional structural data for the NOS isoforms. They crystallised the oxygenase domain (residues 114 to 498) of murine iNOS ($iNOS_{ox}$) and found that these monomers consisted of a single-domain α-β fold which they described as a left-handed baseball catcher's mitt. The 'palm' of the mitt was formed by a large central winged β-sheet whereas the 'fingers' were formed by α-helices and the haem was held in place in the pocket of the glove by the 'thumb'. Using a similarity search program it was found that the polypeptide topology of $iNOS_{ox}$ was not shared by any other known protein structures although the high sequence conservation between the three NOS isoforms indicated that eNOS and nNOS would share the unusual topology. In 1998, further work by Crane and colleagues on the crystals of pterin-loaded and dimeric murine $iNOS_{ox}$ (residues 66-498 expressed with BH_4 + L-arginine/BH_4 + L-thiocitrulline/BH_4 + H_2O) revealed drastic changes to the interface of each NOS subunit with mobile and exposed hydrophobic regions, previously identified in iNOSox Δ114, that refolded to buttress the substrate-binding channel and sequester two molecules of BH_4 within two symmetry-related helical loops at the heart of the extensive dimer interface. However, each subunit maintained the monomer's unusual winged β-sheet that resembled the baseball catcher's mitt. Most recently work done on the haem domain of eNOS (Raman *et al.*, 1998) has confirmed that it also belongs to the α/β protein class with an overall fold similar to that reported by Crane *et al.* In addition, it was this group that observed the presence of the zinc ion, that was tetrahedrally coordinated to pairs of symmetry-related Cys residues, as previously discussed.

In addition to the consensus binding sites present in all three NOS isoforms, nNOS also has a PDZ domain (Cho *et al.*, 1992; Hendricks, 1995). PDZ domains were named after the three proteins they were first identified in; postsynaptic density-95 protein (PSD-95), Drosophila lethal(1)discs-large-1 tumour suppresser protein (Dlg) and the tight junction protein Zo-1. The deduced protein sequence of PSD-95 was found to be highly similar to Dlg (Woods and Bryant, 1991) and similar to p55, a protein associated with human erythrocyte membrane cytoskeleton (Ruff *et al.*, 1991; Bryant and Woods, 1992). All three proteins are associated with the cytoskeleton, contain a carboxy-terminal domain homologous to yeast guanylate kinase and a SH3 domain. The N-terminal domain of PSD-95 was found to contain three PDZ domains. These domains are internal repeat sequences of approximately 90 amino acids and were originally named GLGF repeats after the most conserved amino acid series (Gly-Leu-Gly-Phe) or alternatively, DHR regions after the Discs-large Homologous Region (Cho *et al.*, 1992; Woods and Bryant, 1993). Dlg was also found to have three PDZ domains whereas p55 has only one (Cho *et al.*, 1992). This group also noted that the then recently characterised enzyme NOS contained a single GLGF motif in its extended amino-terminal domain. Hendricks correctly identified that the PDZ domain was specific to nNOS isoforms and proposed that it may be responsible for the direct or indirect association of nNOS with membranes to allow the production of NO close to its potential targets (Hendricks, 1995).

PDZ domains are now known to be common protein-protein interaction domains that help to organise signalling cascades by recognising and binding to carboxy-terminal motifs or internal motifs and by forming head-to-tail oligomers with other PDZ domains

(Hillier *et al.*, 1999). Proteins that have PDZ domains are implicated in ion-channel and receptor clustering, and the linking of receptors and their effector enzymes (Ponting and Phillips, 1995).

With PDZ-peptide interactions, the PDZ domains of membrane-associated guanylate kinases (MAGUKs), one of which is PSD-95, have been shown to recognise a carboxy-terminal E(T/S)XV motif. In comparison, the nNOS PDZ domain has been shown to have a preference for a carboxy-terminal G(D/E)XV motif (Schepens *et al.*, 1997).

A number of proteins have been identified that interact with nNOS via direct PDZ-PDZ interactions. In neurons, nNOS forms heterodimers with PSD-95 through the direct interaction of their PDZ domains (Brenman *et al.*, 1996) and this is thought to couple nNOS to NMDA receptors and therefore NO production to NMDA receptor activation (Dawson *et al.*, 1993; Huang *et al.*, 1993). In muscle cells, nNOS heterodimerises with syntrophin via their PDZ domains to localise nNOS to the dystrophin complex (Brenman *et al.*, 1995), coupling NO production to muscle contraction. In both cases, the tertiary structure of both domains is important and the interaction is not dependent on the recognition of a carboxy-terminal motif. Hillier and colleagues (1999) showed that the structure of the nNOS PDZ domain comprises a receptor face with a peptide binding groove, encompassing residues 1 to 100, and a ligand face that forms a β-hairpin 'finger' and encompasses residues 100-130. They demonstrated that the nNOS PDZ β 'finger' acts as a PDZ ligand for the syntrophin PDZ domain and docks into the peptide binding groove. This also allows the receptor face of the nNOS PDZ domain to be open to interactions with additional proteins. The residues responsible for the primary and tertiary interactions, between the the syntrophin PDZ domain and nNOS, are uniquely

conserved in syntrophin and the second PDZ domain of PSD-95. This may explain the ability of these proteins to form heterodimers with nNOS.

The N-terminal sequence of eNOS has been shown to be responsible for the membrane localisation of this isoform. The mechanism for this is via the myristolation of a glycine residue at position 2 (Sessa *et al.*, 1993; Busconi and Michel, 1993). In addition, it has been found that palmitic acid is attached to cysteine residues at position 15 and 26 (Liu *et al.*, 1995; Garcia-Cardena *et al.*, 1996). Enzymes that are both myristolated and palmitoylated are exclusively membrane-bound and mutation of the cysteine palmitoylation sites prevents the normal caveolae localisation of eNOS. Caveolae are microdomains of the plasmalemma implicated in a variety of cellular functions, including transcytosis of molecules and signal transduction events. The major coat protein of caveolae are the caveolin family of proteins which are thought to play a role in the regulation of signalling pathways. A nine amino acid binding motif for caveolin, residues 350-358, has been identified in bovine eNOS (Garcia-Cardena *et al.*, 1997) and mutation to delete these amino acids produced an active enzyme that was uninhibitible by caveolin-1, demonstrating that caveolins are regulators of eNOS.

Both the nNOS and eNOS isoforms are constitutively expressed as noninteractive monomers which only dimerise after binding FMN and FAD (Feldman *et al.*, 1993; Marletta, 1993; Bredt and Snyder, 1994) followed by haem, BH_4 and L-arginine (Baek *et al.*, 1993). An increase in intracellular calcium levels is then required for the formation of calcium/calmodulin (Ca^{2+}/CaM) complexes which bind to the NOS dimers. A conformational change is induced by the binding of Ca^{2+}/CaM that allows the flow of electrons from NADPH in the reductase domain to the catalytic haem site in the

oxidative domain (Abu-Soud and Steuhr, 1993) for the oxidation of L-arginine to NO and L-citrulline.

In comparison to the constitutive isoforms, iNOS monomers bind CaM even at very low concentrations of intracellular calcium (Cho *et al.*, 1992) and so do not require the stimulation of the Ca^{2+}/CaM signalling pathway in order to be activated. Dimerisation and activation of the iNOS enzyme occurs on the binding of haem, BH_4 and L-arginine, all of which are abundant in most cells. So, unlike nNOS and eNOS, the expression of iNOS as functionally active dimers is not constitutive but induced and regulated at the level of mRNA and protein synthesis and degradation.

1.5 Molecular characterisation of NOS1.

The cDNA encoded by rat NOS1 was the first to be cloned and sequenced (Bredt *et al.*, 1991). This group isolated a cDNA with an open reading frame of 4,287 nucleotides which translated into a protein of 1429 amino acids with a predicted molecular weight of 160.5 kDa. Cloning and sequencing of mouse and human nNOS cDNAs followed soon after (Ogura *et al.*, 1993; Nakane *et al.*, 1993). In 1992 the gene encoding human nNOS was assigned to the q14-qter position on chromosome 12 by Southern blot analysis with a human nNOS cDNA fragment (Kishimoto *et al.*, 1992). The position of the gene was more precisely located to 12q24.2 (Marsden *et al.*, 1993) and 12q24.2-24.3 (Xu *et al.*, 1993) by fluorescent *in situ* hybridisation.

Human NOS1 has the largest and most complex locus of the three NOS genes and as such the structural organisation of human NOS1 was the last to be reported (Hall *et al.*, 1994). NOS1 is present as a single copy in the haploid human genome and spans a

region greater than 160 kb. The major transcription initiation site is located 28 nucleotides downstream of a TATA box and the gene comprises 29 exons and 28 introns. The full-length open reading frame of NOS1 is 4302 bps, with an 11 nucleotide difference as compared to the original sequence of human nNOS cDNA cloned from skeletal muscle (Nakane *et al.*, 1993). However, this is in agreement with the sequence of human nNOS cDNA cloned from a human neuroblastoma cell line (Fujisawa *et al.*, 1994). The translation initiation and termination sites are present in exons 2 and 29 respectively, giving rise to a 160.8 kDa protein of 1434 amino acids.

Both structural and allelic diversity of nNOS mRNA transcripts has been demonstrated. An imperfect microsatellite repeat ($(GT)_{17}A(TG)_{13}$) is present in the 5′-flanking region of exon 1. In addition, intron 2 has a ($(GT)_{16}$) repeat and in the 3′ UTR of exon 29 a ($(TG)_{17}$) repeat is found. A heterozygosity index of 0.34 (4 alleles found in 36 chromosomes) for the exonic microsatellite repeat indicates allelic diversity of the nNOS mRNA transcript (Hall *et al.*, 1994). In addition to allelic diversity, there is also structural diversity of the nNOS mRNA transcript. A 315 bp in-frame cassette deletion of exons 9 and 10 has been reported for human and mouse nNOS which gives rise to a 1329 amino acid protein of 148.8 kDa (Ogura *et al.*, 1993; Hall *et al.*, 1994; Fujisawa *et al.*, 1994). A 175 nucleotide deletion of nNOS exon 10 has also been detected. However, this deletion introduces a stop codon after only 16 nucleotides and the possibility of a novel 561 amino acid NH_2-terminal protein of only 61.8 kDa (Hall *et al.*, 1994; Ogura *et al.*, 1993). A 102 base pair, 34 amino acid, in-frame insertion between exons 16 and 17, which results in a functional Ca^{2+}/CaM dependent protein of 164 kDa, has also been reported for nNOS. This alternative transcript (nNOSμ) has been detected in mouse skeletal muscle and cardiac muscle (Silvagno *et al.*, 1996), rat penile corpus

cavernosa, urethra, prostate, skeletal muscle, cerebellum, pelvic plexus and bladder (Magee *et al.*, 1996). The same alternative transcript has also been found in human skeletal muscle, prostate, placenta, heart, corpus cavernosum, colon and bladder (Larsson and Phillips, 1998).

nNOS is expressed in a broad range of tissue and cell types. Indeed, the nNOS mRNA transcript has been detected in both the central and peripheral nervous system, the non-adrenergic non-cholinergic (NANC) system (Bult *et al.*, 1990; Desai *et al.*, 1991), the adrenal medulla (Dun *et al.*, 1993; Afework *et al.*, 1994), the distal nephron and macula densa of the kidney (Mundel *et al.*, 1992; Terada *et al.*, 1992; *Wilcox et al.*,1992), pancreatic β cells (Schmidt *et al.*, 1992), skeletal muscle (Nakane *et al.*, 1993; Weiner *et al.*, 1994) and the pituitary gland (Wolff and Datto, 1992) among others.

Alternate use of the first exon appears to be responsible for the diverse expression of nNOS (Wang and Marsden, 1995). Although Hall and colleagues (1994) described a single first exon for human nNOS, other nNOS cDNAs cloned and characterised from human brain, skeletal muscle, kidney and adrenal gland have identified independent 5′ cDNA ends which represent different first exons spliced to a common exon 2 (Marsden *et al.*, 1994a). Xie *et al.* (1995) have described two mRNAs with alternative first exons that predominate in the cerebellum relative to non-CNS tissue, whereas Wang *et al.* (1997) located two additional testis-specific alternative first exons, although these were present in intron 3. These testis-specific mRNA transcripts (nNOSγ) give rise to a truncated nNOS protein (TnNOS), predicted to have 1098 amino acids and a calculated molecular mass of 125 kDa. However, this protein is catalytically active and calcium dependent. The total of exon 1 variants used for transcript initiation by human nNOS currently stands at nine. This degree of diversity is unprecedented in any reported

mammalian gene (Wang *et al.*, 1999) and is not confined to human NOS1 alone. Mice generated to carry a targeted disruption of exon 2 (designated nNOS$^{\Delta/\Delta}$) (Huang *et al.*, 1993) were subsequently found to express an alternatively spliced transcript that does not contain exon 2 (Brenman *et al.*, 1996). This N-terminally truncated 136 kDa protein (nNOSβ) has subsequently been detected in wild type mice (Eliasson *et al.*, 1997). In addition, an analogue of human TnNOS exists in mice, which initiates translation from the same ATG in exon 5 as human TnNOS to produce a 125 kDa protein (Brenman *et al.*, 1996). At least five distinct murine nNOS mRNA transcripts have been identified (Brenman *et al.*, 1997). Rat nNOS mRNA transcripts have also been found with alternative first exons and a total of 11 different splice variants have been isolated (Lee *et al.*, 1997; Oberbäumer *et al.*, 1998).

Thus, the expression of nNOS appears to be regulated at three levels (Oberbäumer *et al.*, 1998). i), at the transcriptional level by differential activation of tissue-specific alternative promoters, ii), at the mRNA level via alternative splicing and the putative influence of the 5′ UTRs on translation and iii), at the protein level by the production of different N-termini which may or may not allow interaction of nNOS with certain other proteins.

It is now understood that NO produced from NOS1 in both the central and peripheral nervous systems acts as a neurotransmitter. It is synthesised on demand and not stored in synaptic vesicles to be released by exocytosis upon membrane depolarisation. It diffuses from neuron to neuron to act directly on intracellular components. In addition, the activity of NO is not terminated like conventional neurotransmitters as it is not enzymatically degraded or subject to reuptake mechanisms. Instead, NO terminates its activity when it reacts with a substrate.

Under normal physiological conditions the production of NO results in local modulation of cellular activity through modification of proteins such as sGC which is involved in neurotransmitter release, synaptic plasticity and cerebral blood flow (Dawson and Dawson, 1996). However under pathophysiological conditions such as ischaemic damage, excessive glutamate is released from neurons and as a result, excess NO is produced. In this situation the toxic effects of NO appear to occur through its reaction with superoxide anion to form peroxynitrite and not through its activation of sGC to elevate cGMP levels. This observation has been confirmed by the lack of effect of GC inhibitors or cell permeable analogues of cGMP on NMDA or NO neurotoxicity (Dawson *et al.*, 1993; Lustig *et al.*, 1992).

In the hippocampus NO, formed by stimulation of postsynaptic NMDA-type glutamate receptors, acts on presynaptic neurons to modulate neurotransmitter release. This leads to a modification of synaptic transmission, or long-term potentiation, which is the basis of memory formation (Schuman and Madison, 1991; Bohme *et al.*, 1991). NO has also been implicated in the neurotoxicity seen in MPTP-induced Parkinsonism in baboons. MPTP produces clinical, biochemical and neuropathological changes similar to those seen in idiopathic Parkinson's disease (Hantraye *et al.*, 1996).

Additionally, NO has been implicated in a number of functions including the nonadrenergic noncholinergic (NANC) neurotransmission in several peripheral systems (Rand, 1992). NANC nerve pathways play a particularly important role in producing relaxation of smooth muscle in the gastrointestinal, urogenitial and respiratory tracts and in cerebral circulation (Burnstock, 1981). nNOS has been found in nerve fibres that supply the intestine, retina, adrenal medulla, and blood vessels (Bredt *et al.*, 1990).

1.6 Molecular characterisation of NOS2 and NOS3.

For the purpose of this thesis the molecular characterisation of the NOS2 and NOS3 isoforms are described together.

The gene for human iNOS, NOS2, has been localised to region 17q11.2-q12 of chromosome 17 (Marsden *et al.*, 1994; Chartrain *et al.*, 1994; Xu *et al.*, 1994). It consists of 27 exons and 26 introns (Xu *et al.*, 1996) and spans approximately 37 kb (Chartrain *et al.*, 1994). The sites for translation initiation and termination are found in exons 2 and 27 respectively. The NOS2 full-length open reading frame is 3459 bp which encodes a protein of 1153 amino acids and a predicted molecular mass of 131 kDa (Charles *et al.*, 1993). Many nucleated cell types and tissues are capable of iNOS expression upon stimulation with the appropriate cytokines. These include macrophages (Xie *et al.*, 1992; Lyons *et al.*, 1992; Lowenstein *et al.*, 1992; Lorsbach *et al.*, 1993; Bosca and Lazo, 1994), chondrocytes (Palmer *et al.*, 1993), vascular smooth muscle cells (Koide *et al.*, 1994; Beasley and Elridge, 1994; Sirsjo *et al.*, 1994; Perrella *et al.*, 1994; MacNaul and Hutchinson, 1993), renal tubular epithelium (Markewitz *et al.*, 1993), mesangial cells (Kunz *et al.*, 1994) glia and neurons (Koprowski *et al.*, 1993).

In the immune system NO synthesised by iNOS acts as a primary defence mechanism. Macrophages are highly effective killers of intracellular and extracellular pathogens and generate large amounts of NO when activated by cytokines or lipopolysaccharide (LPS) (Steuhr *et al.*, 1989). However, whilst induced NO production by the immune system is an effective host defence mechanism, increased NO formation is also seen in several inflammatory diseases including arthritis, ulcerative colitis and Crohn's disease (Schmidt and Walter, 1994). In rodent models, symptoms such as erythema and vascular leakiness

are reversed by NOS inhibitors. Also, the unregulated production of NO is self-destructive and is implicated in sepsis, rejection of allografted organs, and some autoimmune disease.

The gene for human eNOS, NOS3, has been assigned to the 7q35-q36 region of chromosome 7 (Marsden *et al.*, 1993; Xu *et al.*, 1994; Robinson *et al.*, 1994). It consists of 26 exons and 25 introns and spans approximately 21 kb of human genomic DNA (Marsden *et al.*, 1993). NOS3 is present as a single copy in the haploid human genome and gives rise to an mRNA of 4052 nucleotides (Marsden *et al.*, 1993; Robinson et al., 1994; Nadaud *et al.*, 1994). With translation initiation and termination sites in exons 1 and 26, respectively, the full-length open reading frame is 3609 bp and encodes a protein of 1203 amino acids. Expression of eNOS is mostly restricted to vascular endothelial cells.

In the cardiovascular system, NO from NOS3 acts as a vasodilator by inducing smooth muscle cell relaxation. NO released by the vascular endothelium is therefore able to regulate both blood flow and pressure by influencing vascular tone (Moncada and Higgs, 1993). Pathologically, oxidised low-density lipoproteins (LDL) have been shown to impair endothelium-derived relaxation (Andrews *et al.*, 1987). In culture medium, oxidised LDL has been shown not only to decrease eNOS mRNA levels but also to bind to and inactivate NO itself (Chin *et al.*, 1992). In addition, reduced NO-mediated vasorelaxation is exhibited by atherosclerotic vessels (Creager *et al.*, 1990; Luscher *et al.*, 1991) and the therapeutic treatment of hypercholesterolaemic individuals with an infusion of L-arginine improves endothelium-dependent vasodilation (Creager *et al.*, 1992).

Thus, whilst it is evident that NO is an integral messenger fundamental to the maintenance of normal human physiology, perturbation of NO levels has been implicated in numerous pathophysiological states. Insights into the regulation of NOS is therefore important for understanding of these disease processes and for the development of therapeutic treatments.

1.7 Evolution of NOS.

Despite being localised to different chromosomes, human NOS1, NOS2 and NOS3 genes show a high degree of similarity in their genomic organisation and intron/exon structure, including the location of splice junctions and exon sizes (Wang and Marsden, 1995). It is thought therefore that the high degree of conservation between isoforms could reflect the duplication and chromosomal translocation of a primitive NOS gene which would have been present in an ancestor common to both vertebrates and invertebrates. Indeed, the existence of the L-arginine: NO pathway has been suggested in several species of an early phylogenetic origin, including the slime mould *Physarum polycephalum* (Werner-Felmayer *et al.*, 1994), hydra (Colasanti *et al.*, 1995), Praying Mantis (D'allessio *et al.*, 1982), the horseshoe crab *Limulus polyphemus* (Radomski *et al.*, 1991) and the starfish *Marthasteria glacialis* (Martinez *et al.*, 1994).

The first direct evidence for NOS in invertebrates came from the work of Regulski and Tully in 1995. They reported the molecular and biochemical characterisation of a *Drosophila* NOS homologue (*d*NOS) which has 43% amino acid identity to rat nNOS and contains putative binding sites for calmodulin, FMN, FAD and NADPH. *d*NOS also has an alternative RNA splicing pattern identical to that for vertebrate nNOS. When

expressed in cell culture, activity of *D*NOS protein is dependent on Ca^{2+}/CaM levels. Other invertebrate NOS genes and cDNAs have since been reported, they are the blood-sucking insect *Rhodnius prolixus* (Yuda *et al.*, 1996), the mosquito *Anopheles stephensi* (Luckhart *et al.*, 1998; Luckhart and Rosenberg, 1999) and the great pond snail *Lymnae stagnalis* (Korneev *et al.*, 1998).

There is also much evidence for the presence of NOS and NO systems in fish. Endothelium-dependent relaxation and contraction in the ventral aorta of the rainbow trout, *Oncorhynchus mickiss*, was demonstrated with the use of acetylcholine (ACh) and the calcium ionophore A23187 (Miller and Vanhoutte, 1986). Further evidence for NOS and NO systems in fish has since emerged. Two studies demonstrated an increase in cerebral blood flow when the brains of two different teleosts, the crucian carp *Carassius carassius* and rainbow trout, were superfused with ACh (Hylland and Nilsson, 1995; Söderström *et al.*, 1995). In addition, these groups demonstrated that NOS inhibitors could completely block the ACh-dependent response and also that administration of the NO donor sodium nitroprusside (SNP) increased cerebral blood flow velocity. Much of the evidence for NOS and NO systems in fish has come through NADPH diaphorase (NADPHd) enzyme histochemistry and NOS immunocytochemistry. In 1993, the presence and distribution of NOS in the CNS of the rainbow trout was demonstrated by NADPHd histochemistry (Schober *et al.*, 1993). Both NADPHd histochemistry and NOS immunocytochemistry were used to localise NOS in the retina of atlantic salmon and the brain of rainbow trout (Holmqvist *et al.*, 1994; Östholm *et al.*, 1994). For the NOS immunocytochemistry, rabbit polyclonal antibodies against the C-terminal fragment of cloned rat nNOS were used. In addition, nitrite release has been demonstrated from a macrophage cell line taken from the goldfish, *Carassius auratus*, following challenge

with LPS and macrophage activation factor (MAF) (Neumann *et al.*, 1995). A similar effect was observed for head kidney cells isolated from the channel catfish following challenge with the live parasite *Edwardsiella ictaluri* (Schoor *et al.*, 1994). More recently RT-PCR has been used to identify iNOS from both the rainbow trout and the goldfish (Grabowski *et al.*, 1995; Laing *et al.*, 1996). iNOS mRNA was detected in the head kidney and gill of rainbow trout two days after it had been challenged *in vivo* with a genetically attenuated ($AroA^-$) fish bacterial pathogen (*Aeromonas salmonicida*). In addition, a partial cDNA of nNOS has been cloned from the atlantic salmon (Oyan *et al.*, 1998). Thus, like mammals, the presence of a NOS/NO system has been demonstrated in the cardiovascular, nervous and immune systems of fish and suggests that they could potentially have all three NOS isoforms.

1.8 Aims of this thesis.

The overall aim of this thesis is to research the NOS genes of an evolutionally distant vertebrate precursor of mammals.

The specific aims of this thesis were:

1) to isolate, sequence and analyse NOS genes from the genome of the Japanese pufferfish *Fugu rubripes*.

2) to isolate NOS cDNA from *Fugu*, to purify and characterise any recombinant NOS protein.

CHAPTER 2

MATERIALS AND METHODS.

CONTENTS.

2.1. Random-primed radiolabelling of cDNA fragments with [α^{32}P] dCTP.

cDNA fragments were labelled with [α^{32}P] dCTP using a commercially available random-priming kit from which all reagents were obtained unless otherwise stated (Boehringer Mannheim). 25 ng of cDNA, in a volume of 12 µl, was heated to 100°C for 10 min in an eppendorf tube and chilled on ice. The denatured cDNA was then mixed with 1 µl each of 0.5 mM dATP, 0.5 mM dGTP and 0.5 mM dTTP, 2 µl of 10 x concentrated hexanucleotide reaction mix, 1 µl Klenow DNA polymerase (2 units/µl) and 2 µl (20µCi) of [α^{32}P] dCTP (specific activity: 3000 Ci/mmol) (Amersham). The reaction was allowed to proceed at 37°C for 30 min and then stopped by the addition of 4 µl of 0.1 M EDTA, pH 8.0. The reaction mixture was loaded onto a 5 ml disposable NICK column (Pharmacia), pre-equilibriated with 5 ml of NICK column buffer (10 mM Tris-HCl, 1 mM EDTA pH 8.0), allowed to move into the gel bed and washed onto the column with 400 µl of NICK column buffer. [α^{32}P] dCTP-labelled cDNA was eluted by the addition of 400 µl NICK column buffer and collected in an eppendorf tube. Unincorporated [α^{32}P] dCTP was retained in the column which was discarded. The labelled cDNA was heated to 100°C for 10 min and chilled on ice prior to its use as a probe.

2.2. Random-primed labelling of cDNA fragments with digoxigenin-dUTP.

cDNA fragments were labelled with digoxigenin-dUTP (DIG-dUTP) using a commercially available random-priming kit from which all reagents were obtained unless

otherwise stated (Boehringer Mannheim). 25 ng of cDNA, in a volume of 15 µl, was denatured by heating to 100°C for 10 min in an eppendorf tube and chilled on ice. This was then mixed with 2 µl of 10 x concentrated hexanucleotide mix, 2 µl dNTP mix (1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM dTTP, 0.35 mM DIG-dUTP, pH 7.5) and 1 µl of Klenow DNA polymerase (2 units/µl). The reaction was allowed to proceed for at least 60 min at 37°C and then stopped by the addition of 4 µl of 0.1 M EDTA, pH 8.0. DIG-dUTP labelled cDNA was precipitated by mixing with 2.5 µl of 4 mM LiCl and 75 µl of ethanol (-20°C), incubating at -70°C for 30 min and centrifuging at 20,000 x *g* for 15 min followed by a 70% ethanol wash and vacuum drying. The DIG-dUTP labelled cDNA was resuspended in 50 µl of TE buffer (20 mM Tris-HCl, pH 7.0, 2.5 mM EDTA), boiled for 10 min at 100°C and chilled on ice prior to use as a probe.

2.3. Screening of a *Fugu* genomic cosmid library.

A *Fugu* genomic library, constructed in the cosmid vector Lawrist 4 as high density gridded filters on Hybond-N^{TM} and Hybond-N+^{TM} nylon membranes (Amersham) by Elgar and Nizetic, was obtained from the MRC HGMP Resource Centre (http://fugu.hgmp.mrc.ac.uk). The filters, which represent an approximately 8-fold coverage of the *Fugu* genome (Yeo *et al.*, 1997), were incubated for 1 hour at 65°C in pre-hybridisation solution (5 x SSPE (Gibco BRL), 5 x Denhardt's solution (BDH), 20 µg/ml denatured salmon testes DNA (Sigma) and 0.5% w/v SDS). At the end of the prehybridisation, a specific [α^{32}P] dCTP labelled DNA probe, prepared as described previously, was added directly to the solution and incubation was allowed to proceed at

65°C overnight. Following incubation with the probe, the library filters were removed from the hybridisation solution and washed twice in 2 x SSPE, 0.1% w/v SDS at room temperature for 10 min each followed by a further wash in 1 x SSPE, 0.1% w/v SDS at 65°C for 15 min. The filters were removed from the last wash, wrapped in Saran-Wrap (Dow Chemical Company) and exposed to X-ray film (Kodak) for 24 hours. Following autoradiography, the film was developed and positive cosmid clones identified. The filters were stripped by incubating with 0.4 M NaOH at 45°C for 30 min followed by a further incubation in 0.1 x SSC, 0.1% w/v SDS, 0.2 M Tris-HCl, pH 7.5 at 45°C for 15 min. The filters were wrapped in Saran-Wrap and exposed to X-ray film for 48 hours. Following autoradiography, the film was developed to ensure the filters had been stripped. The filters were placed in a sealable bag and stored at 4°C until required for re-probing.

2.4. Isolation of *Fugu* cosmid DNA.

Cosmid DNA was isolated using a modification of a previously published method (Birnboim and Doly, 1979). Colonies carrying a putative cosmid of interest were inoculated into 5 ml of terrific broth (TB) (1.2% w/v bacto-tryptone (Gibco BRL), 2.4% w/v yeast extract (Gibco BRL), 0.4% v/v glycerol (BDH), 0.23% w/v KH_2PO_4 (BDH), 1.24% w/v K_2HPO_4 (BDH)) containing 30 µg/ml kanamycin (Sigma) and grown overnight at 37°C with vigorous shaking. At the end of the growth period, glycerol was added, to a concentration of 7.5% v/v, to 1 ml of culture for long term storage at -80°C. 4 ml of culture was transferred to eppendorf tubes and centrifuged at 20,000 x *g* for 2 min. The supernatant was discarded and the pellet resuspended in 200 µl of GTE (25

mM Tris HCl, pH 8, 10 mM EDTA, 50 mM glucose), 5 µl of 10 mg/ml RNase A (Sigma) was added and incubated at room temperature for 10 min. 400 µl of 0.2 M NaOH/1% SDS was added, mixed by inversion and incubated on ice for 5 min. After incubation, 300 µl of 3M Kac was added, mixed by inversion and incubated on ice for 10 min, this was followed by centrifugation at 20,000 x *g* for 10 min. After centrifugation, the supernatant was decanted into a fresh eppendorf tube containing 1 ml of cold ethanol, vortexed briefly and allowed to stand for 1 min before being centrifuged again for 10 min. The supernatant was discarded, the pellet washed in 70% ethanol and allowed to air-dry. 50 µl of TE buffer (25 mM Tris HCl, pH 8, 10 mM EDTA) was added to the pellet and incubated at 4°C overnight to resuspend.

2.5. Isolation of plasmid DNA.

Plasmid DNA was isolated using a commercially available kit (QIAGEN), which uses a modification of a previously published method (Birnboim and Doly, 1979), all reagents were provided with the kit unless otherwise stated. Colonies carrying a putative plasmid of interest following transformation of bacteria were inoculated into 5 ml of LB broth (1% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract and 1% w/v NaCl, pH 7.0) (Imperial) containing 100 mg/ml ampicillin (Sigma) and grown overnight at 37°C with vigorous shaking. At the end of the growth period, 500 µl of the bacterial culture was removed and added to an equal volume of 50% glycerol for long term storage at -80°C. The remainder of the bacterial culture was pipetted into eppendorf tubes and centrifuged for 2 min at 20,000 x *g*. The supernatant was removed and the bacterial pellet

resuspended in 250 µl of resuspension buffer (buffer P1 plus RNase A) by vortexing. 250 µl of lysis buffer (buffer P2) was added and mixed by gentle inversion, the lysis reaction was allowed to proceed for 5 min at room temperature. 350 µl of neutralisation buffer (buffer N3) was added and mixed by gentle inversion, to halt the lysis reaction, followed by centrifugation at 20,000 x *g* for 10 min. After centrifugation, the supernatant was applied to a QIAprep spin column placed in a 2 ml collection tube and centrifuged for 1 min, the flow-through was discarded. The QIAprep spin column was washed by the addition of 750 µl of wash buffer (buffer PE) and centrifugation at 20,000 x *g* for 1 min. The flow-through was discarded and the column centrifuged for a further minute to remove residual wash buffer. Plasmid DNA was eluted from the column by the addition of 50 µl sterile distilled water, incubation at room temperature for 1 min and centrifugation for 1 min.

2.6. Restriction endonuclease digestion of DNA.

A range of different buffer systems and restriction endonucleases (Boehringer Mannheim) were used in these studies to digest DNA into required fragments (see tables 2.1 and 2.2). Restriction digestion of DNA was carried out at 37°C unless otherwise stated. In general, up to 2µg of DNA was digested in the presence of 10-20 units of enzyme in a volume of 20-50µl for 1-4 hours. Plasmid DNA was similarly digested except that protruding sticky ends were dephosphorylated on the 5′ phosphate residue by the addition of 2µl (1 unit/µl) shrimp alkaline phosphatase (Amersham) for the last hour of the digest followed by a 15 min incubation at 65°C to heat inactivate the enzyme.

Table 2.1. Buffer composition used in restriction digests.

(All concentrations are mM and are final in the reaction mixture)

Component	Buffer A	Buffer B	Buffer L	Buffer M	Buffer H
Tris acetate	33	-	-	-	-
Tris.HCl	-	10	10	10	10
Mg-acetate	10	-	-	-	-
$MgCl_2$	-	5	10	10	10
K-acetate	66	-	-	-	-
NaCl	-	100	-	50	100
DTE*	-	-	1	1	1
DTT*	0.5	-	-	-	-
2-Me*	-	1	-	-	-
pH at 37°C	7.9	8.0	7.5	7.5	7.5

*DTE= Diththioerythritol. DTT= Dithiothreitol. 2-Me= β-mercaptoethanol.

Table 2.2. Restriction endonucleases used and their recognition sites.

(In all cases sequences are written 5′ to 3′ on the top strand)

Restriction Endonuclease and Buffer	Recognition Sequence
*Bam*HI B	G↓GATCC CCTAG↑G
*Eco*RI H	G↓AATTC CTTAA↑G
*Eco*RV B	GAT↓ATC CTA↑TAG
*Hind*III B	A↓AGCTT TTCGA↑A
*Kpn*I L	GGTAC↓C C↑CATGG
*Sal*I H	G↓TCGAC CAGCT↑G
*Xba*I H	T↓CTAGA AGATC↑T
*Xho*I H	C↓TCGAG GAGCT↑C

2.7. Agarose gel electrophoresis.

Nucleic acids were routinely analysed by electrophoresis through agarose gels. Gels were made at a range of concentrations from 0.8% to 1.5% depending on the size of the fragment being analysed. Gels were made by dissolving the required amount of dry agarose powder (Ultrapure electrophoresis grade) (Gibco BRL) in 1 x TBE buffer (20 mM Tris-HCl, pH 8.3, 90 mM boric acid, 2.5 mM ethylenediamine tetra-acetic acid (EDTA)) (Gibco BRL) at 100°C. Solutions were allowed to cool to 45°C and ethidium bromide (Sigma) added to a final concentration of 10 ng/ml. Gels were cast into minigel tanks where they were allowed to set. After setting, gels were submerged in 1 x TBE and the well-forming combs were removed. Samples to be electrophoresed were mixed with 10% v/v 10 x BlueJuice™ loading buffer (65% w/v sucrose, 10 mM EDTA, 0.3% bromophenol blue, 10 mM Tris-HCl, pH 7.5) (Gibco BRL) and loaded into the relevant wells on the gel. DNA ladders (100bp: 0.1-1.5 kb, 1kb: 0.5-12.0 kb, 1 kb Plus: 0.1-12.0kb,1 kb Extension: 0.5-40.0 kb, λ *Hind*III: 0.125-23.1 kb) (Gibco BRL), relevant to the size of the fragment being analysed, were similarly treated and run alongside samples to serve as molecular weight markers. Gels were electrophoresed at a constant current of 10 mA/cm gel length until the bromophenol blue marker had travelled the required distance. Gels were then analysed under ultra-violet light and photographed.

2.8. Southern blot analysis of *Fugu* cosmid DNA.

Analysis of *Fugu* cosmid DNA by restriction digestion, gel electrophoresis and labelled probe hybridisation was performed according to a previously published method

(Southern, 1975). 2 µg of cosmid DNA was digested with the appropriate enzymes and subjected to agarose gel electrophoresis as previously described. Following electrophoresis, the gel was photographed alongside a fluorescent ruler under UV light and then incubated in 0.25 M HCl for 15 min at room temperature. 3 sheets of 3MM filter paper were cut to the same width as the gel but left long enough to act as the wick for the transfer, these filter papers were placed on a perspex block (approximately the same size as the gel) in a tray containing 0.4 M NaOH. The 0.4 M NaOH was allowed to saturate the wick by capillary action and the gel placed on top of the wick, whilst ensuring that the gel was not stretched. 1 sheet of GeneScreen plus® nylon hybridisation transfer membrane (NEN) was cut to the same size as the gel and soaked in 0.4 M NaOH. The membrane was placed on top of the gel and lightly rolled over with a glass rod to remove any air bubbles. 3 further sheets of 3MM filter paper were cut to the same size as the gel, soaked in 0.4 M NaOH and placed on top of the nylon membrane. The assembled blot was surrounded by Saran-wrap to prevent transfer buffer flowing around the outside of the blot. A stack of paper towels, approximately 15 cm high, were placed on top of the blot assembly followed by a glass plate and a 500 g weight. Capillary transfer was allowed to proceed overnight at room temperature. At the end of the blot the gel was removed and discarded. The nylon membrane was washed in 2 x SSC for 5 min and allowed to air dry. The blot was incubated for 1 hour at 65°C in prehybridisation solution (4 x SSC, 4 x Denhardt's, 0.1% w/v SDS, 50 µg/ml denatured salmon testes DNA). Following prehybridisation, a specific DIG-dUTP labelled probe, prepared as described previously, was added directly to the solution and incubation allowed to proceed at 65°C overnight. Following hybridisation with the probe, the blot was washed twice in 2 x SSC, 0.1% w/v SDS at room temperature for 15 min each, once

in 1 x SSC, 0.1% w/v SDS at 65°C for 15 min and then twice more in 0.7 x SSC, 0.1% w/v SDS at 65°C for 15 min each. Detection of DIG-labelled nucleic acids was performed using a commercially available kit (Boehringer Mannheim). All reagents were provided unless otherwise stated. All incubations were performed with shaking at room temperature unless otherwise stated. The membrane was equilibrated in buffer 1 (100 mM maleic acid, 150 mM NaCl, pH 7.5) for 1 min and then blocked in buffer 2 (1% w/v blocking reagent dissolved in buffer 1) for 30-60 min. Following blocking, buffer 2 was decanted and replaced with Anti-DIG-alkaline phosphatase diluted 1:10,000 (75 units/ml) in buffer 2. The membrane was incubated with the antibody solution for 30 min and then washed twice in buffer 1 plus 0.3% Tween® 20 (BDH) for 15 min each wash. The membrane was then equilibrated in buffer 3 (100 mM NaCl, 50 mM $MgCl_2$, 100 mM Tris-HCl, pH 9.5) for 2 min. Lumigen™ PPD was diluted 1:100 in buffer 3 and applied to the wet membrane (0.5 ml/100 cm^2) which was then wrapped in Saran-Wrap and incubated at 37°C for 15 min. The membrane was exposed to X-ray film for between 30 min and 24 hours after which time the film was developed and compared to the original photograph of the gel.

2.9. Purification of DNA fragments.

Routinely, DNA fragments produced by restriction digestion or PCR were purified from agarose gels using the commercially available QIAquick gel extraction kit (QIAGEN), from which all reagents were obtained unless otherwise stated. In these cases the DNA was subjected to agarose gel electrophoresis as described previously and the relevant band was excised from the gel using a sterile scalpel. The gel slice was incubated in 3

volumes of buffer QG at 50°C for 10 min with vortexing every 2-3 min to help the gel dissolve. 1 gel volume of isopropanol (BDH) was added, the sample mixed and applied to a QIAquick column, in a collection tube, by centrifuging for 1 min at 20,000 x *g*. The column was then washed with 0.75 ml of buffer PE by centrifuging twice for 1 min, removing flow-through from the collection tube in between spins. The DNA was eluted from the column by adding 30 µl of distilled water (pH range 7.0-8.5) to the column, leaving for 1 min and then centrifuging for 1 min. 3 µl (10%) of purified DNA fragment was then analysed by agarose gel electrophoresis as described previously.

2.10. Ligation and subcloning of restriction endonuclease digested DNA fragments into plasmid vectors.

Following restriction digestion and purification, DNA fragments with compatible sticky ends were ligated using DNA ligase isolated from bacteriophage T4 (Gibco BRL). Fragments were mixed with 50 ng of appropriately digested and purified plasmid DNA (see table 2.3 for plasmids used in these studies) in a 2:1 or 3:1 ratio and incubated with 1 µl of T4 DNA ligase (1 unit/µl) in 1 x ligase buffer (50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT and 5% polyethylene glycol 8000) (Gibco BRL) in a total volume of 10 µl. Control ligations, replacing DNA fragments with water but retaining the plasmid DNA, were set up to calculate the degree of self ligation. Ligations were allowed to proceed overnight at 14°C. No further purification was required prior to bacterial transformation.

2.11. Ligation and subcloning of PCR amplified fragments into plasmid vectors.

Following PCR and any necessary purification, DNA fragments containing 3′ A overhangs, generated by the nontemplate-dependent terminal transferase activity of *Taq* polymerase, were ligated into the plasmid vector pCR®2.1-TOPO (Invitrogen). pCR®2.1-TOPO is supplied linearised with single 3′ T overhangs and topoisomerase for efficient ligation of PCR inserts into the vector. Sterile distilled water was added to 0.5-2 µl of a typical PCR sample, with an average insert size of 400 to 1000 bp, to a final volume of 4 µl. 1 µl of pCR®2.1-TOPO was mixed in with the PCR fragment and the reaction allowed to proceed at room temperature for 5 min. After incubation, the reaction was briefly centrifuged and kept on ice prior to immediate transformation.

2.12. Transformation of subcloning efficiency DH5α *Escherichia coli.*

A heat shock method was utilised to promote the uptake of plasmid vectors into subcloning efficiency DH5α *E. coli* host cells (Gibco BRL). 3 µl of the total ligation mixture was added to 50 µl of competent bacterial cells, previously thawed on ice, mixed by tapping and incubated on ice for 30 min. Positive control transformations were set up replacing the experimental sample with 500 pg of the plasmid pUC19 to assess the efficiency of plasmid uptake. The DNA-cell mixture was then incubated at 37°C for 20 seconds before a second incubation on ice for 2 min. 0.95 ml of LB broth was added to the transformation mix and this was then incubated at 37°C with shaking at 225 rpm for 1 hour. 100 µl and 900 µl of the neat transformation mix, pelleted by centrifugation at

20,000 x *g* for 10 seconds and resuspended in 100 µl of broth, were plated onto 90 mm petri dishes (Sterilin) containing L-agar (1% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract, 1% w/v NaCl and 1.5% agarose, pH 7.0) (Imperial), and the appropriate antibiotic to permit selection of transformants (see table 2.3). Plates were incubated overnight at 37°C and examined the following day for growth of transformed bacteria. In some experiments, and with certain plasmids, the L-agar contained 1 mM isopropyl thio-β-D-galactosidase (IPTG) (Gibco BRL) and 50 µg/ml 5-bromo-4-chloro-3-indolyl-β-D galactosidase (X-gal) (Gibco BRL). This allowed for the selection of bacteria carrying recombinant plasmids by the appearance of white colonies. Those plasmids not carrying an inserted DNA fragment retained the ability to metabolise lactose analogues and therefore produced blue colonies due to the presence of IPTG and X-gal.

2.13. One Shot™ Transformation of TOP10 *Escherichia coli.*

50 µl of TOP10 competent *E.coli* cells (Invitrogen), one vial per transformation, were thawed on ice. Once thawed, 2 µl of 0.5 M β-mercaptoethanol (Invitrogen) was mixed into the cells by gentle stirring with a pipette tip. 2 µl of the TOPO-Cloning™ reaction, section 2.7.2, was added to the cells, mixed by tapping and incubated on ice for 30 min. The cells were then heat shocked at 42°C for 30 seconds followed by a 2 min incubation on ice. 250 µl of room temperature SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) (Gibco BRL) was added, mixed and the cells recovered at 37°C with shaking for 30 min. 100 µl of each transformation was plated onto 90 mm petri dishes containing L-agar and the

appropriate antibiotic (table 2.3) to permit selection of transformants (see table 3.2). Plates were incubated overnight at 37°C and examined the following day for growth of transformed bacteria. In all experiments utilising the pCR®2.1-TOPO plasmid vector the L-agar contained 40 µl of 40 mg/ml X-gal to allow for the selection of bacteria carrying recombinant plasmids by blue/white screening, as discussed previously.

Table 2.3. Plasmids used in these studies for cloning and expression.

Plasmid	Source	Size (kb)	blue/white	Antibiotic	Use
pUC19	Gibco BRL	2.8	Yes	Ampicillin at 100 µg/ml	Positive control for transformation
pBsSKII+	Stratagene	2.9	Yes	"	General subcloning. Sequencing
pCR2.1®-TOPO	Invitrogen	3.9	Yes	Ampicillin at 64 µg/ml	Subcloning of PCR fragments
pVL1393	Stratagene	9.8	No	Ampicillin at 100µg/ml	Expression from insect cells using baculovirus

2.14. Colony PCR.

Following transformation, as an alternative to restriction endonuclease digestion analysis of plasmid DNA, PCR was performed on colonies of interest. A sterile Gilson pipette tip was used to pick colonies and then dipped and mixed into a 20 µl PCR master mix (1 x PCR buffer (Gibco BRL), 2 mM $MgCl_2$ (Gibco BRL), 0.4 mM dNTPs (Promega), 1pmole/µl T3 and T7 primers (Stratagene), 0.125 units of Taq polymerase (Gibco BRL)), the tip was then transferred to 5 ml of LB broth, plus appropriate antibiotics, for

overnight culture at 37°C. The PCR reaction was then incubated at 95°C for 5 min and then subjected to the following cycling parameters: 94°C x 30 seconds; 60°C x 1 min; 72°C x 1 min for 30 cycles followed by 5 min at 72°C. Once cycling was complete, one-tenth of the reaction volume was analysed by agarose gel electrophoresis as described previously.

2.15. PCR amplification of *Fugu* genomic DNA.

Advantage®-GC genomic polymerase mix (Clontech) was used in 50 µl PCR reactions to amplify those regions of the *Fugu* NOS gene not obtained by cloning from cosmid DNA. Advantage®-GC genomic polymerase mix contains two DNA polymerases, one of which provides 3′ - 5′ proof-reading activity, and an automatic "hot start" antibody to increase both the specificity and sensitivity of PCR. 100 ng of *Fugu* genomic DNA (MRC HGMP Resource Centre (http://fugu.hgmp.mrc.ac.uk) was used per reaction with 0.2 µM of each specific primer (see table 2.4 for a list of primers used) plus a PCR master mix (1 x GC genomic PCR reaction buffer, 1.1 mM $Mg(OAc)_2$, 0.2 mM dNTPs). The reaction components were mixed in a 0.2 ml tube and then heated to 95°C for 5 min, 1 µl of 50 x Advantage-GC polymerase mix was then added and the reaction subjected to the following cycling parameters: 95°C x 1 min for 1 cycle; 94°C x 30 seconds, 68°C x 12 min for 30 cycles; 68°C x 12 min for 1 cycle and then a 15°C hold. Once cycling was complete, one-tenth of the reaction volume was analysed by agarose gel electrophoresis as described previously.

Table 2.4. Primers used for the amplification of the *Fugu* NOS1 gene.

Primer	Sequence (5′ → 3′)
Gef1	GCCCGTTGCTCTGTATTTATAAAGCGCCACG
5RACESP1	TTGCAGTGGAGAGGACAGGTTCTC
nNOS2for	GCCCTGGAGATGTTGAAGAATGTGCTGC
GSP1	CATGAGCTGTTGTGCAGTCTCTGGCTAC
Ge19f	CGTCTGACATACACGGCCGAGGCG
Ge20r	ACCTTCTTCTTGTGCACGCCGTAC
Ge20f	AGGTCCATGCAGCAAAGATGCTCG
Ge21r	CCTCCAGCTTATCTATGAGAGCCG
Ge21f	CACCATTCTCGTACGGCTGGACAC
Ge22r	AGGCCTGGTTGATGGTGCAGGGAG
Ge22f	GTAACTGGACCAATGAGACTCGGG
Ge23r	TGCAGGTCTGGAGAGGAGCTGATG
Ge23f	CCTCCATCCAGATGCCTTCTACGC
Ge24r	GAGCCACGACGAACACACTCCATG
Ge24f	CATGGAGTGTGTTCGTCGTGGCTC
Ge25r	GGGAAGCTGGAAGGATGGAGCAC
Ge25F	GTGCTCCATCCTTCCAGCTTCCCA
Ge26r	GCCGTGTACAGCTCCTTGAACACG
Ge26f	TCGAGTCATGCCCAATGATCCTGG
Ge27r	GATGTGTCCTCCCTCCTCCCTCAG
Ge27f	GTGCAGGATGCACTGCGTGAGCAG
Ge28r	CGTAGGTGCGCAGGGTGACTCC
Ge28f	GAGTCACCCTGCGCACCTACGAGG
FUGNOS29	ACAGTATTCTTCAGCCAGCGCCGCCGCC

2.16. Sequencing of DNA templates.

The nucleotide sequence of DNA templates was analysed by a modification of the dideoxy chain termination method (Tabor and Richardson, 1995) using a commercially available kit (Amersham) from which all reagents were taken unless otherwise stated. Approximately 250 ng of plasmid (double-stranded) DNA was combined, on ice, with 4 µl of Thermo Sequenase II reagent premix and 5 µM primer (primers used were either standard M13 forward and reverse primers and universal primers or designed specifically

against sequence data obtained and synthesised by Genosys) in a total volume of 20 µl. The sequencing reaction was placed in a thermal cycler and subjected to the following cycle parameters: 96°C x 30 seconds, 50°C x 15 seconds and 60°C x 1 min for 25 cycles followed by a 4°C hold. Control reactions were performed as described above substituting experminental samples for 1 µl of control DNA (M13mp18, 0.2 µg/µl) and using 1 µl of control primer (universal cycle primer, 23-mer, 0.5 pmol/µl). After cycle sequencing, the termination reaction was transferred to a fresh eppendorf tube containing 2 µl of sodium acetate/EDTA buffer (1.5 M sodium acetate, pH >8; 250 mM EDTA) and 60 µl of cold (-20°C) 100% ethanol and vortexed to mix. The reaction was incubated on ice for 15 min to precipitate the DNA and then centrifuged at 20,000 x *g* for 15 min at room temperature. Once centrifuged, the supernatant was aspirated and the DNA pellet washed with 250 µl of cold (-20°C) 75% ethanol followed by a further centrifugation of 2 min. Again the supernatant was aspirated and the DNA pellet vacuum-dried in a vacuum centrifuge. Pellets were resuspended in 4 µl of formamide loading dye, vortexed for 15 seconds, centrifuged briefly and then heated to 100°C for 2 min before being cooled on ice prior to loading. 2 µl of each experimental sample (1 µl of the control reaction) was electrophoresed through 4% acrylamide gels and run on an ABI 377 automated sequencer (Perkin Elmer Biosystems) according to manufacturers' instructions. Sequence data was printed off in the form of electropherograms and entered into the GCG package of programmes (Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wisconsin.) for analysis.

2.17. Isolation of total RNA from *Fugu* brain tissue.

Total RNA was isolated from *Fugu* brain tissue using TRIZOL® Reagent, a commercially available guanidine isothyocyanate-phenol solution (Gibco BRL). TRIZOL reagent was added to *Fugu* brain tissue (1 ml per 50-100mg of tissue, all volumes quoted here are per 1ml of TRIZOL initially used) which was cut into small pieces using a sterile scalpel and homogenised by passing it through progressively smaller gauged syringe needles, 25 down to 18 gauge. The homogenised sample was transferred into a 1.5 ml eppendorf tube, vortexed briefly and incubated in the TRIZOL reagent for 5 min at room temperature to permit the complete dissociation of nucleoprotein complexes. 0.2 ml of chloroform was added to the sample followed by vortexing for 15 seconds and incubation at room temperature for 2-3 min. After incubation, the sample was centrifuged at 12,000 x g for 15 min at 4°C to separate the RNA into the colourless, upper aqueous phase which was then transferred to a fresh eppendorf tube. 0.5 ml of Isopropyl alcohol (BDH) was mixed with the aqueous phase and incubated at room temperature for 10 min to precipitate the RNA. The sample was centrifuged at 12,000 x *g* for 10 min at 4°C, the supernatant was then removed. The RNA precipitate, often visible as a gel-like pellet, was washed with 1 ml of 75% ethanol and centrifuged at no more than 7,500 x *g* for 5 min at 4°C. The ethanol was aspirated and the RNA pellet air-dried briefly for 5-10 min before being resuspended in 50 µl of RNase-free distilled water. RNase-free water was prepared by adding 0.1% diethyl pyrocarbonate (DEPC) (Sigma) to distilled water for 12 hours followed by autoclaving at 121°C for 15 min. The concentration of RNA was estimated using ultra-violet spectroscopy. A 1 in 100 dilution of the sample was made in DEPC-treated distilled water and the absorbance at

260 nm and 280 nm was measured using quartz cuvettes and a water blank. An absorbance of 1.0 at 260 nm was equal to approximately 40 μg/ml RNA. The RNA was adjudged to be free from contaminants if the absorbance ratio 260 nm/280 nm was greater than 1.8.

2.18. Amplification of *Fugu* cDNA.

Amplification of *Fugu* brain cDNA was performed using the commercially available Marathon™ cDNA amplification kit (Clontech) from which all reagents were obtained unless otherwise stated. 4 μg of total brain RNA was combined with 10 μM cDNA synthesis primer (an oligo d(T) based primer) in a total volume of 5 μl in a sterile 0.5 ml microcentrifuge tube. The contents were mixed, centrifuged briefly and incubated at 72°C for 2 min followed by cooling on ice for 2 min. The tube was centrifuged briefly and the total volume increased to 10 μl with a final concentration of 1 x first-strand buffer, 1 mM dNTP mix, 2 units/μl AMV reverse transcriptase. After mixing, the tube was centrifuged briefly and incubated at 42°C for 1 hour in an air incubator. First-strand synthesis was terminated by incubating the tube on ice. Second-strand synthesis was achieved by mixing the entire first-strand reaction (10 μl) with 1 x second-strand buffer, 0.2 mM dNTP mix, 1 x second-strand enzyme cocktail in a total volume of 80 μl on ice. The tube was centrifuged briefly and incubated at 16°C for 1.5 hours. After incubation, 10 units of T4 DNA polymerase was added, mixed and the tube incubated at 16°C for a further 45 min. Second-strand synthesis was terminated by the addition of 4 μl of the EDTA/Glycogen mix. 100 μl of phenol:chloroform:isoamyl alcohol (25:24:1) (Gibco

BRL) was added to the second-strand reaction and vortexed thoroughly. The tube was centrifuged at 20,000 x *g* for 10 min to separate phases. After centrifugation, the top aqueous layer was transferred to a clean 0.5 ml microcentrifuge tube and 100 µl of chloroform:isoamyl alcohol (24:1) (Gibco BRL) added and vortexed thoroughly. The tube was centrifuged at 20,000 x *g* for 10 min to separate phases and the top aqueous layer transferred to another clean 0.5 ml microcentrifuge tube. One-half volume of 4 M ammonium acetate and 2.5 volumes of room temperature 95% ethanol were added to the top aqueous layer and vortexed thoroughly, the tube was centrifuged at 20,000 x *g* for 20 min at room temperature. The supernatant was carefully removed, the pellet overlaid with 300 µl of 80% ethanol and centrifuged at 20,000 x *g* for 10 min. The supernatant was removed carefully and the pellet allowed to air-dry for 10 min. The pellet was resuspended in 10 µl of sterile distilled water and stored at -20°C. The yield of experimental double stranded cDNA was estimated by analysing 2 µl alongside 2 µl of double stranded cDNA made from positive control human placental RNA on a 1.2% agarose gel as previously described. The Marathon cDNA adaptor was ligated to the double stranded cDNA in a 10 µl reaction volume by combining 5 µl of double stranded cDNA with 2 mM Marathon cDNA adaptor, 1 x DNA ligation buffer and 40 units/µl T4 DNA ligase at room temperature, vortexing and incubating overnight at 16°C. After incubation, the ligase was inactivated by incubating at 70°C for 5 min. The adaptor-ligated double stranded cDNA was diluted in tricine-EDTA buffer, according to the estimation of yield after second-strand synthesis, incubated at 94°C for 2 min, chilled on ice for 2 min, centrifuged briefly and stored at -20°C until ready. Rapid amplification of cDNA ends (RACE) was used to generate 5′ and 3′ cDNA fragments by combining 5 µl

of adaptor-ligated double stranded cDNA with 0.2 mM adaptor primer 1 (AP1), 1 x cDNA PCR reaction buffer (Clontech), 0.2 mM dNTP mix, 1 x Advantage cDNA polymerase mix (Clontech) and 0.2 μM gene specific primer (GSP) (see table 2.5 for a list of gene specific primer used) in a total volume of 50 μl. Thermal cycling was commenced with the following cycling parameters: 94°C x 30 seconds for 1 cycle; 94°C x 5 seconds and 72°C x 4 min for 5 cycles; 94°C x 5 seconds and 70°C x 4 min for 5 cycles; 94°C x 5 seconds and 68°C x 4 min for 25 cycles followed by a 4°C hold. 5 μl of each RACE reaction was analysed on a 1.2% agarose gel, as described previously, alongside control reactions. A second, nested, PCR was carried out for both 5′ and 3′ cDNA fragments using identical conditions and cycling parameters to the first but substituting AP1 and GSPs for adaptor primer 2 (AP2) and nested GSPs. Again, 5 μl was analysed on a 1.2% agarose gel. RACE products were electrophoresed on a 1.2% agarose gel, purified, subcloned, transformed and sequenced as previously described. Sequence data obtained was used to design 5′ and 3′ primers for the amplification of full-length *Fugu* NOS cDNA (*f*NOS) (see table 2.5). 5 μl of adaptor-ligated double stranded cDNA was combined with 2 μM of each of the specific 5′ and 3′ primers, 1 x cDNA polymerase reaction buffer, 0.2 mM dNTP mix and 1 x Advantage cDNA polymerase mix in a total volume of 50 μl. Thermal cycling was commenced with the following cycling parameters: 94°C x 1 min for 1 cycle; 94°C x 30 seconds and 72°C x 6 min for 25 cycles followed by a 4°C hold. 5 μl of each sample was analysed on a 1.2% agarose gel. Gel purification, subcloning into the appropriate vector, transformation and sequencing of the full-length cDNA was carried out as described previously.

Table 2.5. Primers used for the amplification of *f*NOS.

Primer	Sequence (5′→ 3′)	Application
GSP1	CATGAGCTGTTGTGCAGTCTCTGTCATC	5′ RACE
NGSP1	CGTGCAGGTTGGGCATCATAAGGGAGC	Nested 5′ RACE
GSP2	CTCCGACACTCGCAAGTCCTCAAGTGAC	3′ RACE
NGSP2	CTCCGACACTCGCAAGTCCTCAAGTGAC	Nested 3′ RACE
5′NotIfl*	CACTCACAAC**GCGGCCGC**TATGCAAGA GTCCGAGCCTTCCGTG	full-length cDNA amplification
3′NotIfl*	AGGCCCGGAT**GCGGCCGC**CTAGAGCAG AACACCTCATCGGAATC	full-length cDNA amplification
AP1	CCATCCTAATACGACTCACTATAGGGC	5′ and 3′ RACE
AP2	ACTCACTATAGGGCTCGAGCGGC	5′ and 3′ nested RACE

* Primers 5′NotIfl and 3′NotIfl were engineered to incorporate a *Not*I site (indicated by the bold typeface) into the full-length *Fugu* NOS cDNA sequence for subcloning into the plasmid vector pVL1393. Primer 3′NotIfl also contains a STOP codon (indicated by underlining)

2.19. Expression of cloned cDNA in insect cells.

Expression of recombinant cDNA in insect cells was achieved using a baculovirus-insect cell system. Clones of interest were ligated into the *Not*I site of the plasmid pVL1393 (Stratagene), transformed into DH5α *E. coli* and plasmids purified by Qiagen mini-prep as described previously. *Spodoptera frugiperda*-21 (*Sf*21) insect cells, required for the growth of baculovirus, were supplied by Mr. N. Foxwell (The Wolfson Institute for Biomedical research, UCL). *Sf*21 cells were maintained at 28°C as suspension cultures in glass stirrer vessels in complete TC100 (TC100 insect cell medium, 10% fetal calf serum (FCS), 1 x antibiotic/antimycotic and 25 μg/ml gentamicin (Gibco BRL)). For

construction of recombinant baculoviruses, 3 μg of plasmid in 2 μl of distilled water were mixed with 0.5 μg (5 μl) of 'Baculo-Gold' baculovirus DNA (PharMingen) and 43 μl of distilled water. This was mixed with 40 μl of lipofectin (Gibco BRL) (pre-mixed with 10 μl of distilled water) and incubated at room temperature for 10 min. *Sf*21 cells were plated out into a T25 flask (Falcon) to 70/80% confluence and incubated for at least 1 hour at 28°C to allow the cells to settle and adhere to the plastic. The supernatant was then aspirated, the cell monolayer washed twice with serum-free TC100 medium and re-fed with 2 ml of serum-free medium. The DNA-lipofectin mixture was added to the cells in a drop-wise manner, the flask rocked gently to distribute the DNA evenly over the cells and the cells incubated at 28°C for 5 hours. The supernatant was then removed and replaced with 2.5 ml of complete TC100 medium and the incubation continued for a further 3 days. At the end of the incubation period, the supernatant was removed and stored at 4°C wrapped in silver foil, individual recombinant baculovirus were isolated by limiting dilution plaque assay. Six 30 mm petri dishes of *Sf*21 cells, 100% confluence, were set up and washed as described previously. Five log10 dilutions of the 3 day supernatant from the original transfection were made, from neat, in complete TC100 medium. 200 μl of each dilution, and 200 μl of neat supernatant, were added to the cells in each petri dish (after the last wash had been aspirated) and the dishes were gently rocked in order that the entire monolayer was covered. The dishes were incubated for 2 hours at 28°C after which time the monolayer was overlaid with 2 ml of 1% Sea Plaque UltraPure agarose (Gibco BRL), kept liquid at 37°C, (diluted with complete TC100 medium from a 3% stock) and incubated at room temperature for 30 min to allow the agarose to set. 1 ml of complete TC100 medium was added to each dish and the

incubation continued at 28°C for 3 days in a humidified atmosphere. At the end of the incubation period, the overlaid medium was aspirated and the agarose-cell monolayer was stained with 20% v/v neutral red (Sigma) in PBS for 1 hour at room temperature. The stain was decanted and the plates incubated upside down at 4°C for 18 hours. Dishes were then examined for the presence of well separated, light pink plaques (caused by viral infection) against a dark red background. The agarose above individual plaques was removed as a plug, by suction, using 1 mm bore disposable plastic Pasteur pipettes (Alpha) and incubated in 500 µl of complete TC100 medium for 30 min at 37°C. 200 µl of this supernatant was then used to infect a fresh 2 ml culture of *Sf*21 cells prepared as described earlier. This infection was allowed to proceed for 4 days at 28°C after which time the supernatant was removed and 1 ml used to infect a 25 cm^2 5 ml culture of *Sf*21 cells for a further 4 days at 28°C. The infection procedure was repeated using 5 ml of the supernatant from the 25 cm^2 culture to infect a 75 cm^2 20 ml culture of *Sf*21 cells for 4 days. 175 cm^2 50 ml cultures of *Sf*21 cells were then infected for 4 days with 6 ml each of the 75 cm^2 supernatant. Finally, two 175 cm^2 50 ml cultures of *Sf*21 cells were infected with 2.5 ml of supernatant from the previous infection. This final infection yielded 100 ml of culture medium containing high titre recombinant baculovirus (>1 x 10^8 plaque forming units (pfu/ml) and was stored at 4°C. A time-course of infection was carried out using 75 cm^2 20 ml cultures of *Sf*21 cells infected with high titre recombinant baculovirus at a moi of 5. Time points were taken at 0 hours (after cells have been incubated with virus for 2 hours at 28°C), 24, 48, 72 and 96 hours. Non-infected cells were taken at the same time points as a control, 1.5 ml of complete TC100 medium was substituted for the virus. At each time point the cells were collected in PBS and stored at -70°C for analysis by SDS-PAGE and NO synthase activity assay as described below.

2.20. Measurement of nitrite concentrations in culture supernatants.

Nitrite (NO_2^-) concentrations in supernatants were measured by a modification of a chemiluminescent method as previously described (Downes *et al.*, 1976). A known volume of supernatant was injected into a reflux vessel containing 25% v/v 6% sodium iodide in glacial acetic acid at 100°C under nitrogen. Under these conditions NO_2^- (a stable end product of NO synthesis) was converted back to NO, mixed with a stream of ozone gas and the resulting chemiluminescent product quantified using a photomultiplier device (The Wolfson Institute for Biomedical Research, University College London). Sodium nitrite was used to construct standard curves to allow calculation of the NO_2^- concentrations in each sample. Nitrate (NO_3^-) was unaffected in this system, thus this method only allowed the estimation of NO_2^- and not total nitrogen oxides (NO_x) produced by cells. Results were expressed as the concentration of NO_2^- generated/number of cells.

2.21. SDS-PAGE gel electrophoresis of proteins.

SDS-PAGE was performed according to a previously published method (Laemmli, 1970). All gels had an acrylamide to bis-acrylamide ration of 37.5:1. Resolving gels were made at 8% in 190 mM Tris HCl, pH 8.8, 0.1% w/v SDS, 0.1% ammonium persulphate (APS) (BDH) and 0.06% v/v *N, N, N', N'*-tetramethylethylenediamine (TEMED) (Sigma). 5% stacking gels were similarly prepared using 126 mM Tris HCl, pH 6.8 and 0.1% TEMED. All gels were run using a standard electrophoresis buffer containing 25 mM Tris HCl, 0.25 M glycine (BDH) and 0.1% w/v SDS. Samples to be

electrophoresed were mixed with 20% v/v of 5 x concentrated sample buffer (2.5 M Tris HCl, pH 6.8, 0.5 M DTT, 10% w/v SDS, 1% bromophenol blue and 50% v/v glycerol (BDH)), heated to 100°C for 5 min and cooled on ice prior to loading on gels. 'Rainbow' molecular weight standards (range 14 to 200 kDa) (Amersham) were similarly treated prior to electrophoresis and run on all gels. Electrophoresis was carried out at a constant current of 5 mA/cm gel length until the bromophenol blue marker reached the bottom of the gel. At the end of electrophoresis, gels were either silver stained or analysed by western blot according to the methods below.

2.22. Silver staining of SDS-PAGE gels.

Gels were silver stained using a commercially available kit (Pharmacia Biotech) from which all reagents were obtained unless otherwise stated. Following electrophoresis, gels were fixed in 40% ethanol, 10% glacial acetic acid (BDH) for 30 min. Gels were then incubated for 30 min, with shaking, in sensitizing solution (30% ethanol, 0.125% w/v glutardialdehyde, 0.2% w/v sodium thiosulphate, 6.8% w/v sodium acetate) followed by three washes of 5 min each in distilled water. Silver solution (0.25% w/v silver nitrate, 0.0148% w/v formaldehyde) was then added to the gel which was left shaking for 20 min. Two 1 min washes in distilled water were followed by the addition of developing solution (2.5% w/v sodium carbonate, 0.0148% w/v formaldehyde) to gels and incubation with shaking for 2-5 min until bands were sufficiently developed. The development of gels was stopped by incubating gels in stop solution (1.46% w/v EDTA-$Na_2.2H_2O$) for 10 min. Gels were washed in distilled water and 10% glycerol added as a preservative at least 20 min prior to drying gels overnight in gel drying film (Promega).

2.23. Western blot analysis of SDS-PAGE separated proteins.

Four pieces of 3MM filter paper were cut to the exact size of gels selected for western blot analysis and soaked in transfer buffer (25 mM tris HCl, 192 mM glycine, 20% methanol). Two of the filter papers were placed on a pre-wetted sponge in the cassette of a wet electroblotter (Biorad). A sheet of Hybond^{TM} ECL^{TM} nitrocellulose membrane (Amersham), cut to gel size, was pre-wetted with distilled water and transfer buffer and placed on top of the filter papers. Gels, also soaked in distilled water and transfer buffer, were placed on top of the nitrocellulose membrane followed by the remaining two sheets of filter paper. Any air bubbles were removed with a glass rod lightly rolled over the surface of the filter papers. Another pre-wetted sponge was placed on top of the filter papers, the cassette closed and the blotting apparatus assembled. Proteins were transferred to the nitrocellulose membrane by applying a constant current of 1 mA/cm^2 of gel area for at least 90 min. Transfer was verified by assessing the movement of pre-stained molecular weight markers from gel to membrane. At the end of the transfer, the gel and filter papers were discarded and the nitrocellulose was incubated in a solution of 5% non-fat milk in PBS for at least 30 min to block non-specific protein binding sites on the blot. After blocking, blots were washed three times in PBS for 5 min each. Primary antibodies (see table 2.6 for antibodies used in these studies) were diluted with 5% non-fat milk solution in PBS, according to manufacturers instructions, and applied to the blot overnight at room temperature.

Table 2.6. Antibodies used in these studies.

Antibody	Immunogen	Host	Source	Conjugate	Dilution	Use
anti-human nNOS monoclonal	1095-1289	Mouse	Transduction Laboratories	-	1:2500 1:250	WB* IS*
anti-mouse iNOS monoclonal	961-1144	Mouse	Transduction Laboratories	-	1:2500	WB
anti-human eNOS monoclonal	1030-1209	Mouse	Transduction Laboratories	-	1:2500	WB
anti-human nNOS polyclonal	1095-1289	Rabbit	Transduction Laboratories	-	1:500	WB IS
anti-mouse iNOS polyclonal	961-1144	Rabbit	Transduction Laboratories	-	1:5000 1:500	WB IS
anti-human eNOS polyclonal	1030-1209	Rabbit	Transduction Laboratories	-	1:1000 1:500	WB IS
anti-mouse IgG	-	Goat	Sigma	HRPO*	1:5000	WB
anti-rabbit IgG	-	Goat	Sigma	HRPO	1:40,000	WB
anti-mouse IgG	-	Goat	Molecular Probes	Alexa488*	1:250	IS

* where WB = western blot, IS = Immuno Staining, HRPO = horseradish peroxidase, Alexa 488 = a green fluorophore which is used in place of fluorescein due to its increased brightness and photostability.

Following incubation with primary antibodies, blots were washed three times in PBS, for 5 min each wash, and then incubated with the appropriate horseradish peroxidase conjugated secondary antibody diluted in 5% non-fat milk in PBS. After at least one hour incubation at room temperature, the secondary antibody was decanted and blots washed three times in PBS for 5 min each wash. Bands of immunoreactive protein were visualised using SuperSignal® substrate for western blotting (stable peroxide solution and luminol/enhancer solution mixed in a 1:1 ratio) (Pierce). Blots were incubated in substrate for 2 min, excess liquid drained and then covered with Saran wrap. Blots were exposed to Hyperfilm™ ECL™, a high performance chemiluminescence film, (Amersham) for between 10 seconds and 5 min after which time the film was developed.

2.24 Assay for nitric oxide synthase activity.

Assays for NOS activity were based upon a modification of a spectrophotometric method (Feelisch and Noack, 1987). Cytosols were prepared from 500 cm^2 cell monolayers that had been infected with high titer virus at a moi of 5, fed with arginine-free TC100 medium and dialysed fetal bovine serum and incubated for 24 hours at 28°C. Cells were collected in approximately 5 volumes of homogenisation buffer (0.1 M Tris HCl, pH 7.4 (BDH) containing 1 mM dithiothreitol (DTT) (Sigma), 0.1 mM phenyl methyl sulphonyl fluoride (PMSF) (Sigma), 0.5 μM leupeptin (Sigma) and 0.5 μM pepstatin A (Sigma)) and then lysed by sonication at 22 micron amplitudes in three 5 second bursts with 25 seconds cooling on ice in between. Lysates were centrifuged at 105,000 x *g* for 30 min at 4°C to remove cell debris and any residual arginine removed using activated Dowex AG-50W ion exchange resin (Bio-Rad). The resin was prepared by packing it into a column and adding 2N sodium hydroxide (NaOH) until the pH of the run-through reached 14. The resin was rinsed with distilled water until it reached pH 7 and then homogenisation buffer added to the column and left to drain through, the resin was not allowed to dehydrate. The cytosol was incubated with the activated resin for 5 min on ice, vortexing every 30 seconds, followed by centrifugation at 20,000 x *g* for 5 min at 4°C. The supernatant was transferred to a fresh eppendorf tube containing resin and the 5 min incubation on ice and centrifugation repeated. The supernatant was then transferred to a fresh eppendorf and either assayed immediately for NO synthase activity or stored at -80°C. Enzyme assays were performed in quartz cuvettes by incubating a known volume of cell cytosol with 0.1 M HEPES/0.1 mM DTT buffer pH7.5, 100 μM reduced nicotinamide adenine dinucleotide phosphate (NADPH) (Sigma), 5 μM

oxyhaemoglobin (a gift from Neale Foxwell, The Wolfson Institute for Biomedical Research, University College London), 4 μM flavin adenine dinucleotide (FAD) (Sigma) and 5 μM tetrahydrobiopterin (H_4B) in a total volume of 600 μl. Nitric oxide synthesis was initiated by the addition of 30 μM L-arginine (Sigma) and measured using a dual wavelength spectrophotometer (Shimadzu Corporation). The shift in absorbance between the wavelength pair 401-421 nM caused by the conversion of oxyhaemoglobin (oxyHb) to methaemoglobin (metHb) by nitric oxide was measured as a function of time using the equation:

$$\frac{\Delta Abs/min \times 0.02}{100 \times 77200} \times 0.0006$$

where: ΔAbs/min = the change in absorbance in 1 min (% of 100)
0.02 = full scale deflection (i.e. 0-0.02 = 100%)
0.0006 = reaction volume in litres
77200 = extinction co-efficient for haemoglobin in the spectrophotometer used

Results were expressed as moles of NO formed/sample/volume/min or moles of NO formed/mg protein/min.

In some assays the concentrations of L-arginine were varied, while in others the effects of different concentrations of the NOS inhibitors L-N-guanidino monomethyl arginine (L-NMMA) (Boehringer Mannheim), L-Thiocitrulline (CalBiochem) and 1400W (supplied by Mr N. Foxwell, W.I.B.R., UCL) were investigated. Each of the inhibitors was dissolved in 0.1 M HEPES/0.1 mM DTT buffer pH7.5 and 6 μl of each concentration added to the reaction mix prior to the addition of L-arginine. The calmodulin requirements of the enzyme were tested by the addition of 1 mM EGTA to the reaction mix prior to the addition of L-arginine.

2.25. Determination of protein concentration of cell cytosols.

A commercially available kit (Bio-Rad), from which all components were obtained unless otherwise stated, was used to measure the concentration of protein in cell cytosols based on a previously published method (Bradford, 1976). The microassay procedure was followed and dilutions were made of the cell cytosol and a protein standard (4, 8, 16, 20 and 25 μg of protein), 0.2 ml of dye reagent concentrate was added to each dilution, mixed and incubated for 5 min. The OD_{595} was measured versus a reagent blank and the dilutions of the protein standard used to construct a standard curve. The plate reader (Bio-Rad) used to take all measurements then calculated the concentration of the unknowns, in μg, from the standard curve.

2.26. Immunocytochemical staining of *F*NOS in insect cells.

A time-course of *Fugu* NOS infection of insect cells was set up using *Sf*21 cells seeded to 50% confluence on 13 mm coverslips (BDH) and left to adhere at 28°C for 2 hours. Coverslips were prepared by soaking in 70% industrial methylated spirits (IMS) (BDH) and then allowed to air-dry, one coverslip was placed into each of the wells of a 24 well plate (Falcon). 250 μl of high titre recombinant baculovirus-insect cells was added to each well (250μl of complete TC100 medium was used for the control coverslips) and incubated at 28°C for 2 hours. 1ml of complete TC100 medium was added to each well and incubated at 28°C for 0, 24, 48, 72 and 96 hours. When each time point was reached coverslips were removed from their wells, washed twice for 3 min each wash in

PBS and cells fixed in 4% paraformaldehyde (BDH) with gentle shaking for 1 hour at room temperature. Coverslips were then washed once and kept in PBS at 4°C until ready to stain. Cells were permeabilised by incubation in 0.5% Triton X-100 (BDH) in PBS for 10 min and then washed three times for 3 min each time in PBS. The cells were then incubated for 5 min in complete TC100 medium to block any non-specific interactions, this was followed by a 1 hour incubation in primary antibody suitably diluted in complete TC100 medium (see table 2.5). After incubation, the cells were washed three times for 3 min each time in PBS and incubated for 1 hour in an appropriate secondary antibody (see table 2.5), again diluted in complete TC100 medium. The cells were washed three times for 3 min each wash in PBS and then incubated in 170 μg/ml RNase at 37°C for 15 min, this was followed by a further 3 min wash in PBS. Propidium iodide was added to each coverslip and incubated at room temperature for 5 min. After incubation, the coverslips were mounted upside down in 3 μl of Vectashield mountant on a microscope slide, allowed to dry for 20 min at room temperature before using clear nail varnish to seal the edges of the coverslips. All incubations and washes were carried out at room temperature and with gentle shaking except for the RNase and propidium iodide incubations. Images were viewed and captured on a Leica TCS SP laser scanning confocal microscope (Leica).

2.27. Ligand affinity chromatography of *F*NOS on 2′5′ ADP sepharose.

Fugu nitric oxide synthase (*F*NOS) was partially purified from cell lysates by affinity chromatography on 2′5′ adenosine di-phosphate (ADP) sepharose according to a

modification of a previously published method (Steuhr *et al.*, 1991). *Sf*21 Cells from six 175cm^2 cultures, infected for 24 hours with 1.5 ml high titre recombinant baculovirus, were collected in 4ml of lysis buffer (20 mM Tris HCl, pH 7.5, 1 mM DTT, 100 μg/ml PMSF, 10 μg/ml Leupeptin (Sigma), 10 μg/ml Pepstatin (Sigma), 2 μM FAD, 2 μM H_4B, 2 mM L-arginine, 1 mM EDTA, 10% v/v glycerol) and sonicated at 22 amplitude microns in 3 x 5 second bursts, with 25 seconds cooling on ice in between each burst. The sonicated cells were centrifuged at 20,000 x *g* for 15 min at 4°C. The supernatant was transferred to a fresh tube and glycerol added to 10% v/v. 0.5 g of 2′5′ ADP sepharose (Pharmacia) was swollen for 30 min in 50 ml of PBS and equilibriated in 2 changes of 10 ml each of lysis buffer. The supernatant was mixed with the sepharose by gentle stirring at 4°C for 30 min and poured into a disposable 10 ml chromatography column (Bio-Rad). All the following procedures were carried out at 4°C. The gel was allowed to settle under gravity to form a column of approximately 2 ml in volume and then washed in 2 x 5 column volumes of lysis buffer. Non-specifically bound proteins were removed from the column by washing in 5 column volumes of 0.5 M NaCl in lysis buffer and the column was washed again in 5 column volumes of lysis buffer. Specifically bound proteins were eluted from the column by washing in 3 column volumes of 10 mM NADPH in lysis buffer followed by a final wash in 5 column volumes of lysis buffer. 1 ml fractions were collected throughout and were analysed by SDS-PAGE and western blot analysis.

2.28. Ligand affinity chromatography of *F*NOS on calmodulin sepharose.

Following partial purification on 2′5′ ADP sepharose, ligand affinity chromatography on camodulin (CaM) sepharose (Pharmacia Biotech) was performed to purify *F*NOS even further. NOS-positive eluates from the 2′5′ ADP sepharose column were adjusted to 2 mM $CaCl_2$ and 0.1 M NaCl. CaM sepharose was pre-equilibriated in 10 mM Tris, 2 mM $CaCl_2$, 0.1 M NaCl and 10% glycerol, centrifuged at 3,000 rpm for 15 min and the supernatant removed. The CaM sepharose was mixed with the NOS-positive 2′5′ ADP sepharose eluates by gentle stirring at 4°C for 30 min and poured into a disposable 10 ml chromatography column (Bio-Rad). All the following procedures were performed at 4°C. The gel was allowed to settle under gravity to form a column of approximately 1 ml in volume and then washed with 2 x 2 column volumes of buffer (10 mM Tris, 2 mM $CaCl_2$, 0.1 M NaCl and 10% glycerol). A further 10 column volumes of buffer were used to wash the column. To elute specifically bound proteins from the column, 3 column volumes of buffer, minus $CaCl_2$, plus 3 mM EGTA were added to the column and one 0.75 ml fraction collected. Following a 10 min incubation further fractions were collected. 1 column volume of buffer was added to the column and allowed to drain through. 3 column volumes of buffer, minus $CaCl_2$, plus 10 mM EGTA were added to the column and 1 0.75 ml fraction collected. Following a 10 min incubation further fractions were collected. 1 column volume of buffer was added to the column and allowed to drain through. Further elution of the CaM sepharose column was performed first with 3 column volumes of 1 M NaCl in buffer and second with 2 column volumes of 6 M Urea in buffer. 0.75 ml fractions were collected throughout and were analysed by SDS-PAGE and western blot analysis.

CHAPTER 3

MOLECULAR CHARACTERISATION OF A *FUGU* NOS1 GENE.

CONTENTS.

3.1. Introduction.

Nitric oxide (NO) is a potent regulator and mediator of a broad range of physiological and pathophysiological functions that include hypertension, atherosclerosis, stroke and sepsis. NO is synthesized, along with the co-product L-citrulline, from L-arginine by a family of nitric oxide synthase (NOS) enzymes. In mammals, three distinct NOS isoforms have been characterised, neuronal (nNOS; NOS1), inducible (iNOS; NOS2) and endothelial (eNOS; NOS3), named after the tissues in which they were first identified (neurons, cytokine-induced immune cells and vascular endothelium respectively). The genes encoding the three isoforms are localised on different chromosomes. Human NOS1 spans over 160 kb and is localised on chromosome 12 (Marsden *et al.*, 1993; Xu *et al.*, 1993; Kishimoto *et al.*, 1994). The human NOS2 and NOS3 genes are both much smaller than the NOS1 gene and are localised on chromosomes 17 and 7 respectively (Marsden *et al.*, 1993; Chartrain *et al.*, 1994; Marsden *et al.*, 1994; Robinson *et al.*, 1994; Xu *et al.*, 1994). The genes show a high degree of similarity in their genomic organisation and intron/exon structure, this includes the location of splice junctions and exon sizes. Each isoform has consensus binding sites for calmodulin, NADPH, FAD and FMN and has an overall structure that consists of a COOH-terminal reductase domain and an NH_2-terminal oxidative domain. It has been suggested that the high degree of conservation seen between all three isoforms could reflect the duplication and chromosomal translocation of a primitive NOS gene.

At approximately 400 Mb, the genome of the Japenese pufferfish *Fugu rubripes* (*Fugu*) is very compact, around 7.5 times smaller than the human genome. It has been estimated that the gene repertoire of *Fugu* is the same as humans and that the genome averages one

gene every 6-7 kb. The use of *Fugu* as a model genome is based on its compressed genome size which is now known to be due to a reduction in the amount of intronic and intergenic DNA, pseudogenes and dispersed repeats. *Fugu* is an ideal model for the efficient analysis of gene structure and the identification of conserved protein domains and regulatory elements.

The *Fugu* genomic cosmid library used in these studies was constructed by Elgar and Nizetic (Elgar *et al.*, 1999) and was obtained from the MRC HGMP Resource Centre (http://fugu.hgmp.mrc.ac.uk). To create the library, *Fugu* genomic DNA was partially digested with *Mbo*I and cloned into the *Bam*HI restriction endonuclease site of the cosmid vector Lawrist 4 (Figure 3.1). The library was supplied as high-density gridded filters on which cosmid clones were spotted in duplicate. Each cosmid clone of the genomic library was spotted on the filters in such a way that a hybridisation pattern, specific to each clone, would be produced when a probe hybridised to the clone. The clones were spotted on the filters and grown at 37°C overnight on LB agar plates containing 30µg/ml Kanamycin. The filters were treated with SDS, denatured, neutralised and then dried and crosslinked ready for pre-hybridisation. The *Fugu* genomic library consists of around 76,000 cosmid clones, each with an average insert size of 40 kb, this represents ~8 times coverage of the *Fugu* genome (Yeo *et al.*, 1997).

3.2. Identification of a putative NOS-like sequence in cosmid 064O07.

The *Fugu* genomic cosmid library was screened, as described in section 2.3, with two different [α-^{32}P] dCTP labelled probes. Fragments of full-length human iNOS cDNA (4 kb) (Accession number g292241 in the EMBL data library) and 5′ human iNOS cDNA

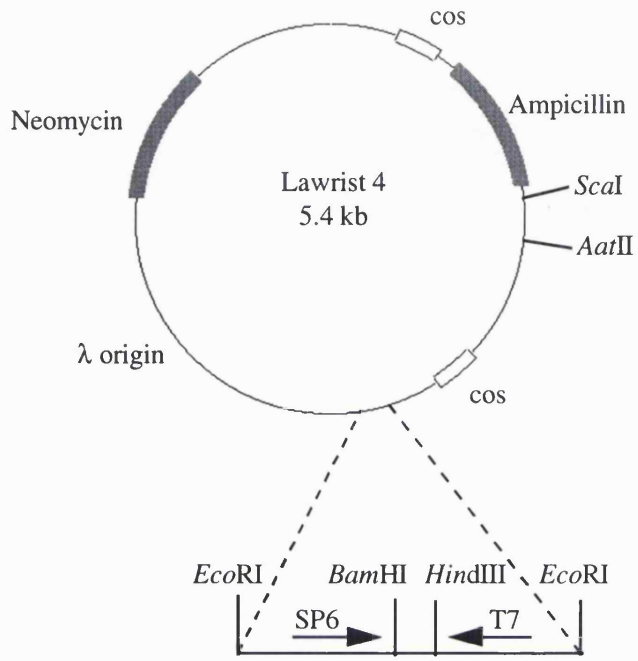
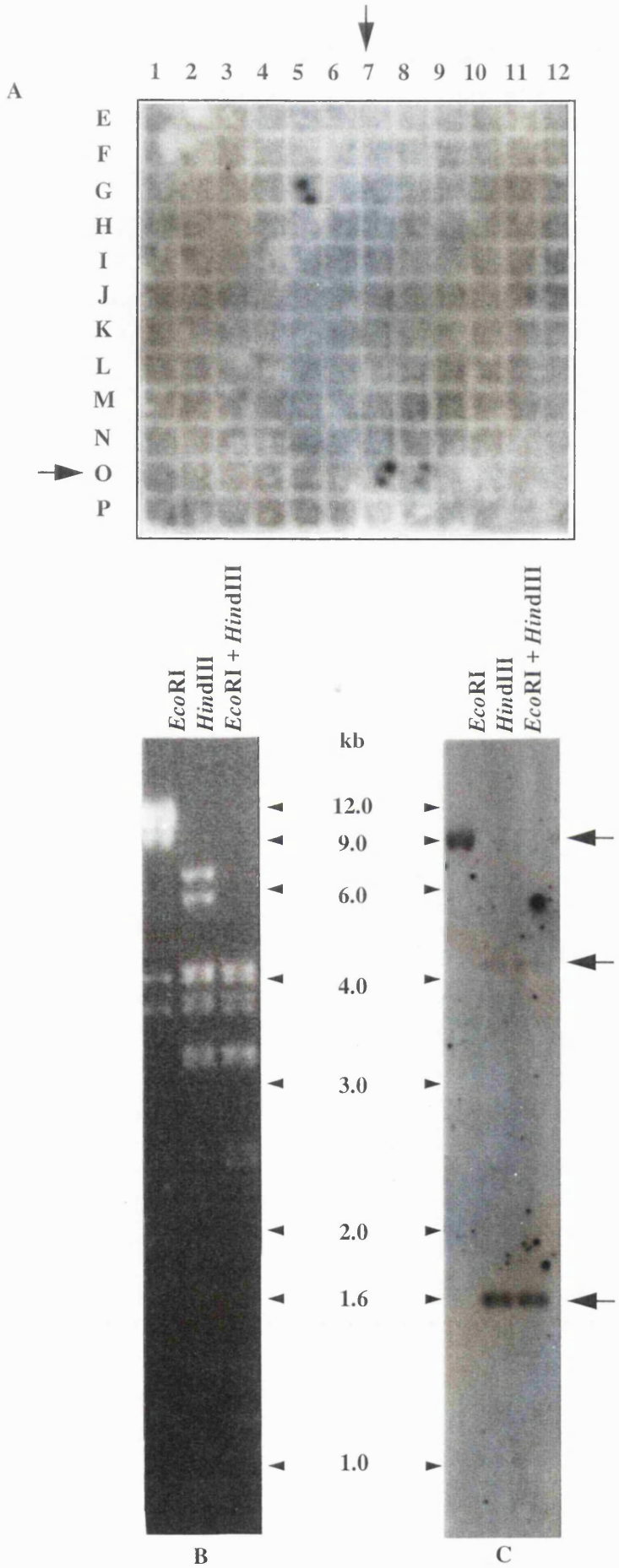


Figure 3.1. Cosmid map of the vector Lawrist 4. *Fugu* genomic DNA was partially digested with the restriction endonuclease *Mbo*I and cloned into the *Bam*HI site of Lawrist 4 to create the *Fugu* genomic cosmid library used in these studies. The ampicillin and neomycin resistance genes are shown as solid boxes. SP6 and T7 primer sites are indicated with arrows.

(1.5 kb), corresponding to nucleotides 229 - 1728 (Charles *et al.* 1993), were prepared using the restriction endonucleases *Eco*RI and *Nde*I (section 2.6) The inserts were gel isolated (see sections 2.7 and 2.9), labelled and purified (section 2.1) for use as probes. Cosmid clones with putative NOS-like sequences were identified by the hybridisation pattern they gave and the filter panel they were on. One cosmid clone which produced a positive signal was cosmid 064O07, as shown in Figure 3.2A. The co-ordinates of those cosmids giving a positive hybridisation pattern were used to obtain the relevant clones from the MRC HGMP Resource Centre. Cosmid DNA was prepared (section 2.4) and restriction endonuclease digestion performed (section 2.6) on twelve cosmid clones that hybridised to the NOS probes. As shown in figure 3.2B and C, Southern hybridisation of cosmid 064O07 digested DNA (section 2.8), with a DIG-[dUTP] labelled 5′ human iNOS cDNA probe (section 2.2), identified a positively hybridising 9.6 kb *Eco*RI fragment, smaller hybridising fragments were also seen for the *Hin*dIII and *Eco*RI plus *Hin*dIII digested DNA. DNA from cosmid 064O07 was digested with *Eco*RI (section 2.6), ligated into the plasmid vector pBsIISK$^+$ (section 2.10) and transformed into subcloning efficiency DH5α cells (section 2.12) to produce plasmid pBs9.6. Figure 3.3 shows a map of pBs9.6 and the cloning strategy used to create it. Colonies containing inserts of 9.6 kb were initially identified by blue-white screening, this insert was verified by restriction endonuclease digestion of plasmid DNA (section 2.6) with *Eco*RI. As illustrated in figure 3.4, pBs9.6 was digested with various restriction endonucleases to identify fragments suitable for subcloning as part of the sequencing strategy. Further subclones of pBs9.6 were created by subcloning *Hin*dIII, *Xba*I, *Bam*HI and *Xba*I plus *Bam*HI fragments into appropriately digested pBsIISK$^+$ plasmid DNA as described previously (sections 2.6, 2.10 and 2.12).

Figure 3.2. Identification of a NOS-like sequence in Cosmid O64O07. (A), A *Fugu* genomic cosmid library was screened with a human iNOS cDNA probe, part of filter panel 2 is shown. Arrows indicate the positively hybridising cosmid 064O07, identified by filter panel number, hybridisation pattern and co-ordinates. (B), DNA from cosmid 064O07 was digested with *Eco*RI, *Hin*dIII and *Eco*RI + *Hin*dIII, separated through an ethidium bromide-stained 1% agarose gel and photographed under U.V. light. (C), DNA was transferred to a nylon hybridisation transfer membrane and probed with a DIG-labelled human iNOS cDNA probe. Arrows indicate positive hybridisation signals. The positions of molecular size markers are indicated.


A
1 2 3 4 5 6 7 8 9 10 11 12
E
F
G
H
I
J
K
L
M
N
O
P
EcoRI
HindIII
EcoRI + HindIII
kb
12.0
9.0
6.0
4.0
3.0
2.0
1.6
1.0
B
C

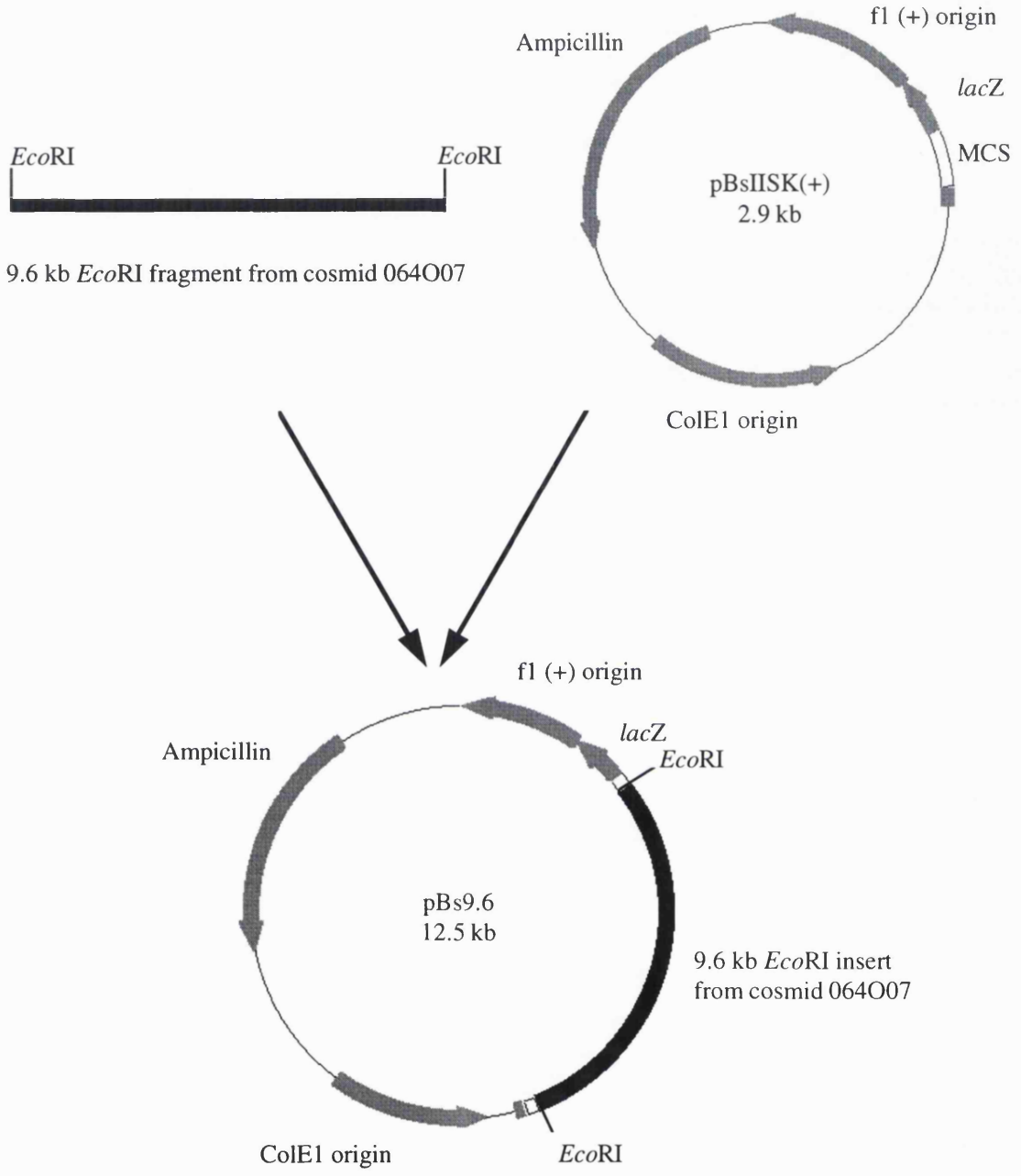


Figure 3.3. Cloning strategy used to create pBs9.6. A 9.6 kb *Eco*RI fragment from cosmid 064O07 was cloned into the *Eco*RI site of the plasmid vector pBsIISK(+) to create pBs9.6.

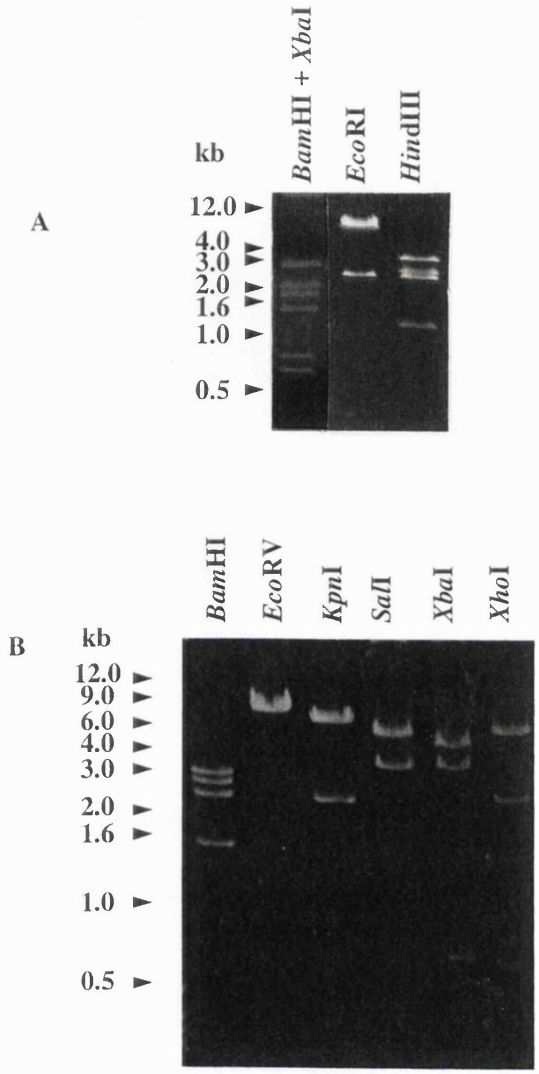


Figure 3.4. Restriction endonuclease digestion of pBs9.6. DNA was digested with restriction endonucleases, separated through ethidium bromide-stained 1% agarose gels and photographed under U.V. light. Fragments suitable for subcloning were identified. (A), restriction endonucleases used were *Bam*HI + *Xba*I, *Eco*RI and *Hin*dIII. (B), restriction endonucleases used were *Bam*HI, *Eco*RV, *Kpn*I, *Sal*I, *Xba*I and *Xho*I. Molecular size markers are indicated.

3.3. Sequence analysis of pBs9.6 and its subclones.

Sequencing of pBs9.6, and its subclones, was initially performed using T7, T3, M13 forward and reverse primers (section 2.16). Sequence specific primers (appendix, table A.1) were designed, from the sequence data obtained, to complete and to confirm the sequence on both strands. The total length of the cloned *Eco*RI fragment was 9,622 nucleotides and BLASTN analysis (Altschul *et al.*, 1990) confirmed the presence of NOS-like sequences. Sequence scanning of cosmid 064O07, performed at the HGMP resource centre, also confirmed the identity of the NOS-like sequences. All sequence analysis was performed using programmes within the Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wisconsin. Greatest sequence similarity was seen with the human neuronal NOS isoform (NOS1), BLASTN analysis identified 9 fragments with 79-90% amino acid identity to human NOS1, 3 with 80-90% identity to human iNOS and 5 with 81-88% identity to human eNOS, exons present within the cosmid were identified as being 5 to 19. Figure 3.5 shows a schematic representation of the 9.6 kb fragment of *Fugu* NOS1, its exons, introns and the positions and lengths of the subclones created as part of the sequencing strategy. Also shown in figure 3.5 is the presence of Lawrist 4 vector sequence which was found at the 3′ end of the 9.6 kb fragment, this confirms that cosmid 064O07 only contains *Fugu* NOS1 sequence up to exon 19. Human NOS1 has 29 exons and 28 introns (Hall *et al.*, 1994), with an average cosmid insert size of 40 kb it would be expected that any upstream exons of *Fugu* NOS1 would be present in cosmid 064O07. Further restriction fragments of cosmid 064O07 were subcloned in an effort to obtain the remainder of the upstream *Fugu* NOS1 but sequence analysis failed to identify any further fragments with obvious NOS sequence similarity. However, sequencing of several of the restriction fragments identified a *Fugu* homologue of the kinase supressor of RAS (*ksr*) gene.

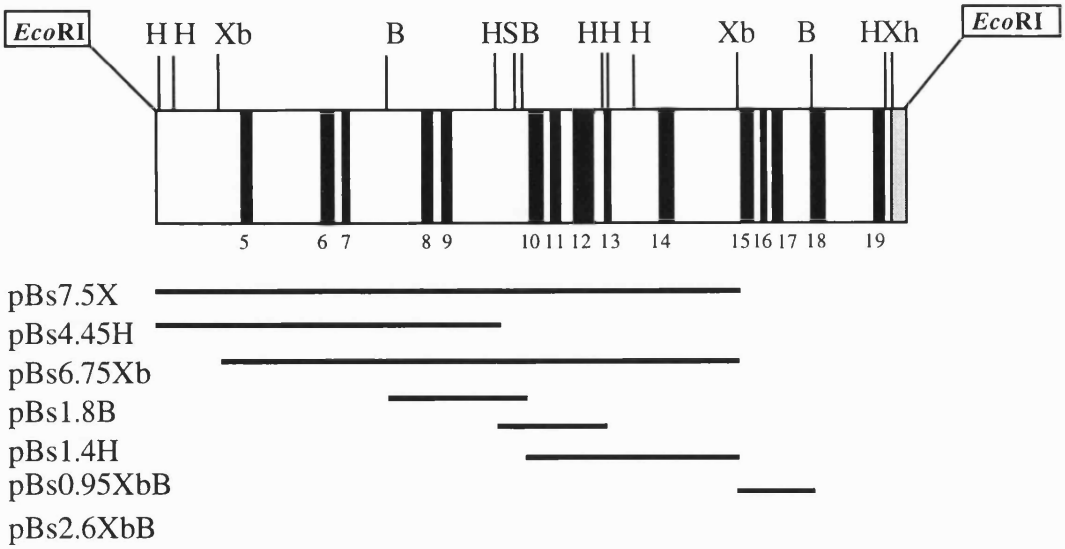


Figure 3.5. Schematic representation of the 9.6 kb *Fugu* NOS1 fragment. A 9.6 kb *Eco*RI fragment of *Fugu* genomic DNA, that positively hybridised to a 5′ human iNOS cDNA probe, was cloned into the *Eco*RI restriction endonuclease site of the plasmid vector pBsIISK$^+$. Restriction endonuclease sites present in the 9.6 kb *Fugu* NOS1 fragment are represented by capital letters and are as follows: H = *Hin*dIII, B = *Bam*HI, S = *Sal*I, Xb = *Xba*I and Xh = *Xho*I. The black boxes represent the exons of *Fugu* NOS1 identified by BLAST analysis and are numbered according to the exons of the human NOS1 gene. The grey box indicates Lawrist 4 vector sequence located 3′ of *Fugu* NOS1 exon 19. Solid black lines represent the subclones of the 9.6 kb *Eco*RI fragment that were created. pBs indicates the pBsIISK$^+$ vector backbone, numbers denote the size of the subcloned fragments in kb and letters represent the restriction endonucleases used to create them.

3.4 Amplification of *Fugu* NOS1 exons 1 to 5 and exons 19 to 29

Fugu genomic DNA obtained from the MRC HGMP Resource Centre (http://fugu.hgmp.mrc.ac.uk) and DNA from cosmid 064O07 were both utilised to complete the *Fugu* NOS1 sequence. A PCR-based approach using sequence specific primers amplified the remainder of the *Fugu* NOS1 gene in overlapping fragments from exon to exon (section 2.15). Figure 3.6 shows that PCR yielded single bands for all except two of the reactions. For exon 28 to 29 a major product was seen at ~1.2 kb, this fragment was selected for cloning and sequencing. PCR amplification of exon 19 to 20 generated multiple bands, three major bands were selected for cloning and sequencing, these fragments had sizes of ~500 bp, 900 bp and 4 kb. To amplify exons 2, 3 and 4 sequence specific primers were designed corresponding to exons 2 and 7, PCR with these primers resulted in the amplification of exons 5 and 6, the sequences of which were already present in cosmid 064O07. The amplified fragments were gel purified, ligated into pCR®2.1-TOPO vector (section 2.11) and transformed into TOP10 competent *E.coli* (section 2.13) to produce the following plasmids: pCRe1-e2; pCRe2-e7; pCRe19-20a,b and c; pCRe20-21; pCRe21-22; pCRe22-23; pCRe23-24; pCRe24-25; pCRe25-26, pCRe26-27, pCRe27-28 and pCRe28-29, where e represents the exon amplified. The cloning of pCRe19-20a-c is illustrated in figure 3.7. Colonies containing inserts of the correct size were identified by blue-white screening and colony PCR (section 2.14). Initial sequencing was performed using T7, T3, M13 forward and reverse primers, whose binding sites were present in the pCR®2.1-TOPO vector sequence, (section 2.16) followed by, where necessary, sequence specific primers (appendix, table A.1). PCR of exons 1 to 5 and exons 19 to 29 was performed with a proofreading enzyme

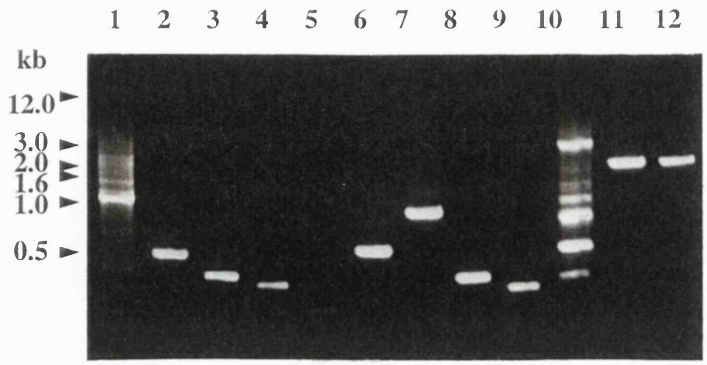


Figure 3.6. PCR amplification of the *Fugu* NOS1 gene exons 1 to 5 and 19 to 29. PCR was performed on *Fugu* genomic DNA and DNA from cosmid O64O07 to amplify the remainder of the *Fugu* NOS1 gene in overlapping fragments. Primer pairs used were as follows: 1 = Ge28f + fugnos29 (for exon 28 to 29), 2 = Ge27f + e28r (for exon 27 to 28), 3 = Ge26f + Ge27r (for exon 26 to 27), 4 = Ge25f + Ge26r (for exon 25 to 26), 5 = Ge24f + Ge25r (for exon 24 to 25), 6 = Ge23f + Ge24r (for exon 23 to 24), 7 = Ge22f + Ge23r (for exon 22 to 23), 8 = Ge21f + Ge22r (for exon 21 to 22), 9 = Ge20f + Ge21r (for exon 20 to 21), 10 = Ge19f + Ge20r (for exon 19 to 20), 11 = Ge1f1 + 5'RACE1 (for exon 1 to 2), and 12 = Ge1f2 + 5'RACE1 (for exon 1 to 2). 1/10th of the PCR reaction volume was separated through an ethidium bromide-stained 1% agarose gel and photographed under U.V. light. Molecular size markers are indicated. A single PCR product was seen for all except two reactions (1 and 10). See text for a more detailed explanation.

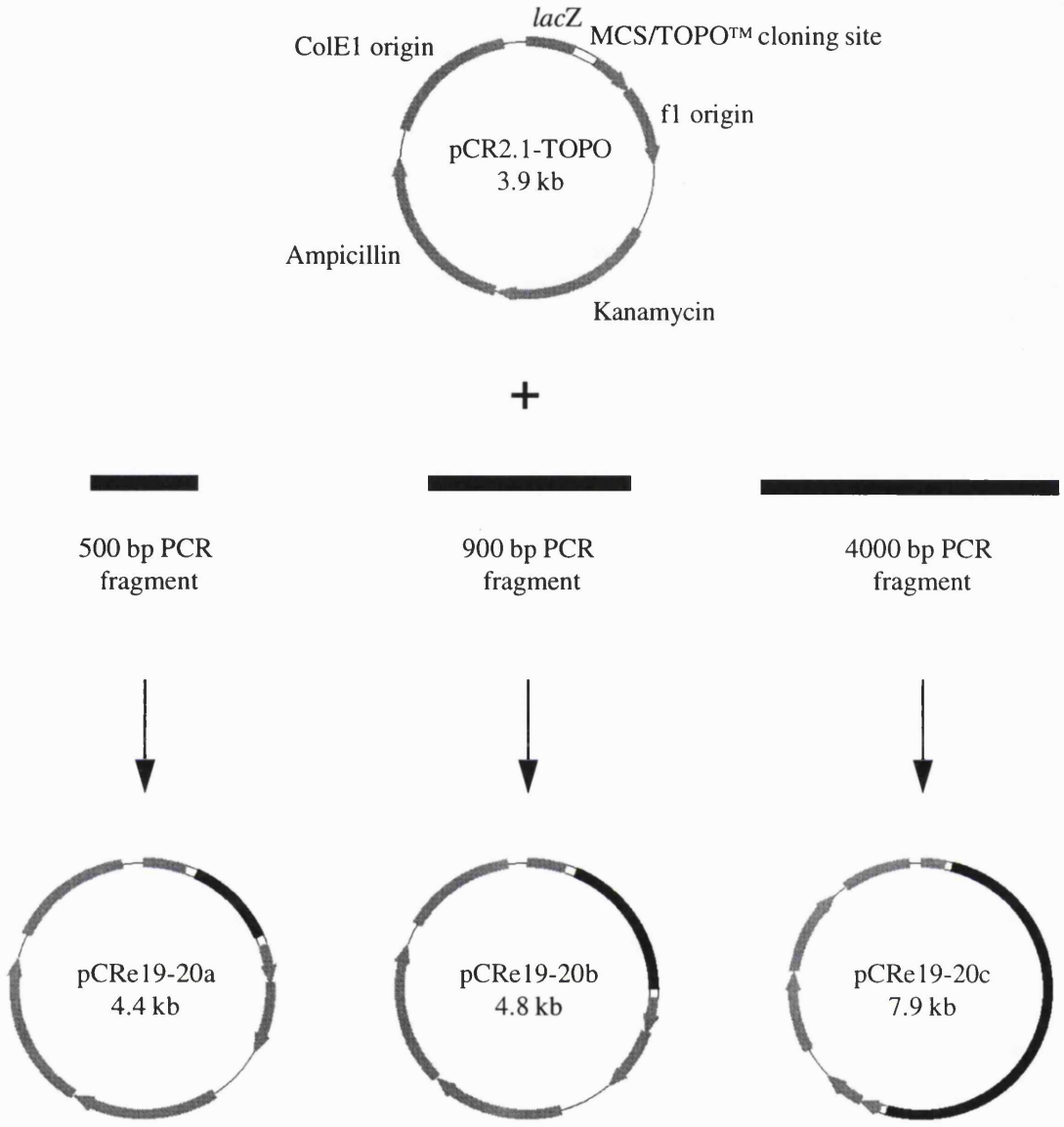


Figure 3.7. Cloning strategy used to create pCRe19-20a-c. Three major PCR products, of 500 bp (a), 900 bp (b) and 4000 bp (c) produced by amplification with primers designed to exons 19 and 20 of *Fugu* NOS1, were TA-cloned into the TOPO™ cloning site of the plasmid vector pCR®2.1-TOPO to create pCRe19-20a, pCRe19-20b and pCRe19-20c.

(section 2.15) and sequencing of multiple isolates was performed to confirm the absence of PCR errors. Sequence analysis was performed using BLASTN analysis (Altschul *et al.*, 1990) and various programmes within the Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wisconsin. Comparison with human NOS1 demonstrated that the missing exons, 1 to 5 and 19 to 29 had been isolated. Plasmid pCRe19-20a, containing the amplified fragment of ~500 nucleotides, was found to contain exon 19 to 20. Using the sequence data from the overlapping fragments a contiguous sequence was constructed for exons 1 to 7 and 19 to 29. Exons 1 to 7 spans approximately 10 kb whereas exons 19 to 29 span only around 4 kb. Figure 3.8 shows a schematic representation of exons 1 to 7 and 19 to 29 and the relative positions of the exons, introns and the primers used for their amplification.

3.5 Sequence analysis of the complete *Fugu* NOS1 gene

A contiguous sequence was constructed for the entire *Fugu* NOS1 gene using the sequence data obtained from pBs9.6, its subclones and the PCR amplified fragments, a schematic representation is shown in figure 3.9. The complete *Fugu* NOS1 sequence encompasses 22,203 nucleotides (appendix, figure A.1) and has 29 exons and 28 introns. The predicted *Fugu* NOS1 coding sequence is 4257 nucleotides long giving rise to a protein of 1418 amino acids. The G + C content of *Fugu* NOS1 is 46.3% in line with the 44.2% G + C content of the *Fugu* genome (Brenner *et al.*, 1993). *Fugu* NOS1 also has microsatellite sequences which include A, C, AC/GT and TTA/TAA repeats, these are among the 20 most commonly found microsatellites in *Fugu* (Edwards *et al.*, 1998).

Figure 3.8. Schematic representation of *Fugu* NOS1 exons 1 to 7 and 19 to 29. Sequence data obtained was used to construct contiguous sequences for exons 1 to 7 and 19 to 29. Arrows indicate the relative positions of primers pairs used for the PCR amplification. Solid black boxes represent the exons of *Fugu* NOS1 which are numbered according to the exons of the human NOS1 gene.

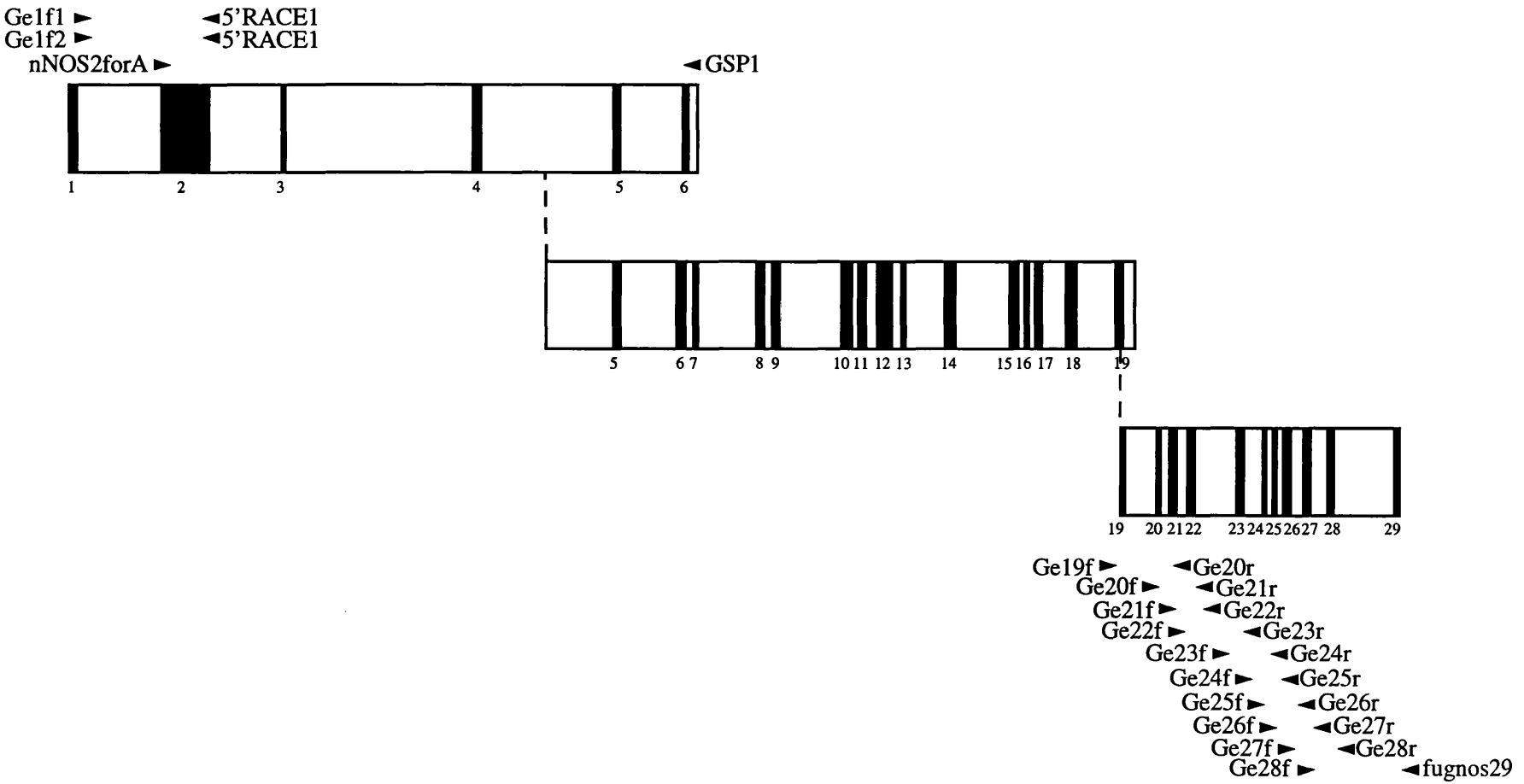

Ge1f1 ►
Ge1f2 ►
nNOS2forA ►
◄5'RACE1
◄5'RACE1
◄GSP1
1 2 3 4 5 6
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
19 20 21 22 23 24 25 26 27 28 29
Ge19f ►
◄Ge20r
Ge20f ►
◄Ge21r
Ge21f ►
◄Ge22r
Ge22f ►
◄Ge23r
Ge23f ►
◄Ge24r
Ge24f ►
◄Ge25r
Ge25f ►
◄Ge26r
Ge26f ►
◄Ge27r
Ge27f ►
◄Ge28r
Ge28f ►
◄fugnos29

Figure 3.9. Schematic representation of the *Fugu* NOS1 gene. Sequence data from pBs9.6, its subclones and the PCR-amplified fragments, was used to construct a contiguous sequence for the entire *Fugu* NOS1 gene. Solid black boxes represent exons which are numbered according to the exons of the human NOS1 gene.


1
2
3
4
5
6 7
8 9
10 11 12 13
14
15 16 17
18
19
20 21 22
23 24 25 26 27 28
29

Table 3.1 shows the percentage nucleotide identity, predicted amino acid identity and similarity of the individual *Fugu* NOS1 exons to their human NOS1 counterparts, as predicted by the GAP algorithm of Needleman and Wunsch (1970). Exon 2 shows limited sequence similarity of 33%, this is also the case for exon 3 at both the nucleotide and amino acid level. Exon 2 of human NOS1 contains a sequence motif for a discs-large homologous region (DHR), also known as 'GLGF repeat' or 'PDZ domain', which spans the amino acid residues 18 - 97 and is responsible for protein: protein interactions (Hendricks, 1995). PDZ domains are typically 80 - 100 amino acids long and proteins that have PDZ domains are implicated in ion-channel and receptor clustering, and the linking of receptors and their effector enzymes (Ponting and Phillips, 1995).

In neurons, nNOS forms heterodimers with postsynaptic density-95 protein (PSD-95) by the interaction of their PDZ domains (Brenman *et al.*, 1996), this is thought to couple nNOS to the N-methyl-D-aspartate (NMDA) receptor and therefore NO production to NMDA receptor activation (Dawson *et al.*, 1992; Huang *et al.*, 1994). In muscle cells, nNOS heterodimerizes with syntrophin via their PDZ domains, to localise nNOS to the dystrophin complex (Brenman *et al.*, 1995), this couples NO production to muscle contraction. The predicted amino acid sequence of *Fugu* NOS1 exon 2 shows only 55% amino acid similarity to its human equivalent, but this is due in part to the differing lengths of these exons, the similarity between the two exons rises to 80% for the first 125 residues and this includes the region covered by the PDZ domain.

The first four exons of *Fugu* NOS1 are the only ones to show any difference in length as compared to the human NOS1 sequence. Table 3.2 shows a comparison of the *Fugu* NOS1 and human NOS1 exons in terms of length, both nucleotide and predicted amino acid, and their derived features.

Table 3.1. Comparison of *Fugu* and human NOS1 exons by GAP alignment

Exon	nucleotide identity (%)	amino acid similarity (%)	amino acid identity (%)
1	32.8	-	-
2	57.7	65.8	54.7
3	36.9	38.5	28.2
4	65.1	74.4	67.4
5	58.2	69.4	57.1
6	77.9	91.1	85.7
7	82.6	100.0	93.3
8	77.5	93.8	91.7
9	74.3	87.2	80.9
10	83.4	91.4	84.5
11	78.4	91.2	85.3
12	82.1	93.8	92.3
13	79.1	89.3	85.7
14	75.9	93.8	91.7
15	79.0	97.1	88.6
16	61.0	75.0	70.0
17	74.4	79.5	74.4
18	78.9	94.8	93.1
19	79.1	91.5	87.2
20	56.9	85.2	74.1
21	72.2	82.8	78.1
22	75.3	78.9	75.4
23	79.1	90.9	83.3
24	69.3	82.8	75.9
25	71.3	87.5	84.4
26	75.2	85.0	77.5
27	69.2	78.5	69.2
28	83.2	95.5	84.1
29	42.5	75.0	75.0

Table 3.2. Comparison of exon size between *Fugu* and human NOS1.

Exon	Exon Size (bp)		Amino Acids		Features[a]
	Fugu	Human	*Fugu*	Human	
1	193	266	-	-	5' UTR
2	759	1145	231	241	5' UTR GLGF
3	102	127	34	42	
4	138	129	46	43	
5	146	146	49	49	
6	163	163	54	54	Haem binding
7	92	92	31	31	
8	142	142	47	47	
9	140	140	47	47	
10	175	175	58	58	
11	102	102	34	34	
12	195	195	65	65	
13	86	86	29	29	Ca^{2+}/calmodulin
14	145	145	48	48	Ca^{2+}/calmodulin
15	105	105	35	35	
16	59	59	20	20	
17	117	117	39	39	
18	175	175	58	58	FMN binding
19	139	139	46	46	
20	79	79	27	26	
21	194	194	64	65	FAD binding
22	170	170	57	57	
23	211	211	70	70	FAD binding
24	88	88	30	29	
25	122	122	40	41	NADPH binding
26	149	149	50	50	
27	195	195	65	65	NADPH binding
28	119	119	40	40	
29	87	2150	4	4	3' UTR

[a] Features; deduced by comparison of the sequence to human neuronal nitric oxide synthase (Hall *et al.*, 1994). Amino acids with interrupted codons were assigned to the exon containing two of the three codon nucleotides.

The highest identity and similarity is seen between exon 7 of *Fugu* NOS1 and human NOS1 however, to date no function has been assigned to the residues of this exon. High similarity is seen for many of the exons notably, exons 6, 13, 14, 18, 21, 23, 25 and 27. These exons have been assigned as binding domains for haem, calmodulin (*Fugu* NOS1 residues 713 - 733), FMN (870 - 901), FAD pyrophosphate (1014 - 1026), FAD isoalloxizine (1159 - 1170), NADPH ribose (1234 - 1252) and NADPH adenine (1331 - 1347) respectively (Bredt *et al.*, 1991; Vorherr *et al.*, 1993; Zhang and Vogel, 1994; Venema *et al.*, 1996). In the haem domain of all known NOS sequences there is a cysteine residue (NOS1 C415, NOS3 C184, NOS2 C194) whose mutation results in the loss of haem binding (Richards and Marletta, 1994; Chen *et al.*, 1994; Sari *et al.*, 1996), at the equivalent position in the predicted *F*NOS sequence, residue 404, is also a cysteine residue. In addition, in exon 5 of the *F*NOS predicted amino acid sequence there is another cysteine residue (C319) which is conserved in all known NOS sequences. The mutation of the equivalent cysteine residue in human eNOS (C99) to alanine was found to result in a large loss of BH_4 binding as measured by enzyme activation (Chen *et al.*, 1995), the same mutation of the corresponding residue in rat nNOS (C331) had a similar effect (Martásek *et al*, 1998). More recent work, on the crystal structure of bovine eNOS (Raman *et al.* 1998), has shown that C101 of bovine eNOS (C99 of human eNOS) is part of a Zinc binding motif (C - $(X)_4$ - C) which is also conserved in the predicted *F*NOS sequence. The two cysteine residues of the zinc binding motif co-ordinate the zinc atom which functions both to stabilise the BH_4-binding site and facilitate recognition of stereospecific pterin. Also present in the *F*NOS predicted amino acid sequence is a region described as a dihydrofolate reductase (DHFR) homology fragment, in rat nNOS this fragment (residues 564 - 715) was identified as the L-arginine

binding domain as it was found to bind the L-arginine analogue N^G-nitro-L-arginine (NNA) (Nishimura *et al.*, 1995). Residues 830 to 870 of rat nNOS were identified as a calmodulin inhibitory sequence which is present in constitutive NOS isoforms but is absent from inducible isoforms (Salerno *et al.*, 1996), this region is present in the predicted *F*NOS sequence.

At 22,203 nucleotides long, *Fugu* NOS1 is approximately 7.2 times smaller than human NOS1 (≥160 kb; Hall *et al.*, 1994), this is in line with the overall 7.5 times reduction of the *Fugu* genome (400 Mb) as compared to the human genome (3000 Mb) (Brenner *et al.*, 1993). The predicted coding sequence of *Fugu* NOS1 is of comparable size to the human nNOS coding sequence, which at 4302 nucleotides translates into 1434 amino acids. The reduction in the size of the *Fugu* NOS1 coding sequence can be attributed to the reduction in intron size as shown in Table 3.3, this is in agreement with data confirming that the *Fugu* genome is 7.5 times smaller than the human genome but has the same amount of coding sequence (Brenner *et al.*, 1993). All of the *Fugu* NOS1 introns are reduced in size, except introns 4 and 7. The greatest reduction in intron size is seen for intron 8, which at 92 nucleotides long is 130.4 times smaller than its human equivalent. 80% of *Fugu* introns are less than 150 bp in size although there is at least one large intron (greater than 400 bp) present in *Fugu* genes (Elgar *et al.*, 1996). Of the *Fugu* NOS1 introns, fifteen are less than 400 bp in size, with ten being less than 150 bp, eight are between 400 and 1000 bp and five are greater than 1000 bp. Also shown in Table 3.3 is the intron/exon boundary type and in all cases the boundary type is identical to those of human NOS1. Table 3.4 shows that the proposed intron/exon boundaries for *Fugu* NOS1 conform very closely to the Mount consensus for splice junction sequences

Table 3.3. Comparison of introns between *Fugu* and human NOS1.

Intron	Type [a]		Size (bp)		Reduction[b]
	Fugu	Human	*Fugu*	Human	
1	-	-	1322	≥25000	≥19.0
2	II	II	1177	≥25000	≥21.0
3	0	0	3062	20000	6.5
4	0	0	2312	2000	↑1.2
5	II	II	886	1800	2.0
6	0	0	105	800	7.6
7	II	II	945	800	↑1.2
8	0	0	92	12000	130.4
9	II	II	1161	2700	2.3
10	0	0	87	4000	47.1
11	0	0	77	2300	29.9
12	0	0	199	1400	7.0
13	II	II	545	3300	6.1
14	0	0	978	1300	1.3
15	0	0	80	700	8.8
16	II	II	75	5000	66.7
17	II	II	395	6000	15.2
18	0	0	747	4000	5.4
19	I	I	478	600	1.3
20	II	II	111	8000	72.1
21	I	I	116	6000	51.7
22	0	0	654	1100	1.7
23	I	I	386	700	1.8
24	II	II	112	1500	13.4
25	I	I	91	2000	22.0
26	0	0	159	300	1.9
27	0	0	275	2000	7.3
28	II	II	989	2800	2.8

[a] Intron type; 0 indicates a splice junction between codons, I indicates a splice junction after the first nucleotide of a codon, and II indicates a splice junction that occurs after the second nucleotide. Amino acids with interrupted codons were assigned to the exon containing two of the three codon nucleotides. [b]Reduction; as compared to the length of corresponding introns from human NOS1 (Hall *et al.*, 1994). ↑ indicates an increase in the *Fugu* NOS1 intron size relative to its human equivalent.

Table 3.4. Predicted splice junction sequences in the *Fugu* NOS1 gene sequence.

Intron	Exon	5′ Donor	3′ Acceptor	Exon
1	CCCG	GTGAGT	TCTATTTCTTTTCAG	AAAC
2	CCTG	GTAATT	CTGTTTGACTCCCAG	GGGC
3	CAAG	GTAAAA	TTTGTGCCCTTGCAG	CCTC
4	CAAG	GTGAGA	ATGGTTTGTTTACAG	ACGC
5	AAAG	GTAAGC	GTAATGTGACTGCAG	GTAC
6	ACAG	GTGAAA	GCGCTCACATTTCAG	GTTT
7	TGAG	GTAAAA	CTTGATTTGTTGCAG	GTCA
8	TGAG	GTGCGA	TGCTGTCGCCATTAG	ATCT
9	CAAA	GTGAGT	TGTACCCCTTCCAAG	GTAT
10	GGAG	GTAACG	CTGTTTACTGCTTAG	GAGG
11	CCAG	GTAAAC	GTGCTTCATTCTCAG	TCAT
12	CCAG	GTGACC	TCCCATATGCAAAAG	CTTG
13	CCAA	GTGAGC	CCCCCCTCCTTGCAG	AGCG
14	AAAG	GTACCA	TTGTCGTCTCTGTAG	GTCA
15	AGAG	GTATGC	TGCCCGTCTCTGTAG	AAAT
16	GAAA	GTGAGT	GTTCCGTCCCTTCAG	GAGC
17	TCAG	GTTAAT	TGTGTCTCTGTTTAG	GTTC
18	TAAG	GTCGGT	TTTCCTGTGACGCAG	GCCG
19	AAAG	GTAAAT	ATGTTCTGCCAACAG	CTTT
20	CCAA	GTAAGT	CACTTGTTTCTGCAG	TCGT
21	CTGG	GTGGGT	CCTTCTTCCCCTCAG	GTGT
22	CAAG	GCAAGT	GGCTTCCGTCCACAG	GGCT
23	AGAG	GTGAGG	AACCATCCTTTTCAG	ATGG
24	GAAG	GTAAAT	TTTTTTGTTTTGAAG	TGCT
25	AACG	GTGAGA	TTTACCCCTCTGTAG	GCAT
26	AAAG	GTACTA	CTGTCTCCTTGTCAG	AAAT
27	TCGG	GTACGA	TACTACCTAACACAG	GATG
28	ATGA	GTAAGT	GTCGTCTCCCCTCAG	GGTG

Mount consensus for splice junction sequences

MAG	GTRAGT	Y(n)NYAG	G

(Mount, 1982). In all cases except one, intron 23, the 5′ donors begin with GT and all of the 3′ acceptors terminate with AG.

3.6 Summary

A NOS1 gene has been identified and cloned from the genome of the Japanese pufferfish *Fugu rubripes* by library screening and PCR amplification of genomic DNA. *Fugu* NOS1 is approximately 7.2 times smaller than human NOS1, but the predicted coding sequence is of a comparable size to the human NOS1 coding sequence, the reduction in the size of *Fugu* NOS1 is due to a reduction in intronic DNA. *Fugu* NOS1 contains sequences that are known to code for functional binding domains in all NOS enzymes including: haem, L-arginine, calmodulin, FMN, FAD and NADPH. Specific residues, within these domains, are known to be essential for BH_4 binding and haem binding and have been identified in all known NOS sequences, they are conserved in the *Fugu* NOS1 sequence. The calmodulin inhibitory sequence element, seen only in constitutive NOS enzymes, is also present in the *Fugu* NOS1 sequence. Overall, *Fugu* NOS1 has greatest similarity to the mammalian NOS1 isoforms, and in particular human NOS1, which is evident from the presence of sequence coding for over 200 amino acids more at the N-terminal end, as compared to either iNOS or eNOS isoforms. Within the 200 amino acids is a PDZ domain, a region known only to be found in neuronal NOS isoforms. *Fugu* NOS1 contains the predicted coding sequence of an enzyme which, by comparison to known NOS sequences, should be produced constitutively; Ca^{2+}/CaM dependent and able to be localised to cell-cell junctions through its PDZ domain.

CHAPTER 4

CLONING AND SEQUENCING OF FULL-LENGTH *FUGU* NOS1 cDNA (*fNOS*)

CONTENTS

4.1. Introduction.

To try and isolate the full-length *Fugu* NOS1 cDNA (*f*NOS), RNA had to first be isolated. Due to a lack of material available from *Fugu rubripes*, the pufferfish used for the genomic study, another Japanese pufferfish *T. poecilonotus* (Figure 4.1) was used instead. Three different species of pufferfish were collected from Yoshihama Bay, Iwate-ken, Japan by Dr. Shigero Sato, of the Kitasato University School of Fisheries Sciences in Japan. The pufferfish brains were removed and snap-frozen ready for RNA isolation, they were then shipped over on dry ice. The three different species of pufferfish brain sent were *T. pardalis*, *T. vermicularis* and *T. poecilonotus*. Pufferfish are phylogenetically closely related species and even the freshwater pufferfish *T. fluviatilis* is estimated to have diverged from *Fugu rubripes* only approximately 18-30 million years ago (Crnogorac-Jurcevic *et al.*, 1997). *T. poecilonotus*, or Komonfugu, was chosen for this study due to its even closer evolutionary position to *Fugu rubripes* (Table 4.1). Brain was chosen as the starting material for the RNA isolation because the *Fugu* NOS1 gene sequence showed most similarity to the human NOS1 isoform, as described in chapter 3, and mRNA transcripts for nNOS have been detected in both the central and peripheral nervous system.

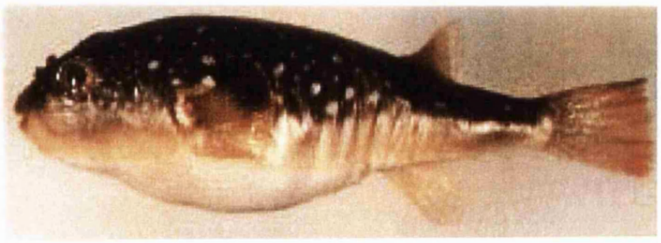


Figure 4.1. The Japanese pufferfish *T. poecilonotus*.

Table 4.1. Partial classification of the Tetraodontiformes order of fishes.

Superfamily	Family	Subfamily	Genus	Species & common name
Tetraodontoidea	Tetraodontoidae	Tetraodontinae	Takifugu	***T. pardalis* (Higanfugu)**
				***T.vermicularis* (Nashifugu)**
				T. snyderi (Syosaifugu)
				T. porphyreus (Mafugu)
				T. chrysops (Akamefugu)
				***T. poecilonotus* (Komonfugu)**
				T. alboplumbeus (Komondamashi)
				T. stictonotus (Gomsfugu)
				T. exascurus (Mushifugu)
				T. oblongus (Takifugu)
				T. niphobles
				T. basilevskianus (Kouraifugu)
				T. reticularis (Amimefugu)
				***T. rubripes* (Torafugu)**
				T. chinensis (Karasu)
				T. obscurus (Mefugu)
				T. flavidus (Sansaifugu)
				T. bimaculatus (Fitatsuboshifugu)
				T. pseudommus (Nameradanshi)
				T. xanthopterus (Shimafugu)
				T. ocellatus (Meganafugu)

The three species sent from Japan are highlighted in bold type along with *Fugu rubripes*, demonstrating their evolutionary positions relative to one another.

4.2. Isolation of total RNA from *Fugu* brain.

Total RNA was isolated from a *T. poecilonotus* brain, which weighed 164 mg, as described in section 2.17. A 1 in 100 dilution of the total RNA isolated was made in DEPC-treated distilled water and the yield of total RNA estimated using ultra-violet spectroscopy and the following equation:

$$OD_{260\,nm} = 1 \text{ for approximately } 40\ \mu g/ml \text{ RNA}$$

$$\textit{Fugu} \text{ total RNA} = \frac{0.345 \times 100 \times 50}{1000}$$

$$= 69\ \mu g \text{ of total RNA}$$

where: 0.345 = $OD_{260\,nm}$ of the sample
100 = dilution factor
50 = original sample volume (µl)
1000 = 10^3 µl per ml

The OD_{260}/OD_{280} ratio was calculated as 1.64.

4.3. Amplification of *f*NOS.

3.5 µg of *Fugu* total RNA was used to amplify double-stranded cDNA (section 2.18). The yield of experimental double-stranded cDNA generated was estimated as being half that of the double-stranded cDNA made from the positive control human placental RNA. Figure 4.2 shows the agarose gel electrophoresis of $1/5^{th}$ of the reaction volume of each sample which was used to estimate yield. Once adaptor-ligated, the double-stranded cDNA was diluted according to its estimated yield. Positive control, adaptor-ligated and double-stranded cDNA was diluted 1 in 250, the *Fugu* adaptor-ligated double-stranded

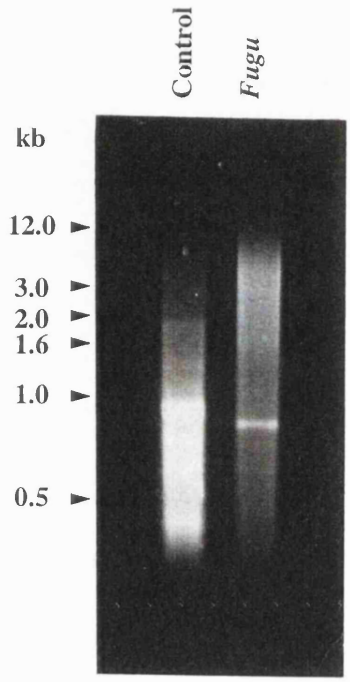
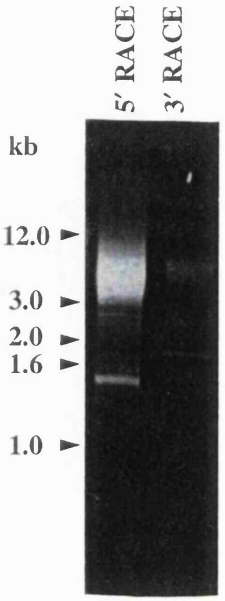


Figure 4.2. Amplification of *Fugu* brain double-stranded cDNA. Double-stranded cDNA was amplified from 3.5µg of total RNA isolated from the brain of the pufferfish *T. poecilonotus*. $1/5^{th}$ of the double-stranded cDNA was electrophoresed through an ethidium bromide-stained agarose gel and photographed under U.V. light. The yield of the *Fugu* cDNA was estimated to be half that of the positive control cDNA. Molecular size markers are indicated.

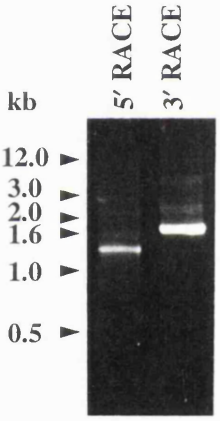
cDNA was therefore diluted 1 in 125. 5′ and 3′ RACE-PCR was performed using adaptor primers and gene specific primers designed against the *Fugu* NOS1 gene sequence (section 2.18). 1/10th of the RACE-PCR reactions were electrophoresed on an agarose gel (section 2.7) as shown in figure 4.3A. Nested RACE-PCR was performed on 5 µl of the primary RACE-PCR product which had been diluted 1 in 250 with TE buffer (section 2.18). Again, 1/10th of the reaction volume for each sample was analysed by agarose gel electrophoresis (section 2.7) as shown in figure 4.3B. The original and nested RACE-PCR products were gel purified (section 2.9), ligated (section 2.11) and transformed into TOP10 *E. coli* (section 2.13). Colony PCR (section 2.14) was performed to identify colonies containing plasmids with inserts. Plasmid DNA was isolated (section 2.5) and inserts sequenced using T7, T3, M13 forward and reverse primers, this was followed by sequencing with specific primers (section 2.16). The 5′ RACE sequence data revealed the presence of 260 nucleotides upstream of the predicted ATG, these corresponded to 193 nucleotides of exon 1 and 67 nucleotides from exon 2 (appendix, figure A.2). The 3′ RACE sequence data revealed the presence of 71 nucleotides downstream of the proposed stop codon TAG, also seen in exon 29 of the *Fugu* NOS1 sequence (appendix, figure A.2). The 5′ and 3′ RACE sequence was used to design primers for the amplification of full-length *Fugu* NOS1 cDNA (*f*NOS) which is shown in figure 4.3C. The primers designed were engineered to contain *Not*I restriction endonuclease sites so that these sites would be incorporated at either end the full-length cDNA. Sequencing demonstrated that there was no *Not*I site present in the predicted coding sequence of *Fugu* NOS1 making it suitable for the cloning of full-length *f*NOS into the multiple cloning site of the baculovirus expression vector. Full-length *f*NOS was

Figure 4.3. Amplification of full-length *f*NOS cDNA. (A), Primary 5′ and 3′ RACE-PCR was performed on double-stranded *Fugu* cDNA. $1/10^{th}$ of each RACE-PCR reaction was electrophoresed through an ethidium bromide-stained agarose gel and photographed under U.V. light. (B), Nested 5′ and 3′ RACE-PCR was performed on 5µl of a 1 in 250 dilution of each respective primary RACE-PCR product. $1/10^{th}$ of the nested RACE-PCR products were separated through an ethidium bromide-stained agarose gel and photographed under U.V. light. (C), Full-length *f*NOS was amplified by PCR with primers designed from sequence information obtained from the 5′ and 3′ RACE-PCR products. $1/10^{th}$ of the PCR product was electrophoresed through an ethidium bromide-stained agarose gel and photographed under U.V. light. Molecular size markers are indicated.

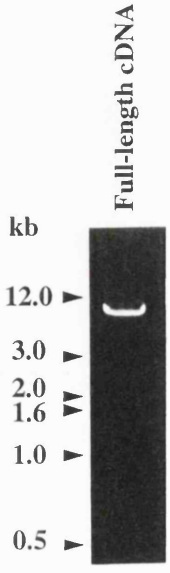
A


5′ RACE
3′ RACE
kb
12.0
3.0
2.0
1.6
1.0

B


kb
5′ RACE
3′ RACE
12.0
3.0
2.0
1.6
1.0
0.5

C


Full-length cDNA
kb
12.0
3.0
2.0
1.6
1.0
0.5

digested with *Not*I (section 2.6) and ligated into vector pVL1393 (section 2.10) to produce pVL*f*NOS, as illustrated in figure 4.4. Transformation into subcloning efficiency DH5α *E. coli* (section 2.12) was followed by colony PCR to identify colonies containing pVL*f*NOS (section 2.14) and plasmid DNA was isolated from those colonies (section 2.5). The orientation of inserts was checked with diagnostic digests (section 2.6). Figure 4.5 shows the correct orientation of one insert.

4.4. Sequence analysis of *f*NOS.

Sequencing of full-length *f*NOS coding cDNA sequence was achieved using primers designed from the *Fugu* NOS1 gene sequence (appendix, table A.1). Including the stop codon, *f*NOS was found to be 4257 nucleotides long. When translated and analysed using the PEPTIDESORT programme in the Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wisconsin, the *f*NOS sequence gives rise to a predicted protein (*F*NOS) of 1419 amino acids with a molecular weight of 159.7kDa. The complete sequence of *f*NOS and the predicted amino acid sequence of *F*NOS are shown in the appendix (figure A.2).

4.5. Comparative analysis of *f*NOS.

BLASTN analysis of the *f*NOS nucleotide and *F*NOS predicted amino acid sequences confirmed greatest similarity to the vertebrate NOS1 isoforms. Table 4.2 shows the percentage similarity and identity of *f*NOS and *F*NOS, to the complete sequences of other known NOS isoforms, as calculated by GAP alignment to each individual

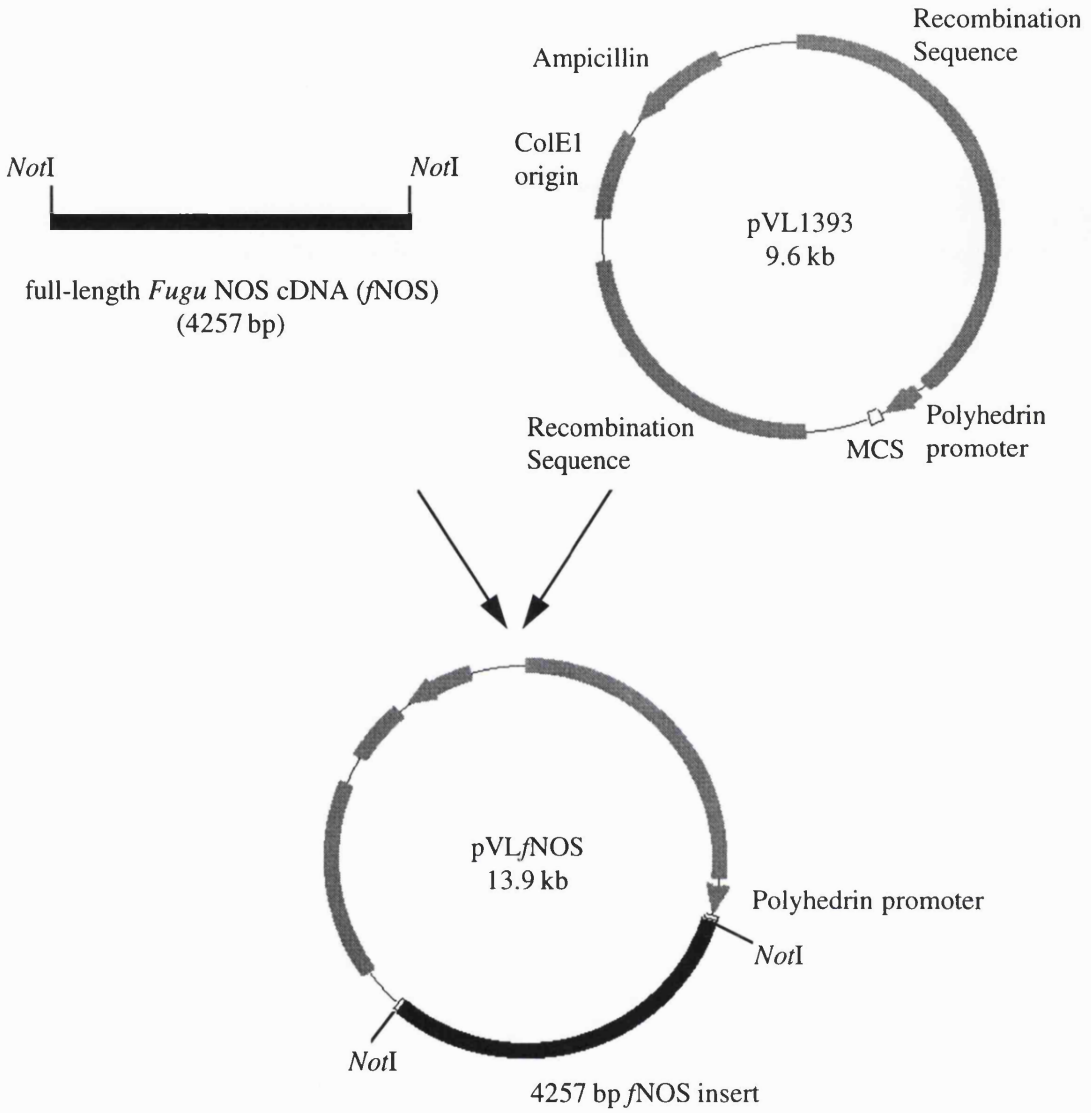


Figure 4.4. Cloning of *f*NOS into the plasmid vector pVL1393 to create pVL*f*NOS. Full-length *Fugu* NOS1 cDNA (*f*NOS) was digested with the restriction endonuclease *Not*I and cloned into the *Not*I site of the baculovirus expression vector pVL1393 to create pVL*f*NOS.

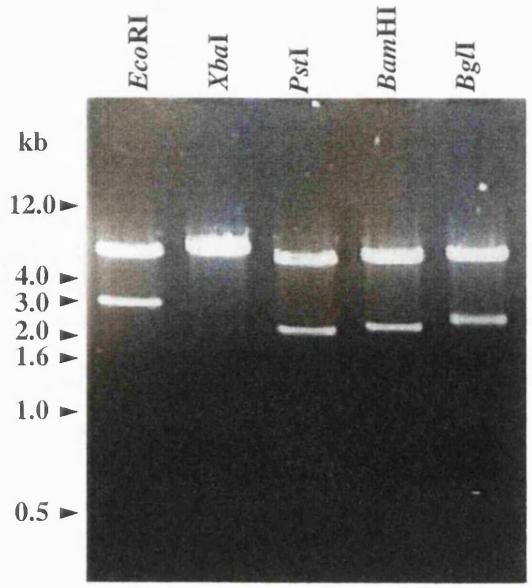


Figure 4.5. Orientation of the *f*NOS fragment in pVL*f*NOS. DNA was digested with *Eco*RI, *Xba*I, *Pst*I, *Bam*HI and *Bgl*I, separated through an ethidium bromide-stained agarose gel and photographed under U.V. light. Molecular size markers are indicated. The approximate sizes of the restriction fragments obtained were compared to the sizes predicted for a correctly orientated fragment (calculated from the mapping of sequence data obtained for the *f*NOS cDNA sequence). The *f*NOS fragment shown was correctly orientated.

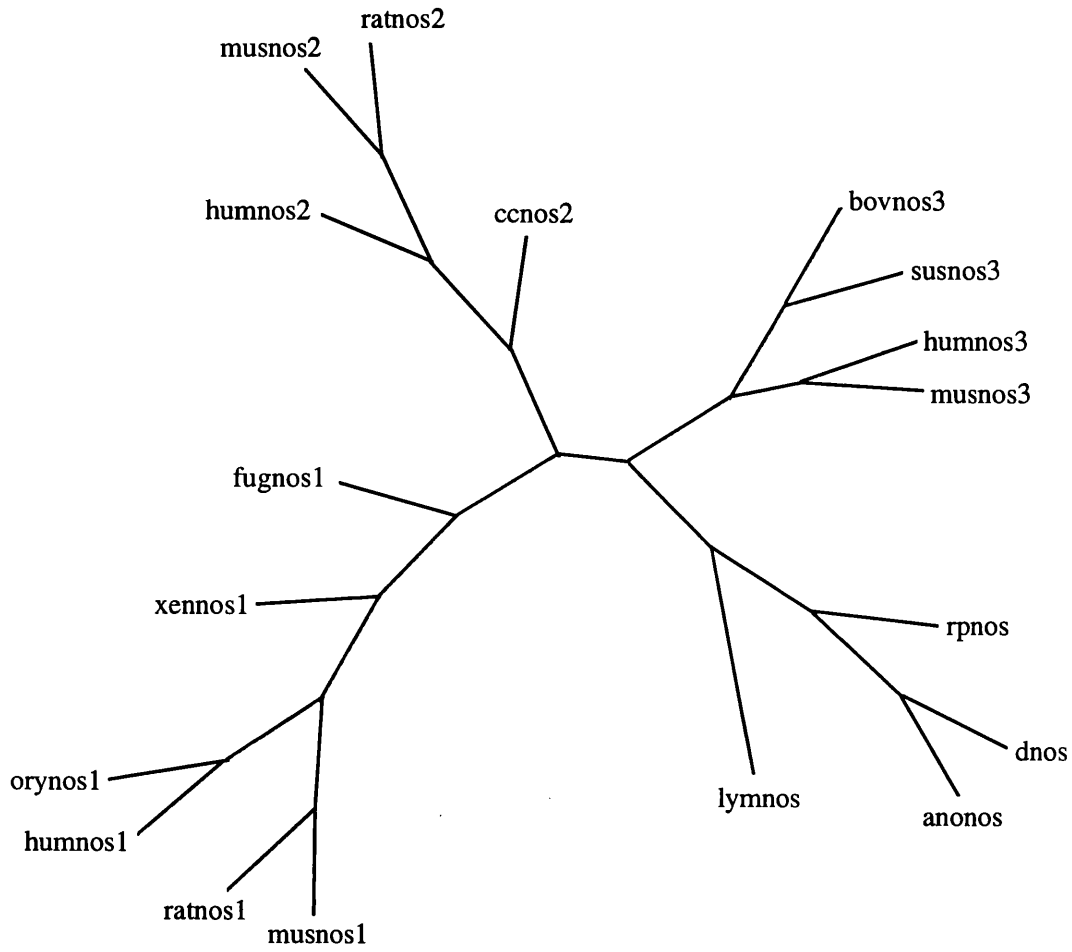
sequence. All of the nNOS isoforms, including the *Xenopus laevis* sequence, have very high percentage similarities and identities to the predicted *F*NOS sequence. Human nNOS has the highest amino acid similarity (82%) and identity (75%). Lower amino acid similarity is seen between *F*NOS and the endothelial and inducible isoforms, dropping to between 61 - 66%, and also to the reported invertebrate NOS sequences.

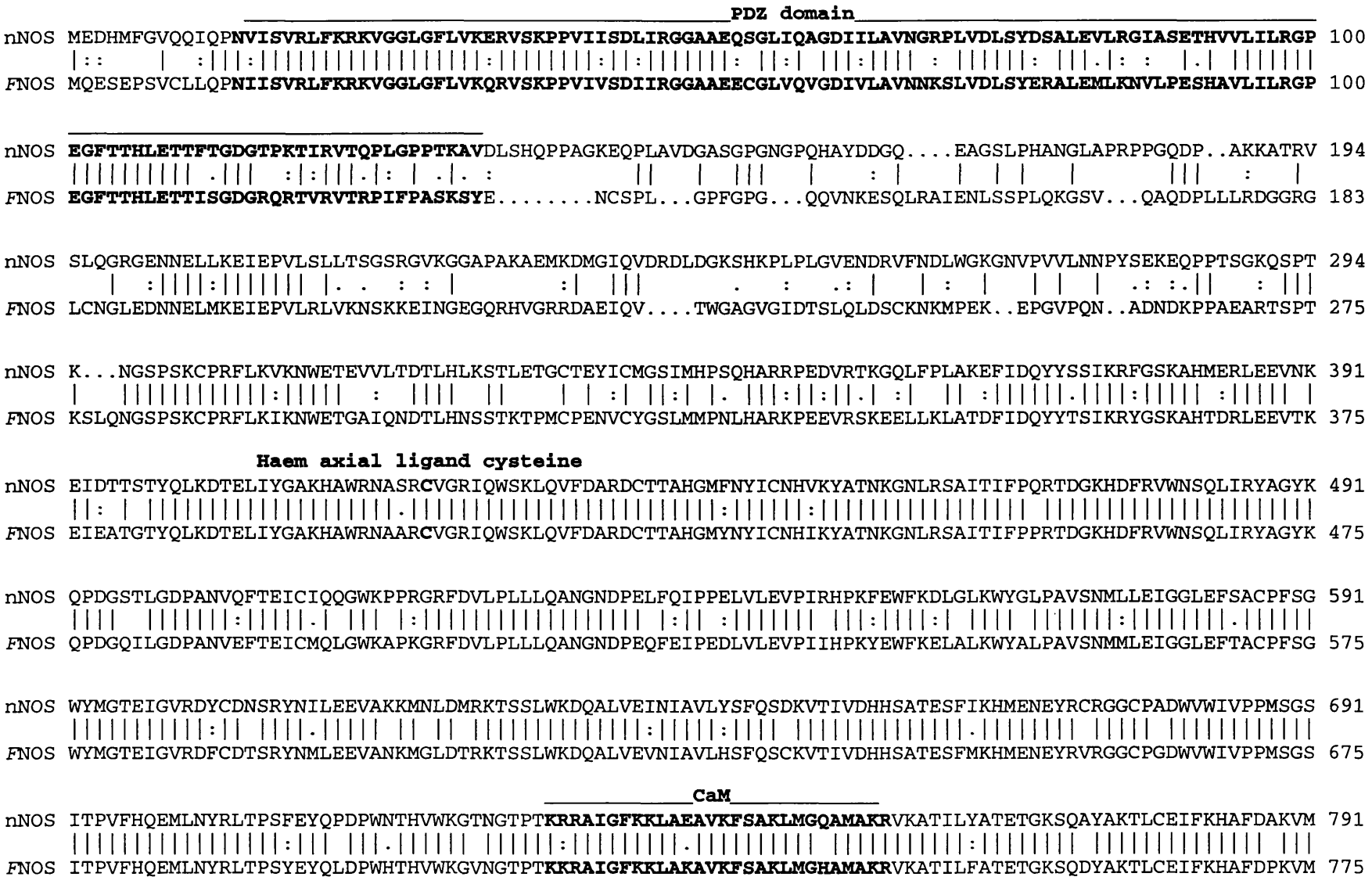
Table 4.2. Sequence comparisons of the *f*NOS nucleotide and predicted *F*NOS amino acid sequence to other known NOS isoforms using the GAP alignment programme.

Species and Isoform	Nucleotide Identity (%)	Amino Acid Similarity (%)	Amino Acid Identity (%)
Human, neuronal	72	82	75
Rabbit, neuronal	73	81	75
Mouse, neuronal	72	82	74
Rat, neuronal	72	81	74
Xenopus laevis, neuronal	68	82	74
Human, endothelial	62	66	58
Mouse, endothelial	61	66	58
Bovine, endothelial	62	66	58
Porcine, endothelial	62	65	58
Rat, inducible	59	62	54
Human, inducible	62	61	54
Mouse, inducible	59	62	54
C. carpio, inducible	57	61	52
R. prolixus	54	60	52
L. stagnalis, neuronal	50	60	50
A. stephensis	53	58	48
D. melangoster, neuronal	52	56	47

CLUSTAL W (1.74) alignment of *F*NOS (Thompson *et al.*, 1994), with the amino acid sequences of other NOS isoforms (appendix, figure A.3), was followed by phylogenetic analysis using the PIE programme available through the MRC HGMP Resource Centre, the results of this analysis are shown in figure 4.6. Again this analysis confirmed the status of *F*NOS as being most like an nNOS isoform. Figure 4.7 shows the GAP alignment of the predicted *F*NOS amino acid sequence to human nNOS and marks the

Figure 4.6. Phylogenetic comparison of *F*NOS with other known NOS isoforms. Full-length NOS amino acid sequences were aligned using the CLUSTAL W (1.74) multiple sequence alignment programme and a phylogenetic tree constructed by the PROPARS programme (available through the Phylogeny Interface Environment (PIE) at http://www.hgmp.mrc.ac.uk/Registered/Webapp/pie/) with 1000 bootstrap replications. Where anonos = *Anopheles stephensi* NOS (Luckhart *et al.*, 1998; accession number AF053344), bovnos3 = bovine eNOS (Lamas *et al.*, 1992; accession number M89952), ccnos = Cyprinus carpio iNOS (Saeij *et al.*, 1999; accession number AJ242906), dnos = *Drosophila melangoster* nNOS (Regulski and Tully, 1995; accession number U25117), humnos1 = human nNOS (Hall *et al.*, 1994; accession number U17327), humnos2 = human iNOS (Geller *et al.*, 1993; L09210), humnos3 = human eNOS (Janssens *et al.*, 1992; accession number M93718), Lymnos = *Lymnaea stagnalis* nNOS (Korneev *et al.*, 1998; AF012531), musnos1 = murine nNOS (Ogura *et al.*, 1993; accession number D14552), musnos2 = murine iNOS (Kone *et al.*, 1995; accession number U43428), musnos3 = murine eNOS (Gnanapandithen *et al.*, 1996; U53142), orynos1 = rabbit nNOS (Jeong and Yim, 1997; accession number U91584), ratnos1 = rat nNOS (Magee *et al.*, 1996; accession number U67309), ratnos2 = rat iNOS (Kosuga *et al.*, 1994; accession number D44591), rpnos = *Rhodnius prolixus* NOS (Yuda, 1996; accession number U59389), susnos3 = porcine eNOS (Zhang *et al.*, 1996; accession number U59924), xennos = *Xenopus laevis* nNOS (Sheinker *et al.*, 1998; accession number AF053935).


ratnos2
musnos2
humnos2
ccnos2
bovnos3
susnos3
humnos3
musnos3
fugnos1
xennos1
rpnos
dnos
orynos1
humnos1
lymnos
anonos
ratnos1
musnos1


PDZ domain
nNOS MEDHMFGVQQIQPNVISVRLFKRKVGGLGFLVKERVSKPPVIISDLIRGGAAEQSGLIQAGDIILAVNGRPLVDLSYDSALEVLRGIASETHVVLILRGP 100
FNOS MQESEPSVCLLQPNIISVRLFKRKVGGLGFLVKQRVSKPPVIVSDIIRGGAAEECGLVQVGDIVLAVNNKSLVDLSYERALEMLKNVLPESHAVLILRGP 100
nNOS EGFTTHLETTFTGDGTPKTIRVTQPLGPPTKAVDLSHQPPAGKEQPLAVDGASGPGNGPQHAYDDGQ....EAGSLPHANGLAPRPPGQDP..AKKATRV 194
FNOS EGFTTHLETTISGDGRQRTVRVTRPIFPASKSYE.....NCSPL...GPFGPG...QQVNKESQLRAIENLSSPLQKGSV...QAQDPLLLRDGGRG 183
nNOS SLQGRGENNELLKEIEPVLSLLTSGSRGVKGGAPAKAEMKDMGIQVDRDLDGKSHKPLPLGVENDRVFNDLWGKGNVPVVLNNPYSEKEQPPTSGKQSPT 294
FNOS LCNGLEDNNELMKEIEPVLRLVKNSKKEINGEGQRHVGRRDAEIQV....TWGAGVGIDTSLQLDSCKNKMPEK..EPGVPQN..ADNDKPPAEARTSPT 275
nNOS K...NGSPSKCPRFLKVKNWETEVVLTDTLHLKSTLETGCTEYICMGSIMHPSQHARRPEDVRTKGQLFPLAKEFIDQYYSSIKRFGSKAHMERLEEVNK 391
FNOS KSLQNGSPSKCPRFLKIKNWETGAIQNDTLHNSSTKTPMCPENVCYGSLMMPNLHARKPEEVRSKEELLKLATDFIDQYYTSIKRYGSKAHTDRLEEVTK 375
Haem axial ligand cysteine
nNOS EIDTTSTYQLKDTELIYGAKHAWRNASRCVGRIQWSKLQVFDARDCTTAHGMFNYICNHVKYATNKGNLRSAITIFPQRTDGKHDFRVWNSQLIRYAGYK 491
FNOS EIEATGTYQLKDTELIYGAKHAWRNAARCVGRIQWSKLQVFDARDCTTAHGMYNYICNHIKYATNKGNLRSAITIFPPRTDGKHDFRVWNSQLIRYAGYK 475
nNOS QPDGSTLGDPANVQFTEICIQQGWKPPRGRFDVLPLLLQANGNDPELFQIPPELVLEVPIRHPKFEWFKDLGLKWYGLPAVSNMLLEIGGLEFSACPFSG 591
FNOS QPDGQILGDPANVEFTEICMQLGWKAPKGRFDVLPLLLQANGNDPEQFEIPEDLVLEVPIIHPKYEWFKELALKWYALPAVSNMMLEIGGLEFTACPFSG 575
nNOS WYMGTEIGVRDYCDNSRYNILEEVAKKMNLDMRKTSSLWKDQALVEINIAVLYSFQSDKVTIVDHHSATESFIKHMENEYRCRGGCPADWVWIVPPMSGS 691
FNOS WYMGTEIGVRDFCDTSRYNMLEEVANKMGLDTRKTSSLWKDQALVEVNIAVLHSFQSCKVTIVDHHSATESFMKHMENEYRVRGGCPGDWVWIVPPMSGS 675
CaM
nNOS ITPVFHQEMLNYRLTPSFEYQPDPWNTHVWKGTNGTPTKRRAIGFKKLAEAVKFSAKLMGQAMAKRVKATILYATETGKSQAYAKTLCEIFKHAFDAKVM 791
FNOS ITPVFHQEMLNYRLTPSYEYQLDPWHTHVWKGVNGTPTKKRAIGFKKLAKAVKFSAKLMGHAMAKRVKATILFATETGKSQDYAKTLCEIFKHAFDPKVM 775

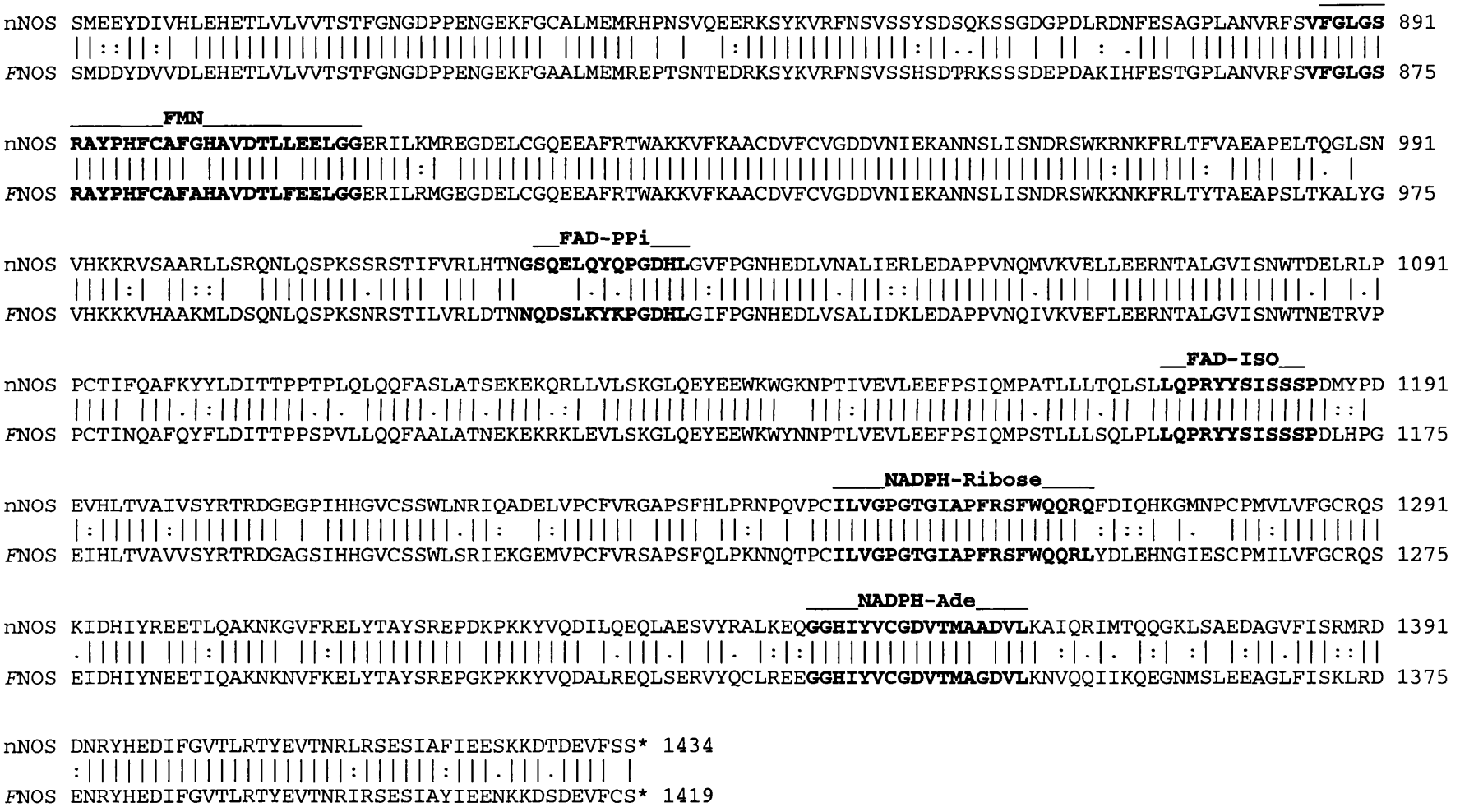


Figure 4.7. Alignment of the *Fugu* NOS1 predicted amino acid sequence (*F*NOS)with human nNOS. Domains are highlighted in bold and underlined. | represents identical amino acids and : represents a conservative substitution.

positions of binding domains, previously identified in the neuronal isoform, which as discussed previously in chapter 3, appear to be conserved in the *F*NOS sequence. Interestingly, the invertebrate *D*NOS amino acid sequence shows 47% identity to *F*NOS and between 46% and 48% identity to the human NOS isoforms, as measured by CLUSTALW alignment The identity shared by *F*NOS and human NOS1 is much higher than this and would suggest that while the ancestral *D*NOS has common core sequences the sequences of the vertebrate NOS isoforms have evolved to become more specialised. Figure 4.8 shows the GAP alignment of the two *F*NOS amino acid sequences as predicted from the genomic and cDNA nucleotide sequence data. In total, 27 nucleotide changes are present between the *Fugu rubripes*, *Fugu* NOS1 gene sequence and the *T. poecilonotus*, *f*NOS cDNA sequence (appendix, figure A.4). Of the nucleotide changes, 19 are silent, with no effect on the predicted amino acid composition of *F*NOS, whereas 8 result in amino acid differences. Sequencing of multiple isolates of the full-length *f*NOS cDNA, and the *Fugu* NOS1 gene, was performed to ensure no PCR errors had been introduced, for the same reason a proof-reading *Taq* polymerase was used for all PCR amplifications. The predicted amino acid sequences are 99.4% identical and the 0.6% difference most likely represents a species difference for this sequence.

4.6. Summary.

A full-length *Fugu* NOS1 cDNA (*f*NOS) was isolated from the total brain RNA of the Japanese pufferfish *T. poecilonotus* and cloned into the plasmid vector pVL1393. Sequencing of multiple isolate of *f*NOS verified the *Fugu* NOS1 gene structure discussed in chapter 3 and revealed a species difference for this sequence between the two Japanese pufferfish *T. poecilonotus* and *Fugu rubripes*. GAP alignment and phylogenetic analysis of *f*NOS sequence also verified the identification of the *Fugu* NOS1 gene and *f*NOS as neuronal NOS isoforms.

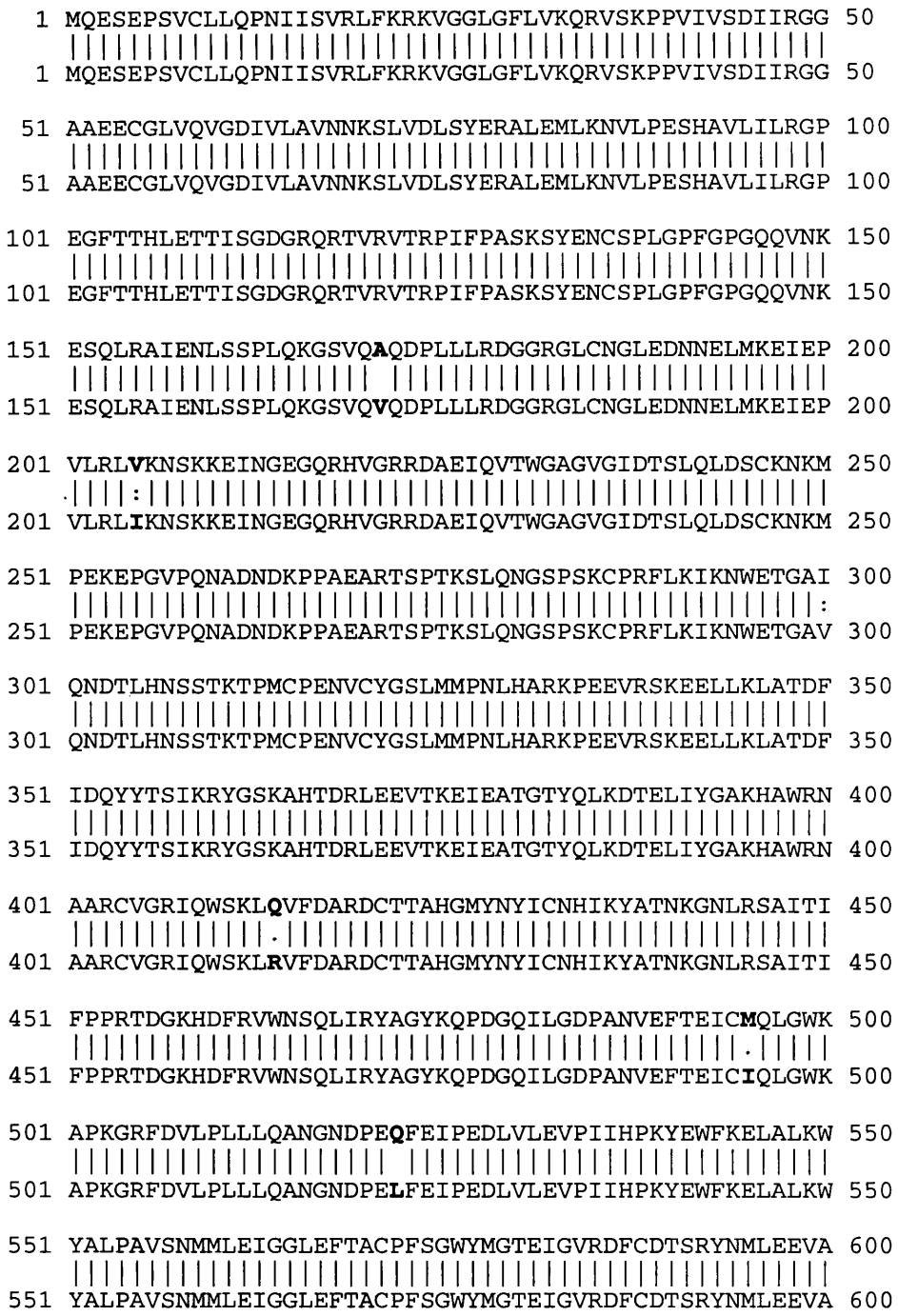


Figure 4.8. Alignment of the predicted *F*NOS amino acid sequences from genomic and cDNA sequence data. Genomic sequence data from *Fugu rubripes* and cDNA sequence data from *T. poecilonotus* were translated into predicted amino acid sequences and aligned using the GAP alignment programme. Amino acid sequence differences are highlighted in bold type.

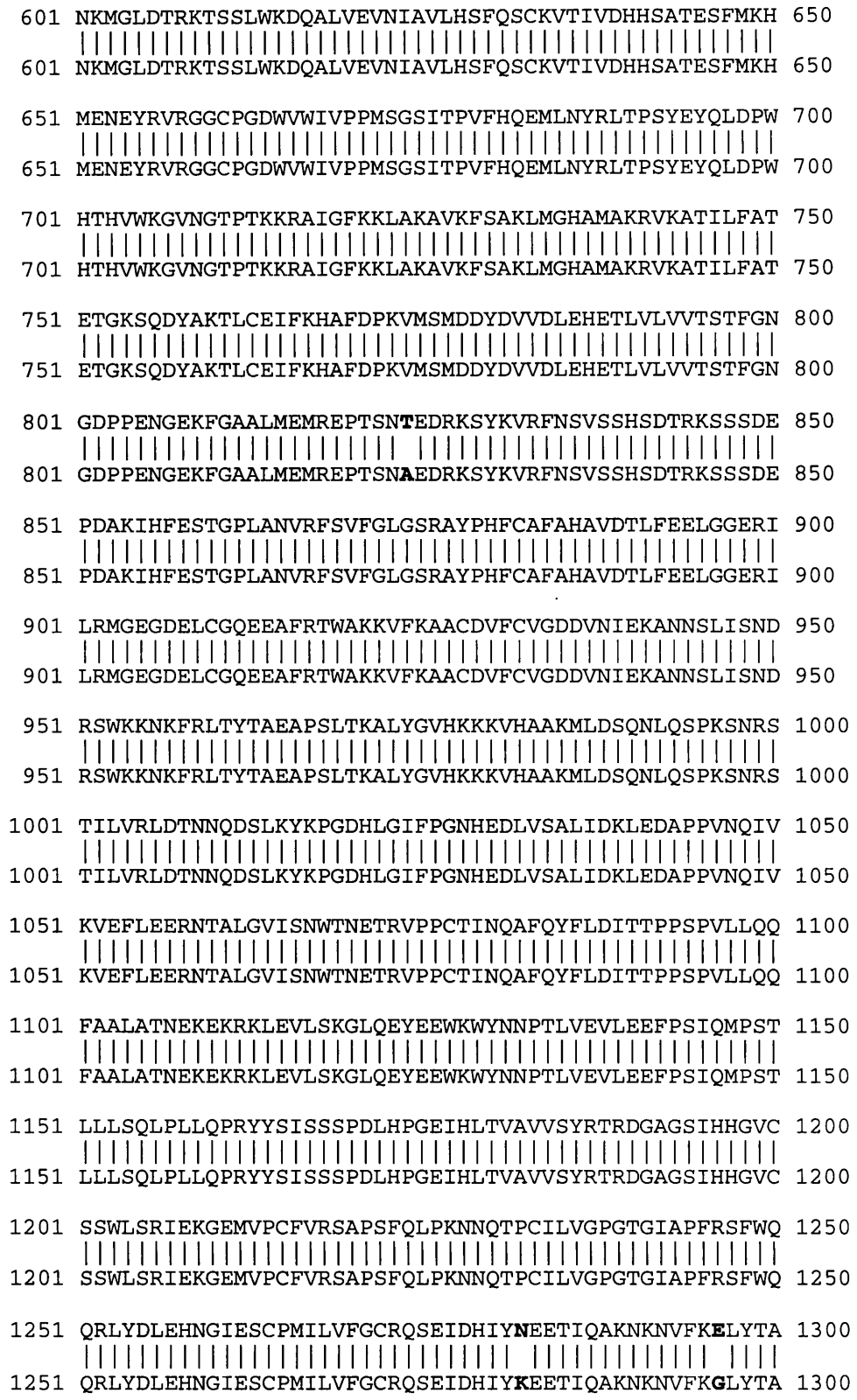


Figure 4.8. continued.

```
1301 YSREPGKPKKYVQDALREQLSERVYQCLREEGGHIYVCGDVTMAGDVLKN 1350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1301 YSREPGKPKKYVQDALREQLSERVYQCLREEGGHIYVCGDVTMAGDVLKN 1350

1351 VQQIIKQEGNMSLEEAGLFISKLRDENRYHEDIFGVTLRTYEVTNRIRSE 1400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1351 VQQIIKQEGNMSLEEAGLFISKLRDENRYHEDIFGVTLRTYEVTNRIRSE 1400

1401 SIAYIEENKKDSDEVFCS* 1419
     ||||||||||||||||||
1401 SIAYIEENKKDSDEVFCS* 1419
```

Figure 4.8. continued.

CHAPTER 5

EXPRESSION AND ACTIVITY OF RECOMBINANT *FUGU* NOS (*F*NOS)

CONTENTS.

5.1. Introduction.

For the purpose of biochemical characterisation, the *Fugu* NOS enzyme (*F*NOS) was expressed in a baculovirus-insect cell system. This was the system of choice because recombinant proteins are processed, modified, targeted and generally expressed at a higher level than in other eukaryotic expression systems. Also, the expression of mammalian NOS enzymes has been well established in this system (Charles *et al.*, 1993; Moss *et al.*, 1995; Riveros-Moreno *et al.*, 1995, Charles *et al.*, 1996). The baculovirus used for this study was the *Autographa californica* Nuclear Polyhedrosis Virus (AcNPV). AcNPV infects lepidopteran species including the fall armyworm *Spodoptera frugiperda*. In most cases, clone 21 (*Sf* 21) cells are used for expression. cDNAs to be expressed must first be integrated into the AcNPV genome, this is not possible by direct cloning due to the large size of the AcNPV genome (128 kb). Instead, cDNAs are cloned into the multiple cloning site of plasmid-like 'transfer-vectors', such as pVL1393, which direct the *in vivo* homologous recombination of the cDNA into the AcNPV genome. pVL1393 is a transfer vector which, when integrated into the AcNPV genome, complements the lethal deletion of the polyhedrin gene present in the vector used for this study (Baculo Gold linearised baculovirus DNA; Pharmingen). Polyhedrin is a protein highly expressed by AcNPV in the second half of its biphasic-life cycle, it encapsulates mature-virus particles within the nucleus of infected cells and is essential for lateral-viral transmission. In tissue-culture cells, polyhedrin is not required for viral transmission and can therefore be deleted and replaced by heterologous proteins, in this case *F*NOS.

5.2. Identification of recombinant baculovirus producing nitrite.

Recombinant baculoviruses, containing the full-length coding sequence for *f*NOS, were constructed by transfecting *Sf*21 cells with pVL*f*NOS (figure 4.4, chapter 4) and linearised baculovirus DNA as described in section 2.19. Four individual recombinant baculovirus were identified by limiting dilution plaque assay. High titre recombinant baculovirus stocks were generated by infecting approximately 5 x 10^7 *Sf*21 cells with 2.5 ml of viral supernatant, from the previous round of bulking up, for 96 hours.

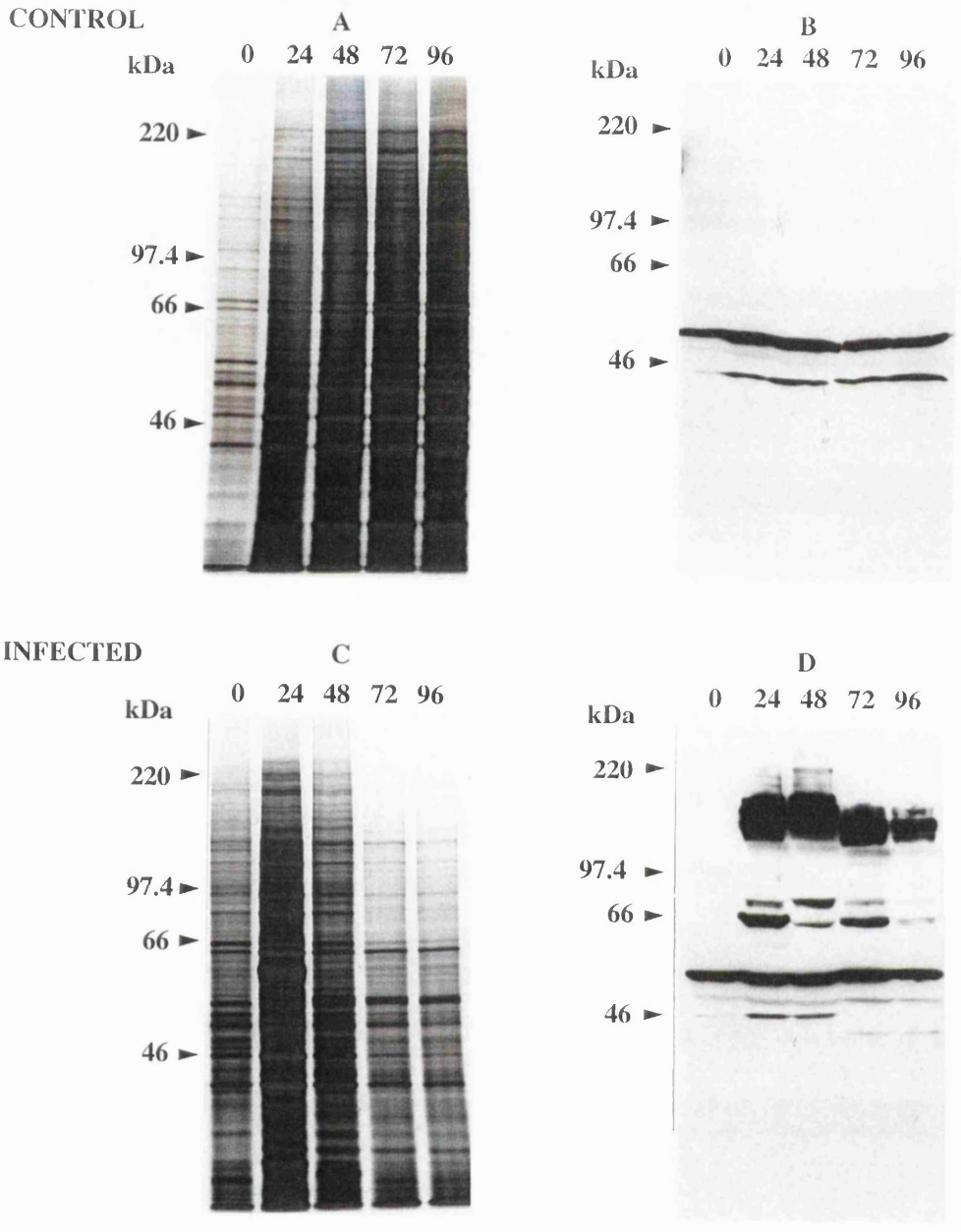
In biological systems NO has a very short half life and is susceptible to oxidation (Henry *et al.*, 1991; Ignarro *et al.*, 1993; Radi *et al.*, 1993; Beckman *et al.*, 1994). However, as a product of NO oxidation, the presence of nitrite (NO_2^-) can be reliably used as an indication of NO synthesis. Therefore, the NO_2^- concentration in each of the culture supernatants was measured using a chemiluminescence assay (section 2.20) (Downes *et al.*, 1976). In this assay, NO_2^- is first reduced back to NO and measured by the light emitted by the gas-phase reaction of NO with ozone. This can then be used to calculate the amount of NO_2^- present in the supernatants by comparison to a $NaNO_2$ standard curve. The chemiluminescence assay demonstrated that each of the insect cell cultures infected with recombinant baculovirus described above were producing NO_2^- in approximately the same quantities, clone 1 had generated ~160 μM, clone 2 ~145 μM, clone 3 ~170 μM and clone 4 ~165 μM. The presence of nitrite in the supernatants indicated that *F*NOS was being functionally expressed in the insect cells. Clone 1 was chosen for the further biochemical analysis of *F*NOS.

5.3. Expression of *F*NOS.

5.3.1. Time-course of *F*NOS expression.

A time-course of *F*NOS expression after 0, 24, 48, 72 and 96 hours of infection was performed as described in section 2.19. SDS-PAGE and western blot analysis of the lysates from the time-course (Figure 5.1) revealed that highest level of *F*NOS expression was seen after 24 hours. There was no obvious *F*NOS expression band seen on silver stained SDS-PAGE gels (section 2.21 and 2.22) at the expected molecular weight of ~159kDa. This is in contrast to the 160 kDa band visualised for recombinant rat nNOS expressed in the baculovirus-insect cell system (Charles *et al.*, 1993) but comparable to recombinant murine iNOS expressed in the same system (Moss *et al.*, 1995). At the 24 and 48 hour time points, western analysis with a mouse anti-human nNOS monoclonal primary antibody and a goat anti-mouse IgG secondary antibody (section 2.23), detected two major expression bands of approximately the right size. At 72 and 96 hours only the smaller of the two bands was detected. The presence of two bands at 24 and 48 hours may represent the processing or degradation of *F*NOS in insect cells. In addition, in the infected cells several lower molecular weight bands were detected by western analysis, these may also represent breakdown products of the full-length *F*NOS. A common protein was detected at each time-point for the infected and control cells and was therefore considered to be non-specific.

Figure 5.1. SDS-PAGE and western blot analysis of time-course of *F*NOS protein production from control and pVL*f*NOS-infected *Sf*21 cells. SDS-PAGE profiles (A & C) and corresponding western blots (B & D) of a time-course of pVL*f*NOS-infected and control *Sf*21 cells. Cytosols were electrophoresed through 8% acrylamide gels at a constant current of 5 mA/cm gel length. Gels were silver stained, dried and photographed. Blots were performed, using an anti-nNOS primary antibody, and developed as described. Lane numbers indicate hours after infection. The positions of molecular weight markers are indicated.


CONTROL
A
kDa
0 24 48 72 96
220
97.4
66
46
B
kDa
0 24 48 72 96
220
97.4
66
46
INFECTED
C
kDa
0 24 48 72 96
220
97.4
66
46
D
kDa
0 24 48 72 96
220
97.4
66
46

5.3.2. Antibody specificity of *F*NOS.

Further western blot analysis, using monoclonal and polyclonal antibodies raised against all three NOS isoforms, showed that detection of *F*NOS expressed in insect cells was only possible with anti-nNOS antibodies (Table 5.1) and that the monoclonal antibody was more sensitive than the polyclonal antibody. This data was in agreement with the genomic and cDNA sequence which identified the *Fugu* NOS gene as NOS1.

Table 5.1. Western blot detection of *F*NOS with anti-NOS antibodies.

Isoform	monoclonal	polyclonal
neuronal	++	+
inducible	-	-
endothelial	-	-

5.4. Nitric oxide synthase activity of *F*NOS.

The spectrophotometric method described in section 2.24 (Feelisch and Noack, 1987) was used to measure the conversion of oxyhaemoglobin to methaemoglobin by NO between the wavelengths 401 and 421 nm.

5.4.1. Time-course of *F*NOS activity.

As a preliminary experiment, the activity of *F*NOS from 0 to 96 hours was measured using the spectrophotometric assay for NOS activity (section 2.24). However, as the time-course was not performed in arginine-free medium the absolute activity could not be calculated. The activity at each time point was calculated as a percentage of the maximum activity seen. NOS activity was absent in lysates of uninfected control cells, however it was seen in lysates of infected cells, with maximum activity occurring 24 hours after infection (figure 5.2).

5.4.2. *F*NOS dose-response to L-arginine and inhibition.

A large-scale 24 hour infection of *Sf*21 cells, grown in arginine-free medium, was performed specifically for the accurate measurement of *F*NOS activity. A cytosolic preparation of the infected *Sf*21 cells was treated with Dowex AG50 anion exchange resin to remove any residual arginine (section 2.24). The protein concentration of the cytosol was calculated as ~11 mg/ml (section 2.25) and this value was used to convert all assay results from µl sample to mg protein. The reaction volume used for the assay was 600 µl and each sample always contained a final concentration of 0.1 M HEPES/0.1 mM DTT buffer pH7.5, 100 µM reduced NADPH, 5 µM BH_4, 5 µM oxyhaemoglobin and 4 µM FAD. For each reaction, 5 µl of cytosol was added.

NO synthesis was initiated by the addition of L-arginine and measured over the two wavelengths 401 and 421 nm using a dual wavelength spectrophotometer. Figure 5.3 shows the *F*NOS arginine dose-response curve which was constructed by initiating NO

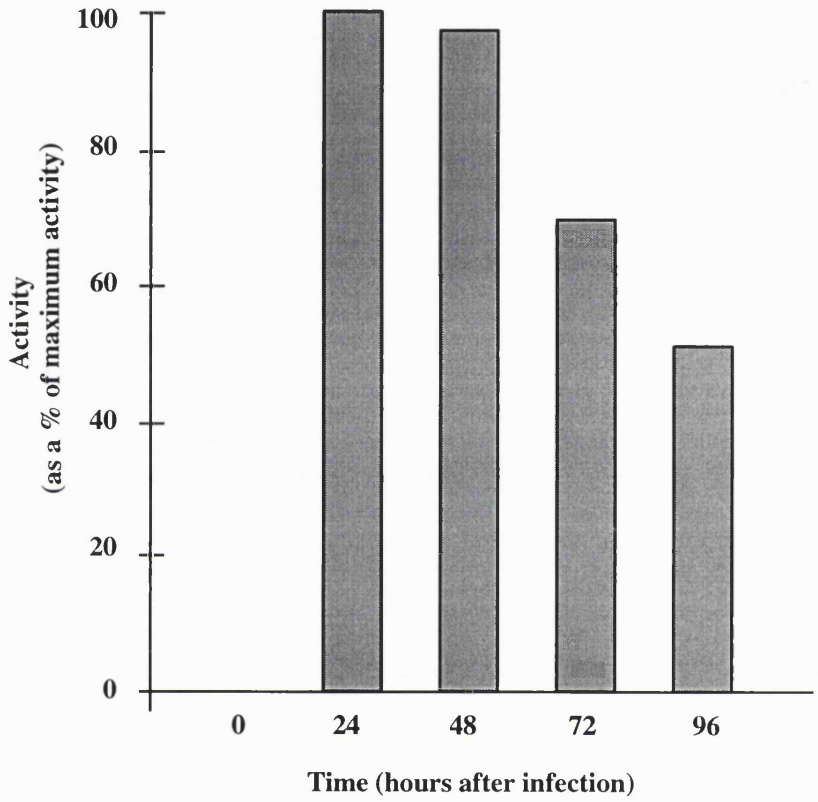


Figure 5.2. Time-course of *F*NOS activity. Spectrophotometric measurement of *F*NOS activity of cytosols from pVL*f*NOS-infected *Sf*21 cells over a 96 hour time-course. For this experiment, complete medium was used, therefore activities were calculated as a percentage of the maximum activity (24 hours) and the results are n = 1.

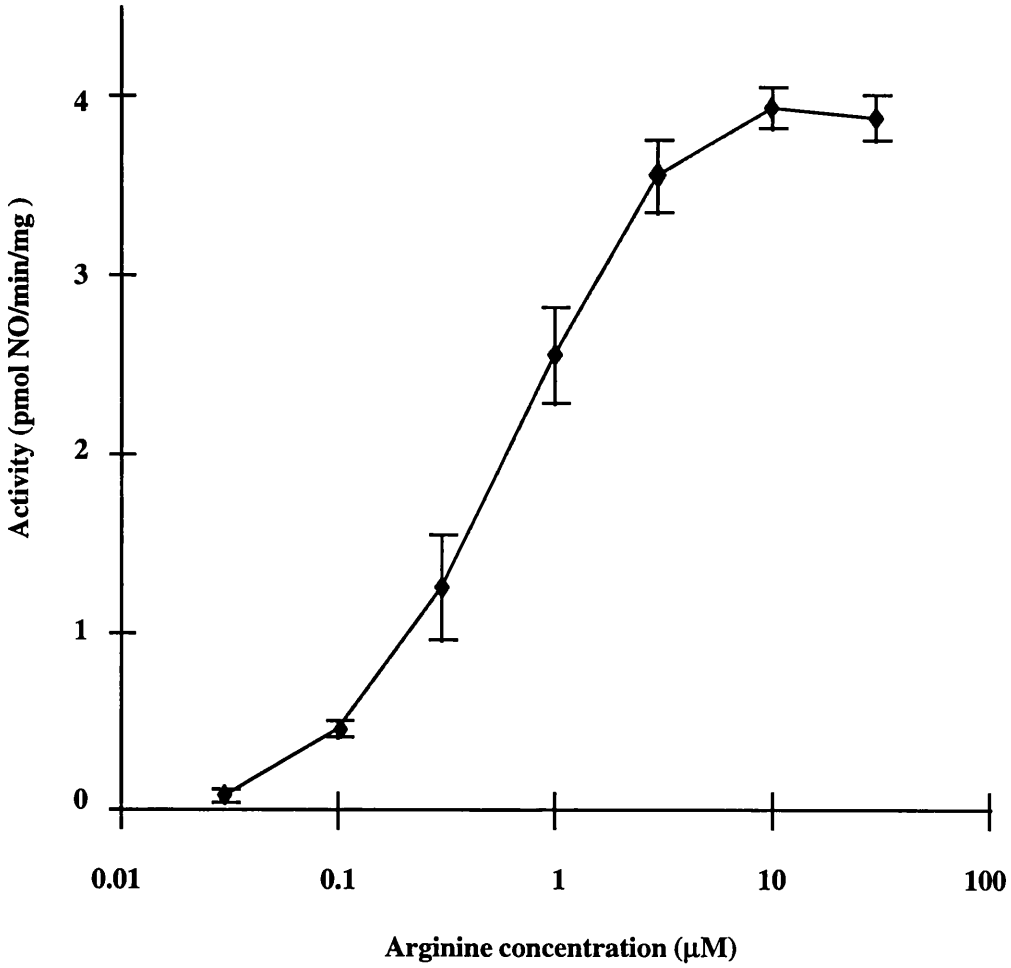


Figure 5.3. L-arginine requirements of *F*NOS from pVL*f*NOS-infected *Sf*21 cells.

Response of *F*NOS to increasing concentrations of L-arginine. All measurements are the mean of three independent observations +/- one standard error of the mean (SEM).

synthesis with varying concentrations of L-arginine (final concentrations of 0.1 μM to 30 μM). Maximum *F*NOS activity was seen at a final concentration of 10 μM L-arginine and a half-maximal enzyme rate was seen at a final concentration of ~0.79 μM L-arginine. Exogenous calmodulin (CaM) did not have to be added for *F*NOS activity to be initiated. This Ca^{2+}/CaM-independent behaviour is normally associated with iNOS and not the constitutive NOS isoforms however, the addition of 1 mM EGTA, a calcium chelator, completely abolished *F*NOS activity (data not shown). This data suggests *F*NOS is Ca^{2+}-dependent, as is reported for the constitutive mammalian NOSs.

The effect of the addition of NOS inhibitors on NO synthesis by *F*NOS, in the presence of 30 μM L-arginine, is shown in Figure 5.4. All three inhibitors tested showed complete inhibition of *F*NOS activity at a final concentration of 300 μM. This is in contrast to the different concentrations of each inhibitor required to reduce *F*NOS enzyme activity by 50% (IC_{50}). L-NMMA had an IC_{50} of ~0.63 μM, 1400W ~1.53 μM and L-Thiocitrulline ~15.3 μM.

The response of *F*NOS to the addition of L-arginine, and its inhibition by known NOS inhibitors, demonstrates that recombinant *F*NOS is being expressed as a functional enzyme and supports the result of the chemiluminescence assay which simply indicated the presence of NO_2^- (section 5.2).

5.5. Immunocytochemical detection of *F*NOS expressed in insect cells.

The localisation of NOS has traditionally been performed using the histochemical technique of NADPH-diaphorase (NADPHd) staining which relies on the electron transfer from NADPH to nitroblue tetrazolium (NBT), by the C- terminal portion

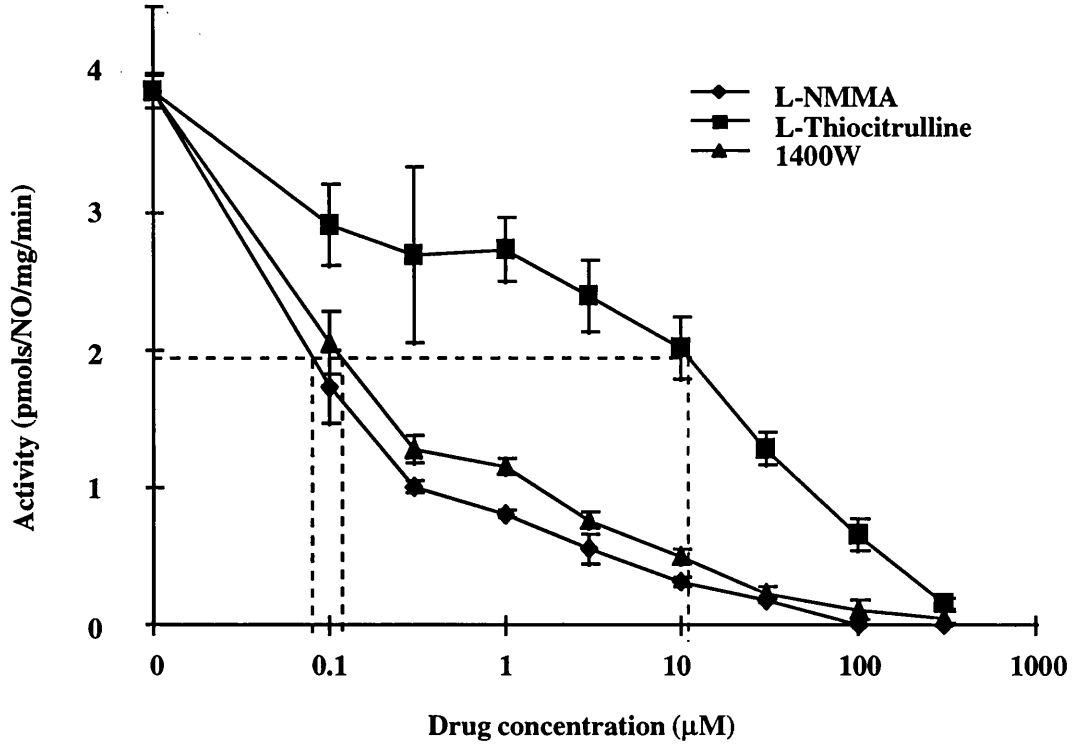


Figure 5.4. Inhibitor analysis of *F*NOS from pVL*f*NOS-infected *Sf*21 cells. Activity of *F*NOS in response to L-NMMA (filled diamonds), L-Thiocitrulline (filled squares) and 1400W (filled triangles) in the presence of 30μM L-Arginine. All measurements are the mean of three independent observations +/- one standard error of the mean (SEM). Dashed lines represent the IC_{50} value of each drug for *F*NOS.

of NOS (Norris *et al.*, 1994). The reduction of NBT results in the formation of a insoluble purple formazan product which is characteristic of NADPHd staining. However, although there is good correlation between sites of NADPHd activity and nNOS in the CNS (Bredt *et al.*, 1991; Dawson *et al.*, 1991), the ability to reduce tetrazoliums is not exclusive to NOS and outside the CNS NADPHd staining is a much less reliable NOS marker. Immunohistochemical detection of NOS with antibody probes has allowed increased specificity of NOS localisation. The antibodies used are described in table 2.6, section 2.23. The secondary antibodies used, goat anti-mouse or anti-rabbit IgG, were conjugated to Alexa 488 in place of fluorescein. Alexa 488 is a green fluorophore with the same spectral characteristics as fluorescein but with increased brightness and photostability.

A time-course of infection with *F*NOS was set up using *Sf*21 cells adhered onto 13 mm glass coverslips, as described in section 2.26. Cells were fixed with paraformaldehyde, permeabilised with Triton X-100 and non-specific protein interactions blocked with insect cell medium containing 10% FCS. Figure 5.5 shows the detection of *F*NOS in *Sf*21 cells by the mouse anti-human nNOS monoclonal antibody. The uninfected control cells were negative for NOS immunofluorescence over the 96 hour time-course. NOS immunoreactive protein was detected in the infected *Sf*21 cells 24 to 96 hours after infection but not 0 hours after infection. After 24 hours infection, when the cells still looked healthy and intact, NOS immunofluorescence was at a maximum and appeared to be distributed throughout the cytosol. By 96 hours post-infection, when there was membrane disruption of the cells, the localisation of the NOS immunoreactive protein had moved to the nucleus. The co-localisation of the green NOS immunofluorescence with the red propidium iodide nuclear stain accounts for appearance of the yellow

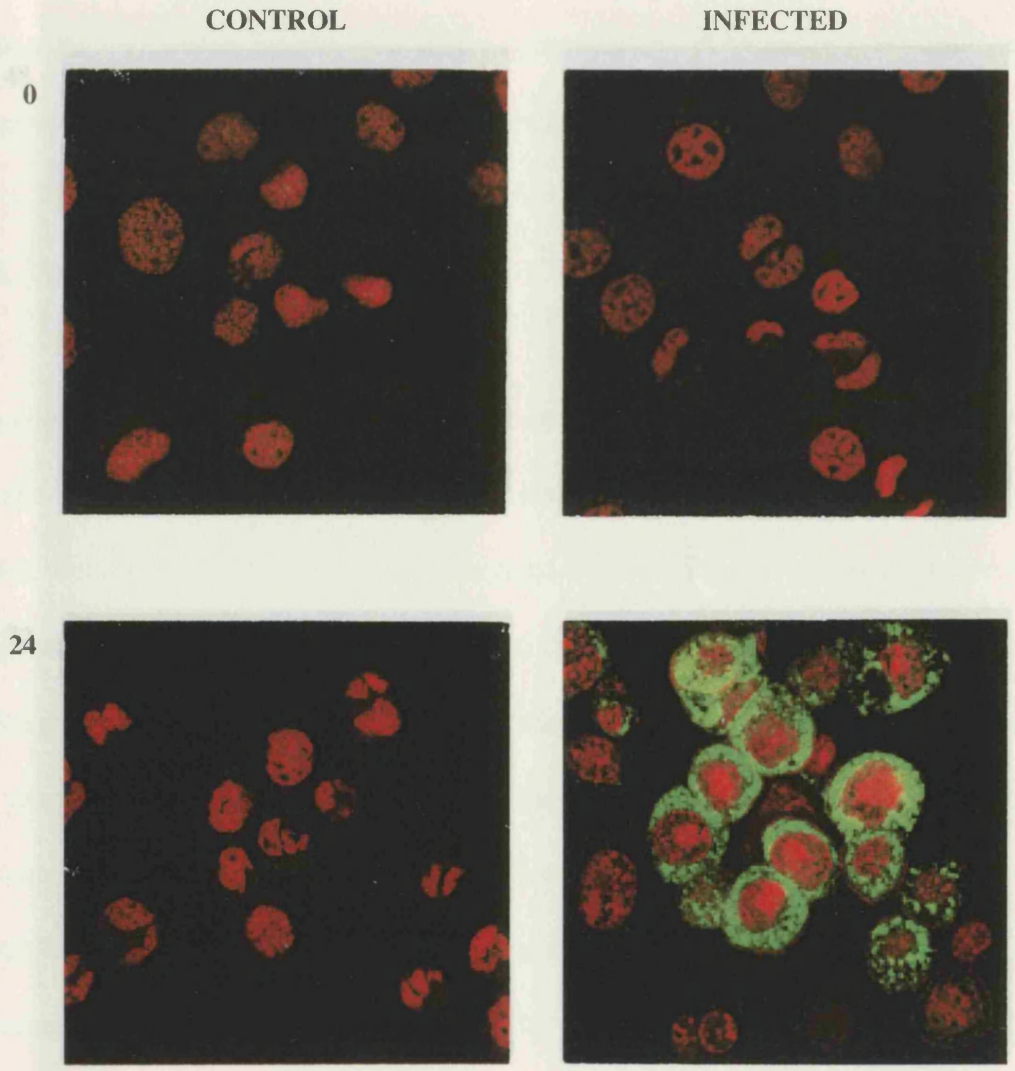


Figure 5.5. Immunocytochemical detection of *F*NOS in control and pVL*f*NOS-infected *Sf*21 cells over a 96 hour time course. Sf21 cells were adhered to 13mm coverslips, fixed and permeabilised as described. *F*NOS staining is visualised by the Alexa 488 green fluorophore. Cell nuclei are stained red by propidium iodide. Numbers indicate the time in hours that the cells were fixed following infection

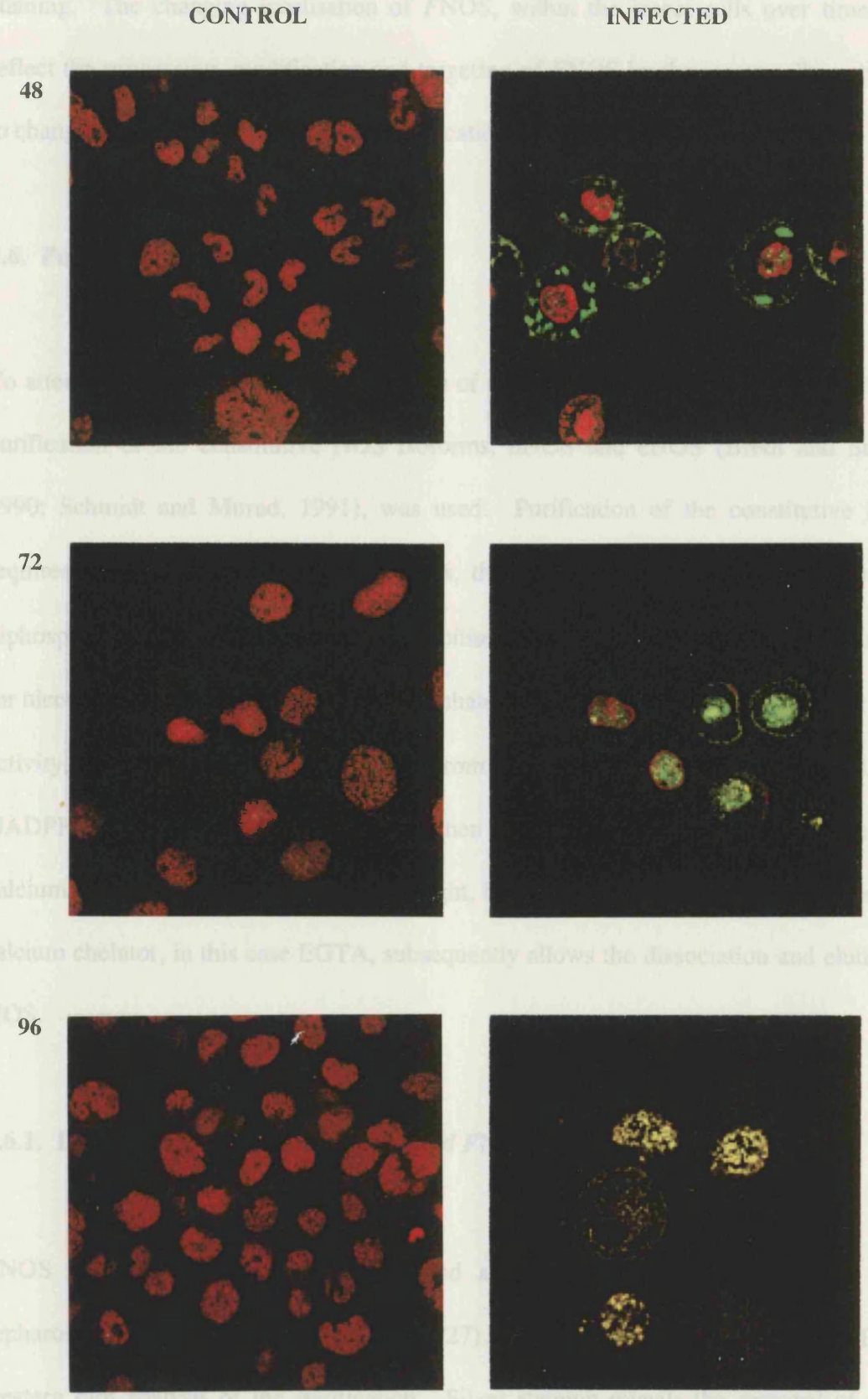


Figure 5.5. continued.

staining. The changing localisation of *F*NOS, within the insect cells over time, may reflect the processing, modification and targeting of *F*NOS by the insect cells or be due to changes directed by the baculovirus replication within the cell.

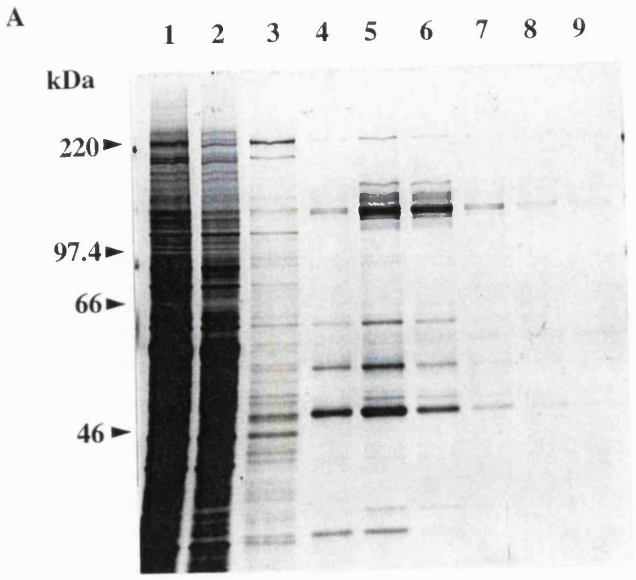
5.6. Purification of *F*NOS.

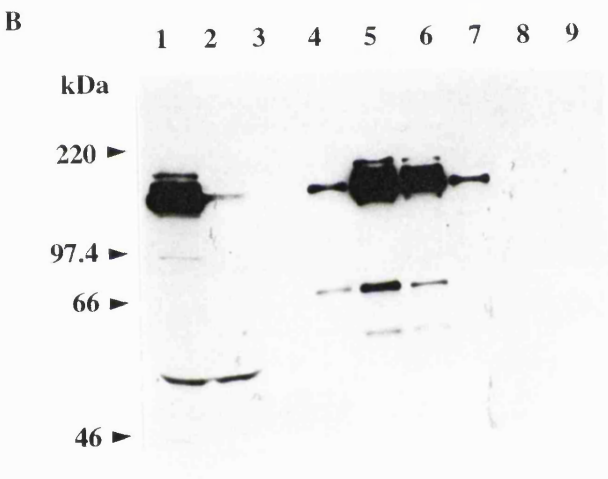
To attempt to purify *F*NOS, a modification of the procedure originally described for the purification of the constitutive NOS isoforms, nNOS and eNOS (Bredt and Snyder, 1990; Schmidt and Murad, 1991), was used. Purification of the constitutive NOSs requires two affinity-chromatography steps, the first using immobilised 2′5′ adenosine diphosphate (ADP) and the second, immobilised CaM. 2′5′ ADP is a structural mimic for nicotinamide adenine dinucleotide phosphate (NADPH), a cofactor required for NOS activity, and therefore elution of NOS from 2′5′ ADP resin can be achieved with NADPH. Constitutive NOS isoforms are then bound to a CaM resin in the presence of calcium, which allows the formation of a tight, but reversible complex. The addition of a calcium chelator, in this case EGTA, subsequently allows the dissociation and elution of NOS.

5.6.1. Ligand affinity-chromatography of *F*NOS on 2′5′ ADP sepharose.

*F*NOS was partially purified using ligand affinity-chromatography with 2′5′ ADP sepharose as the affinity matrix (section 2.27). Figure 5.6 shows the SDS-PAGE and western blot analysis of the purification. Silver staining reveals the removal of many proteins away from *F*NOS, seen as flow-through in lanes 1 and 2 on figure 5.6, some of

Figure 5.6. SDS-PAGE and western blot analysis of fractions isolated from 2′5′ ADP sepharose chromatography of cytosol from pVL*f*NOS-infected *Sf*21 cells. Samples were electrophoresed at a constant current of 5 mA/cm gel length through 8% acrylamide gels. Gels were silver stained, dried and photographed (A). Blots were performed with an anti-nNOS primary antibody, developed as described and photographed (B). **Lane 1:** crude cytosol. **Lane 2:** column wash-through (fraction 1). **Lane 3:** 0.5 M NaCl eluate (fraction 5). **Lanes 4-9**: 10 mM NADPH eluate (fractions 8-13). The positions of molecular weight markers are indicated.


A
1 2 3 4 5 6 7 8 9
kDa
220
97.4
66
46


B
1 2 3 4 5 6 7 8 9
kDa
220
97.4
66
46

which cross-reacted with the monoclonal anti-nNOS antibody used for the western analysis. This included a band of ~159 kDa, the expected size for *F*NOS, which indicated that the 2′5′ ADP sepharose column used for the purification was fully saturated by *F*NOS and any other proteins with affinity to 2′5′ ADP. Washing the affinity column with 0.5M NaCl removed further proteins (lane 3) but this did not include an immunoreactive protein of ~159 kDa. Elution of *F*NOS from the column was seen on addition of NADPH (lanes 4-9) along with a number of other proteins none of which cross-reacted to the anti-nNOS monoclonal antibody. Of the proteins eluted with NADPH, only 4 bands were detected by western blot analysis, including a major band seen at ~159 kDa, a slightly larger band and two smaller bands (less than 66 kDa).

5.6.2. Ligand affinity-chromatography of *F*NOS on CaM sepharose.

To further purify *F*NOS, following the 2′5′ ADP sepharose purification, a second ligand affinity-chromatography step was performed using calmodulin (CaM) sepharose. Fractions containing *F*NOS, as determined by SDS-PAGE and western analysis (figure 5.6), were loaded onto a CaM sepharose column as described in section 2.28. The calcium chelator EGTA was added to the column at a concentration of 3 mM to try and elute *F*NOS from the column and this was then followed by 10 mM EGTA. SDS-PAGE of the eluted fractions suggested that *F*NOS was still bound to the column and subsequent SDS-PAGE analysis of a sample of column material itself verified that this was the case (data not shown). Even after the addition of 1 M NaCl and 6 M urea, *F*NOS remained bound to the CaM sepharose column (data not shown).

5.7. Summary.

A functional *F*NOS enzyme has been expressed in an insect cell-baculovirus expression system. The expression, activity and immunocytochemical detection of *F*NOS, over a 96 hour time-course, demonstrated that *F*NOS expression was at its highest level 24 hours after infection. Activity of *F*NOS, in the presence of NADPH, BH_4 and HEPES:DTT buffer, was initiated by L-arginine but did not require the addition of exogenous calmodulin. The addition of EGTA completely abolished NOS activity suggesting that *F*NOS was Ca^{2+} dependent but already bound to calmodulin. The most potent inhibitor of *F*NOS was L-NMMA and surprisingly, the iNOS selective inhibitor, 1400W was a more potent inhibitor than the nNOS selective inhibitor L-Thiocitrulline. Partial purification of *F*NOS was achieved by 2′5′ ADP ligand affinity-chromatography but further purification with CaM ligand affinity-chromatography was not possible.

CHAPTER 6

DISCUSSION

CONTENTS.

6.1. Molecular characterisation of a *Fugu* NOS1 gene.

The experiments described in chapter 3 detail efforts to clone and characterise NOS genes from the genome of the Japanese pufferfish *Fugu rubripes*. Screening of the *Fugu* genomic cosmid library resulted in the isolation of a *Fugu* NOS1 gene.

The NOS1 gene isolated from *Fugu* was found to have the same organisation as the human NOS1 gene previously cloned and characterised by Hall and colleagues in 1994. With the exception of exon 1, all of the *Fugu* NOS1 exons were identified by direct comparison to the human NOS1 gene sequence. Exon 1 was identified by comparison of the genomic sequence data upstream of exon 2 with 5′ RACE sequence data obtained as part of the experiments detailed in chapter 4. The 5′ donors of all the *Fugu* NOS1 introns begin with the consensus GT, except for intron 22 which begins with GC, and all the 3′ acceptors terminate with the consensus AG, and in all cases the intron type, whereabouts the splice junction occurs in a codon, is conserved between the *Fugu* and human NOS1 genes.

The *Fugu* NOS1 gene spans only 22.2 kb, when compared to the approximately 160 kb covered by the human NOS1 gene (Hall *et al.*, 1994). This represents a reduction in size of at least 7.2 times which is in line with the overall reduction in the size of the *Fugu* genome (Brenner *et al.*, 1993). At 4257 bp, the length of the predicted coding sequence of *Fugu* NOS1 is comparable to the 4302 bp of human NOS1 and this is consistent with many of the *Fugu* genes that have been cloned and characterised to date (Mason *et al.*, 1994; Baxendale *et al.*, 1995; Yeo *et al.*, 1997). Differences in the lengths of exons 2, 3 and 4 account for the 45 bp, 15 amino acid, size difference between the *Fugu* and human NOS1 coding sequences. Therefore, as for other *Fugu* genes, the majority of the size

reduction of the *Fugu* NOS1 gene is due to a reduction in the amount of intronic DNA (Elgar *et al.*, 1996).

It was reported previously that 80% of *Fugu* introns are less than 150 bp long (Elgar *et al.*, 1996). In contrast to this, only 36% of the *Fugu* NOS1 gene introns are less than 150 bp long. This phenomenon has been reported for several other *Fugu* genes including the *Fugu* TUPLE1/HIRA, Huntington's disease (*FrHD*) and Neurofibromatosis Type 1 (NF1) genes, where only 50, 70 and 75% of introns respectively are less than 150 bp long (Baxendale *et al.*, 1995; Kehrer-Sawatzki *et al.*, 1998; Llevadot *et al.*, 1998).

To substantiate this finding, Villard and colleagues (1998) demonstrated that AT-rich genes (40% GC) have 3-fold larger introns than GC-rich genes. Furthermore, evidence from isochore analysis, that AT-rich regions are the most gene poor and GC-rich regions are the most gene dense, was used to predict that the base composition of a human gene could be used to estimate the relative compaction of the *Fugu* homologue. To illustrate this they cloned and characterised the *Fugu* homologue of the beta-amyloid precursor protein (APP) gene (Villard *et al.*, 1998). The human APP gene is an AT-rich (40% GC) gene which is mutated in some forms of familial Alzheimer's disease. *Fugu* APP was found to be approximately 30 times smaller than the human APP gene with a 50-fold average reduction in intron size. In comparison, the human HD gene is 45% GC-rich and the overall reduction in the size of the *FrHD* gene is approximately 7.4-fold. Another example, where the reduction in *Fugu* gene size is reflected by the GC content of the human equivalent, is the 47% GC-rich human p55 gene. In humans, the p55 gene spans 30 kb whereas in *Fugu*, the distance covered is approximately 6 kb and represents an overall reduction of between only 4 and 6-fold (Elgar *et al.*, 1995). There is no information available on the entire GC content of the human NOS1 gene, due to the lack

of sequence data for the large introns present in the gene, but the ~7.2-fold reduction in the size of the *Fugu* NOS1 gene suggests that it is GC-, rather than AT-, rich.

Although most *Fugu* introns are small, it appears to be a characteristic of *Fugu* genes that the first intron is large (Elgar *et al.*, 1996). Large introns are described as those over 400 bp long and in addition to the first large intron, a further one is also normally found located towards the middle of *Fugu* genes (Elgar *et al.*, 1995). The *Fugu* NOS1 gene actually has 13 large introns and 5 of these are over 1000bp. Consistent with other *Fugu* genes the first intron is 1322 bp long and another of 1161 bp is found towards the middle of the gene. It has previously been suggested that when introns are significantly longer than 100 bp in length an intron-specific function may be present (Cecconi *et al.*, 1996). However, database homology searches failed to identify any known regulatory elements within *Fugu* NOS1 intronic DNA. Additionally, comparison of the non-coding sequence of the human and *Fugu* NOS1 genes failed to detect any conserved regulatory elements between the two. The dramatic reduction in size of several of the *Fugu* NOS1 introns also suggests that these regions may not contain any essential function, for example intron 8 is reduced by 130-fold from 12 kb in the human NOS1 gene to just 92 bp in the *Fugu* NOS1 gene.

Originally, only exons 5 to 19 of the *Fugu* NOS1 gene were isolated from cosmid 064O07. In an attempt to isolate the remainder of the *Fugu* NOS1 gene upstream of exon 5, additional restriction fragments from cosmid 064O07 were cloned and sequenced (section 3.3). Database homology searches failed to identify any further restriction fragments with obvious NOS1 homology. Instead, coding-sequence from the *Fugu* homologue of the kinase suppressor of RAS (*ksr*) gene was identified. The *ksr* gene

encodes a protein kinase (KSR) that is an evolutionarily conserved component of RAS-dependent signalling pathways (Michaud *et al.*, 1997). It has been isolated from *Drosophila melangoster* and *Drosophila virilis*, *C. elegans*, humans and mice (Kornfeld *et al.*, 1995; Sundaram and Han, 1995; Thierren *et al.*, 1995).

These findings indicate that the *Fugu ksr* gene lies on the same cosmid as the *Fugu* NOS1 gene. Assuming that cosmid 064O07 contains a contiguous stretch of DNA, and taking into consideration that the *Fugu* NOS1 gene is not present beyond exon 19 in cosmid 064O07, then the *ksr* gene must lie upstream of NOS1. Additionally, if cosmid 064O07 contains an insert of 40 kb, the average for the *Fugu* genomic library constructed (Elgar *et al.*, 1999), then the *ksr* gene would have to lie within approximately 20 kb of the NOS1 gene. Unlike the human NOS1 gene, which has been mapped to chromosome 12q24.2, the chromosomal location of the human *ksr* gene is unknown. Considering the conserved synteny that is seen across large regions of the human and *Fugu* genomes, it is possible that the *ksr* gene may also map to human chromosome 12. However, evidence for conserved linkage and gene order is not so compelling and therefore it is not possible to predict whether the *ksr* gene would also lie upstream of the NOS1 gene in the human genome.

Thierren and colleagues (1995) identified the mouse homologue of *ksr* by a database homology search. The partial open reading frame was named *hb* and they used this sequence as a probe to screen a mouse cDNA library. They isolated a 4 kb cDNA clone that encoded a protein of 873 amino acids which they named mKSR1 (Thierren *et al.*, 1995). The *hb* sequence had originally been isolated as part of an exon-trapping strategy to establish the transcription map of a 1 Mb region around the mouse NF1 locus (Nehls *et al.*, 1995; Genbank accession number X81634). The murine NF1 gene maps to

chromosome 11 and the human NF1 gene to chromosome 17q11.2. In contradiction to the idea that the human *ksr* gene would be present in the vicinity of the NOS1 gene on chromosome 12, this data suggests that the human *ksr* gene should also lie on chromosome 17. Interestingly, the NOS2 gene maps 0.46 cM upstream of the NF1 locus on mouse chromosome 11 and in humans it has been mapped to 17q11-q12. Therefore, the genes for NF1 and NOS2 form a synteny group across the mouse and human genomes. A *Fugu* NF1 gene has also been characterised (Kehrer-Sawatzi *et al.*, 1998). Analysis of the NF1 locus in *Fugu* identified three additional genes, AKAP84, BAW and WSB1, 3′ of the NF1 gene (Kehrer-Sawatzki *et al.*, 1999). The human homologues of these genes are not in the immediate vicinity of the NF1 gene but do map to human chromosome 17. None of these genes is a *Fugu* NOS2 homologue however it is possible that a *Fugu* NOS2 homologue would lie 5′ of NF1 as it does in the mouse genome.

Taken together this data questions the conservation of synteny between *Fugu* and humans for the NOS1 and NF1 regions and the evolution of these genes and their respective loci. To be more certain it would be necessary to map the human *ksr* gene, to isolate a *Fugu* NOS2 homologue and see if it maps in the vicinity of the *Fugu* NF1 gene. Interestingly, the 148 kb stretch of genomic DNA analysed around the wnt1 locus of *Fugu rubripes* identified eight genes which have all been mapped to human chromosome 12q13 upstream of NOS1 (Gellner and Brenner, 1999). The wnt1 and wnt10b genes identified are tightly linked within 2.3 cM on murine chromosome 15 and the *Drosophila* wnt1 homologue (wingless) and Dwnt4 are closely linked. In addition, the human wnt3 and wnt15 are both present on chromosome 17q21 within 125 kb of each other. These authors suggested that the linkage of the wnt1 genes occurred once in evolution and that

in the evolution of vertebrates this pattern was duplicated on several chromosomes. Taking this suggestion into consideration, and the fact that the wnt genes lie on both human chromosomes 12 and 17 as do human NOS1 and NOS2, it could be inferred that there was a single NOS gene that was also duplicated onto several chromosomes.

Analysis of the *Fugu* NOS1 sequence identified putative co-factor binding sites characteristic of the NOS family of enzymes. A more in-depth analysis of individual exons revealed differences in the degree of identity seen between the human and *Fugu* sequences. Five of the *Fugu* NOS1 exons show over 90% amino acid identity when compared to their human counterparts. Unsurprisingly, this includes exons which are known to have functional roles in the activity of the NOS enzyme.

Exon 12 has 92% amino acid identity and forms part of the dihydrofolate reductase (DHFR) homology fragment. The DHFR homology fragment binds the L-arginine analogue N^G-nitro-L-arginine (NNA) and is the location of the L-arginine binding site (Nishimura *et al.*, 1995). The remainder of the DHFR homology fragment is comprised of exons 10, 11 and part of exon 13. Although all three of these exons have over 80% identity to their human counterparts, no individual residues of importance for L-arginine binding have been identified in exons 11 and 13. In exon 10, however, six amino acid residues have been singled out as important for binding, as well as one in exon 8 which is also conserved. All except one of these residues, identified by Raman *et al.* (1998), are predicted to be present in the *Fugu* NOS1 gene, they are W358, E363, Q249, V338, F355 and Y359. The residue that differs between the bovine eNOS studied and the predicted *F*NOS amino acid sequence is N368 which was identified as important for amino acid carboxylate binding. In *Fugu* the equivalent residue is D368, this aspartic

acid residue is also present for the other vertebrate nNOS isoforms and the vertebrate iNOS isoforms. It is only an asparagine residue for the eNOS isoforms.

Further comparison has shown that exon 14 has 92% amino acid identity with human nNOS exon 14 and comprises part of the Ca^{2+}/CaM binding domain, whilst exon 18 has 93% amino acid identity and forms the region responsible for FMN binding. In addition, exons 7 and 8 show over 90% identity but to date have no defined functional role. These exons lie next to the haem binding domain, which is present in exon 6, and may therefore be structurally important.

Another 10 exons are over 80% identical at the amino acid level and include the remaining part of the Ca^{2+}/CaM binding domain and the FAD isoalloxizine and NADPH ribose-binding domains. Again the remainder have no defined functional role. Of the eight exons that have over 70% amino acid identity to their human counterparts, exons 16 and 17 form the CaM inhibitory sequence and exon 21 the FAD pyrophosphate-binding domain.

Two exons with ~60% amino acid identity are exons 4 and 27. Exon 27 contains the sequence for the NADPH adenine-binding domain and within this defined region the conservation of amino acids is good. In comparison, for exon 4 there are very few conserved amino acid residues. Of note are the aspartic acid and threonine residues which are conserved (D303 and T304). These residues are the equivalent of rat nNOS D314 and T315 which have been identified as important for the catalytic activation of molecular oxygen by rat nNOS (Sagami and Shimizu, 1998). Activation of molecular oxygen is required for the conversion of the NO intermediate N^G-hydroxy-L-Arg (NHA) to NO and L-citrulline. Site-directed mutagenesis of the D314 residue in rat nNOS

completely abolished NO formation activity and it was therefore proposed as a residue crucial for catalysis.

Exon 5 of *Fugu* NOS1 shows a predicted amino acid identity to human nNOS exon 5 of only 57%. However, this exon encompasses the zinc binding motif, C - $(X)_4$ - C, first identified by Raman *et al.* (1998), and further characterised by Miller and colleagues (1999). The two cysteine residues are conserved in all known NOSs and are indeed present in the predicted *F*NOS amino acid sequence in the same C-$(X)_4$-C arrangement. Also conserved in *F*NOS is a serine residue in the equivalent position of the bovine eNOS S104, this residue has been shown to form a direct hydrogen bond to the pterin side chain hydroxyl. However, the valine residue at position 106 in bovine eNOS, which has been shown to form a direct non-bonded contact with pterin, is not conserved in *F*NOS. The equivalent position is filled by a methionine residue which is also present in all other nNOS and iNOS isoforms.

The other exon showing low amino acid identity with its human equivalent, just over 50%, is exon 2. As is evident from the alignment of the predicted amino acid sequence for *F*NOS, translated from *f*NOS, and human nNOS this low conservation is due in part to the differing lengths of these two exons. Exon 2 in human NOS1 contains the sequence for the PDZ domain and this domain is also shown to be present in the *Fugu* sequence. Over the first 125 amino acids of *Fugu* and human nNOS sequences, which encompasses the PDZ domain, the amino acid similarity between the two rises to 80% and the identity to 67%. This suggests that *F*NOS has a PDZ domain which can form protein-protein interactions as described for other nNOS isoforms. This is not the first PDZ domain to be identified in *Fugu*. The human p55 gene is an erythrocyte membrane protein and was one of the first proteins to be identified as having a PDZ domain, like

nNOS it has a single PDZ domain which is conserved in the *Fugu* p55 homologue (Elgar *et al.*, 1995).

At 33%, exon 1 of *F*NOS shows limited amino acid identity to the human nNOS exon 1 reported with the gene sequence by Hall *et al.* (1994).

The lowest amino acid identity seen between the two sequences is for exon 3 which differs in *Fugu* for length and sequence, in total there are only 6 conserved amino acids.

A major theme in the regulation of nNOS expression is the structural and allelic mRNA diversity which has been found for human, mice and rats. Allelic diversity of human nNOS was reported for exon 1 where 4 alleles of an imperfect microsatellite were found on chromosome 12 from 36 individuals (Hall *et al.*, 1994). The structural variations found for nNOS include cassette exon deletions, cassette exon insertions, multiple use of exon 1 and promoters and the use of alternate polyadenylation signals.

To date nine exon 1 variants have been reported for human nNOS (Wang *et al.*, 1999). All nine variants arise from the use of alternative promoters of the single NOS1 gene, making NOS1 the most structurally diverse human gene, in terms of promoter usage, described thus far. These nine exon 1 variants, designated 1a, 1b, 1c, 1d, 1e, 1f, 1g, 1h, and 1i according to their 5′ to 3′ order over 105 kb upstream of exon 2, are all spliced to the common exon 2. The exon-intron boundaries of each variant conform to the GT/AG consensus and there are stop codons present prior to the AUG translation initiation codon in exon 2. Therefore, the protein sequence encoded is not affected and the diversity present is within the 5′ UTR of the nNOS mRNA. Analysis of the *Fugu* gene sequence upstream of exon 2, which totals 1515 bp and includes the exon 1 sequence identified by 5′ RACE PCR, did not identify any sequences homologous to the exon 1

variants. One possible explanation for this is the much smaller distance covered, taking into consideration the reduction in genome size, the equivalent of 105 kb of human DNA would be approximately 13 kb in *Fugu*. In addition to this, the single *Fugu* exon 1 identified by 5′ RACE PCR is smaller than and shows only 32.8% nucleotide identity to the exon 1 first reported for human nNOS (Hall *et al.*, 1994). This also presents the possibility of a more simple regulation of *f*NOS transcription as compared to mammalian nNOS.

The first cassette exon deletions described for human nNOS mRNA transcripts were 315 bp of exon 9 and 10 and 175 bp of exon 10 respectively which both result is a truncated nNOS protein (Ogura *et al.*, 1993; Hall *et al.*, 1994; Fujisawa *et al.*, 1994). Another example of a truncated nNOS protein is the one encoded by the nNOSβ transcript which entirely lacks exon 2 (Brenman *et al.*, 1996). Obviously, from the *Fugu* NOS1 gene sequence it is not possible to tell if any of these alternative nNOSs, resulting from cassette exon deletions, exist as that would only be possible by looking at mRNA from the various tissues where these transcripts have been found to be expressed.

The nNOSγ transcript, first identified in mice, lacks exon 2 but uses an alternative exon 1 (Brenman *et al.*, 1996). Translation is initiated from an ATG in exon 5 and results in a 125 kDa protein. Further studies demonstrated that in humans a testis-specific mRNA transcript existed, also named nNOSγ, and that translation was also initiated from an ATG codon in exon 5 to encode a 125 kDa protein (TnNOS) (Wang *et al.*, 1997). The additional exons used by the human nNOSγ transcript (Tex1, Tex1b and Tex2) were all found located in intron 3. Interestingly, exon 5 of the predicted *F*NOS amino acid sequence does contain 3 methionine amino acid residues from which translation of a

truncated *F*NOS could be initiated. However, whilst the *Fugu* NOS1 intron 3 is large, over 3 kb, comparison to the sequence of the alternative exons identified for human nNOSγ, by GAP alignment, did not identify any regions of significant similarity. There was 69% nucleotide identity over 36 bp when compared to Tex1, 65% identity over 52 bp for Tex1b and 83% identity over only 18 bp for Tex2. Therefore, although *f*NOS has the potential for an alternative exon 1 to be spliced to exon 4 and translated from an ATG in exon 5, it would appear that a *Fugu* nNOSγ transcript does not exist.

An alternative nNOS transcript that results from a cassette exon insertion is the nNOS transcript named nNOSμ. nNOSμ has been identified in humans, mice and rats (Silvagno *et al.*, 1996; Magee *et al.*, 1996; Larsson and Phillips, 1998) and has an insertion of 102 bp between exons 16 and 17, that translates into an additional 34 amino acids. In addition, in humans two further alternative insertions of 42 and 67 bp were also identified, although both of these result in the introduction of a stop codon and thus a truncated nNOS protein. Analysis of the *Fugu* NOS1 gene sequence for intron 16 reveals that it is highly unlikely for *Fugu* to have this alternative nNOSμ transcript because the total length of intron 16 is only 75 bp. The sequence of the 102 bp, 34 amino acid insert is well conserved between human, mouse and rat. However, comparison of the nucleotide and possible amino acid sequences of *Fugu* NOS1 intron 16 did not reveal any regions with similarity to the 102 bp, 34 amino acid insert or to the smaller 42 and 67 bp human inserts.

Together, the potential lack of splice variants and reduced number of exon 1 variants may reflect a less complex regulation of the *Fugu* NOS1 gene than that seen for the mammalian NOS1 genes.

6.2. Characterisation of *f*NOS.

The aim of the experiments detailed in chapter 4 was to isolate full-length *Fugu* NOS1 cDNA (*f*NOS). For this, material from the Japanese pufferfish *Takifugu poecilonotus* was used due to a lack of material available from *Fugu rubripes*. Problems of this nature have been experienced by other groups who have used the freshwater pufferfish *T. fluviatilis* as an alternative because it is easier to keep in a laboratory environment and therefore obtain material from (Crongorac-Jurcevic *et al.*, 1997). The use of *T. fluviatilis* has been proposed on a complementary basis to *Fugu rubripes* and several groups have now used *T. fluviatilis* for their studies. From data obtained from mitochondrial DNA, *T. fluviatilis* is estimated to have diverged from *Fugu rubripes* approximately 18-30 million years ago (Crongorac-Jurcevic *et al.*, 1997). The Japanese pufferfish *T. poecilonotus* and *Fugu rubripes* share the same habitat and are even more closely related. For this reason, and the availability of material, *T. poecilonotus* was used for this part of the research project.

Total RNA was isolated from *T. poecilonotus* brain and double-stranded cDNA synthesised. *f*NOS was successfully amplified and cloned into the baculovirus-transfer vector pVL1393. Sequencing of *f*NOS confirmed the structure of the *Fugu* NOS1 gene, as deduced in chapter 3 by comparison to the human NOS1 gene i.e. length of exons and positions of splice-junctions. Amplification and sequencing of *f*NOS also identified the non-coding exon 1, which as discussed in section 6.1 does not have any identity to the sequences of known variants of mammalian NOS1 exon 1.

The predicted amino acid sequence of *F*NOS, translated from the *f*NOS sequence, was used to construct a phylogenetic tree with all the known full-length NOS amino acid

sequences. This phylogenetic analysis was performed using the PIE package available through the bioinformatics web site provided by the HGMP Resource Centre. The PIE package aims to simplify the generation of phylogenetic trees from multiple sequence alignments. In this case CLUSTAL W (Thompson *et al.*, 1994) was used to perform the multiple sequence alignment of all the known full-length NOS sequences including the invertebrate NOSs isolated from *Drosophila melangoster*, *Anopheles stephensis*, *Rhodnius prolixus* and *Lymnaea stagnalis* (appendix, figure A4).

There are several different methods that can be used to create phylogenies, or evolutionary trees, such as the PROTML and PROTDIST programmes. PROTDIST uses a distance matrix programme to generate a phylogenetic tree from distance measurements computed by maximum likelihood estimates for the protein sequences. In comparison, the PROTML maximum likelihood method produces a phylogenetic tree which is the most obvious explanation for the observed sequences, taking into account that some mutations are more likely to occur than others. All methods have their advantages and disadvantages according to the size of alignment file used or how difficult the data set. The phylogenetic programme used to create the unrooted phylogenetic tree for the NOS multiple alignment was the PROTPARS protein sequence parsimony method. PROTPARS is a combination of the methods of Eck and Dayhoff (1961) and Fitch (1971) that produces the most parsimonious tree, that is one where the fewest number of mutations during the evolution of the sequences has occurred. It insists that any amino acid changes that occur are consistent with the genetic code. The assumptions of this method have been described by Felsenstein (http://www.hgmp.mrc.ac.uk/Menu/Help/phylip-protpars.html) as follows:

"1. Change in different sites is independent.

2. Change in different lineages is independent.

3. The probability of a base substitution that changes an amino acid sequence is small over the lengths of time involved in a branch of the phylogeny.

4. The expected amounts of change in different branches of the phylogeny do not vary by so much that two changes in a high-rate branch are more probable than one change in a low-rate branch.

5. The expected amounts of change do not vary enough among sites that two changes in one site are more probable than one change in another.

6. the probability of a base change that is synonymous is much higher than the probability of a change that is not synonymous."

It is also possible to test the robustness of phylogenetic trees created using PROTPARS by bootstrapping. Bootstrapping works by resampling the multiple sequence alignment file (i.e. in this case the CLUSTAL W multiple sequence alignment file) to test whether a feature of the tree is likely to be real. The new data set is the same size as the original with randomly sampled characters replaced. The multiple trees produced by bootstrapping are then resolved to produce a single unrooted phylogenetic tree as seen in figure 4.6, chapter 4. In this case, 1000 bootstrap replications were specified and the *f*NOS predicted amino acid sequence was grouped with the nNOS isoforms. For example, out of 1000 trees, *F*NOS, all of the iNOS and eNOS isoforms and the invertebrate NOSs grouped to the right of murine, rat, rabbit, human and *Xenopus* nNOS 953.5 times. The iNOS, eNOS and invertebrate NOS isoforms then all grouped to the right of *f*NOS 1000 times (appendix, figure A5).

As previously described in section 6.1, one of the primary aims of this research project was to isolate and characterise NOS genes from the model vertebrate genome of the Japanese pufferfish *Fugu rubripes*. The initial library screening was performed with a full-length human iNOS cDNA probe and identified many positive cosmid clones. This

number was significantly reduced by screening the library for a second time with a 5′ iNOS cDNA probe. After the secondary screening, 12 cosmid clones were selected for further study and sequence scanning performed at the HGMP resource centre. However, out of the twelve, only two of the clones were positively identified as containing NOS-like sequences, these were cosmids 064O07 and 066F06. Both cosmid 064O07 and 066F06 showed the highest degree of similarity to the nNOS isoforms. As already discussed, cosmid 064O07 was chosen for further study and the *Fugu* NOS1 gene exons 1 to 19 characterised.

The isolation of a *Fugu* NOS1 using an iNOS probe is not totally surprising when one considers the high degree of sequence identity seen between the three NOS isoforms in mammals. In addition, as from the outset it was not known how similar the *Fugu* and mammalian NOS sequences would be, the stringency of the hybridisation was reduced and therefore may not have been high enough to discriminate between the three isoforms. Although in this case a *Fugu* NOS1 was isolated it does not mean that the other positive cosmid clones identified did not contain *Fugu* NOS2 or NOS3s. Furthermore, despite the fact that sequence scanning only identified two cosmid clones out of the twelve selected cosmids as containing NOS-like sequences does not necessarily mean that the other 10 didn't contain NOS-like sequences. The method of sequence scanning performed at the HGMP Resource Centre involves sonication of cosmid DNA and subcloning into the plasmid vector pBsKS(+) for sequencing from the M13 forward and reverse primer sites. This may not have picked any other NOS genes up because the average *Fugu* cosmid has an insert size of 40 kb and approximately 8 genes on it and the sequence scanning was not totally exhaustive. The other possibility is that *Fugu* does not have either a NOS2 or NOS3 gene but only has a NOS1 gene.

Searching the *Fugu* genomic cosmid library for NOS-like sequences through the HGMP bioinformatics web pages does identify putative NOS2 and NOS3 sequences. However, on closer inspection the regions aligning to the NOS2 and NOS3 isoforms are very limited in their length. The alignments to NOS2 and NOS3 are poor over regions common to both the constitutive and inducible isoforms and the same regions show higher identity to other non-NOS sequences. This is unlike the alignments produced by 064O07 and 066F06 where the regions of identity span several hundred bases and are of high quality i.e. they demonstrated the highest degree of similarity to NOS sequences. Also, the cosmids identified by this search as containing putative NOS2 and NOS3 sequences were not positively identified by either the primary or secondary screening of the *Fugu* genomic cosmid library with the human iNOS cDNA probes.

Evidence for the existence of other NOS isoforms in teleosts has been shown by the partial nNOS cDNA sequence isolated from the Atlantic salmon *Salmo salar*, another teleost fish (Oyan *et al.*, 1998, unpublished: accession number AJ006). It covers the CaM binding domain and the CaM inhibitory sequence and on GAP alignment shows 80% nucleotide identity, 93% amino acid similarity and 89% amino acid identity to the same region of the *Fugu* NOS1 coding sequence. In addition to the *Cyprinus carpio* full-length iNOS mRNA sequence, partial iNOS sequences have been isolated from the rainbow trout *Onchorhyncus mykiss* and the goldfish *Carassius auratus* (Grabowski *et al.*, 1995; Laing *et al*, 1996). Both were identified in macrophage cell lines that had been immunologically challenged. In addition, endothelium-dependent responses in the ventral aorta of the rainbow trout have also been demonstrated (Miller and Vanhoutte, 1986). This supports the hypothesis that these teleost fish have all three NOS isoforms.

As *Fugu* is a teleost fish, as are all of these fishes described above, this suggests that a NOS2 and NOS3 isoform should be present in the *Fugu* genome. However, none of the fish that these isoforms have been identified in belong to the same order as *Fugu*, i.e. the tetraodontiformes, nor do they have such small genomes as *Fugu rubripes*. Fish from the *Cypriniformes* order, which includes the goldfish and crucian carp, have a haploid DNA content of approximately 2pg per cell, those from the *Salmoniformes*, including both the atlantic salmon and the rainbow trout, have a haploid DNA content of 3-3.3pg per cell (Hinegardner 1968; Hingardner and Rosen, 1972). These values are more comparable to the 3pg per cell haploid DNA content of humans. Hinegardner (1968) showed that the more highly specialised, or evolutionarily advanced, a fish was then the less DNA per cell it would have as compared to the more generalised, or less evolved, fishes. Hinegardner and Rosen (1972) went on to discuss the distribution of DNA within the teleostei. They suggested that during evolution the amount of DNA present could remain constant and be modifed by either recombination or mutation, that it could increase by duplication or that it could decrease. They concluded that all three changes had occurred, alone and in combination with each other. They went on to suggest that the loss of DNA, driven by evolution, causes or is caused by, more and more specialisation. Eventually a species may have so little DNA left that can be modified without affecting essential components, that almost any environmental change could lead to its extinction.

This would suggest that, with the smallest vertebrate genomes, the tetraodontiformes order of fishes have undergone significant specialisation and have been left with a minimal amount of DNA. Therefore, if there are differences to be seen in terms of the number of NOS isoforms present in fish then it could be argued that it would be most

likely for it to occur in this order of fish. If indeed *Fugu* was to have only a single NOS gene, with highest identity to the other vertebrate NOS1 isoforms, then it would be required to fulfil the roles played by the NOS2 and NOS3 isoforms.

To date all of the NOS isoforms described for invertebrates have also shown the highest level of homology to the vertebrate nNOS isoforms (Regulski and Tully, 1995; Yuda *et al.*, 1996; Korneev *et al.*, 1998; Luckhart *et al.*, 1998; Luckhart *et al.*, 1999).

The first invertebrate NOS sequence to be reported was the Drosophila (*d*NOS) gene which encodes *D*NOS, a 152 kDa protein of 1350 amino acids (Regulski and Tully, 1995). *D*NOS shows 43% amino acid similarity to rat nNOS but does not have a PDZ domain characteristic of the vertebrate nNOSs. In place of the PDZ domain, *D*NOS has 214 N-terminal amino acids that contain a stretch of 24 glutamine residues. These regions have been implicated in protein-protein interactions, and the regulation of the activation of transcription, in *Drosophila*. Therefore, this domain in *D*NOS may act to localise and/or regulate *D*NOS activity by forming protein-protein interactions, possibly in the same way as the PDZ domain does for vertebrate nNOS isoforms. In common with the vertebrate constitutive NOS isoforms, *D*NOS was also shown to be dependent on Ca^{2+}/CaM for its activity. Regulski and Tully propose that, with the presence of a NOS gene in *Drosophila*, that this gene must have existed for at least 600 million years in an ancestor common to vertebrates and arthropods. Furthermore, in their study it was reported that multiple NOS transcripts were produced in *Drosophila* and that there was preliminary data to suggest that NOS2 and NOS3 isoforms were present in the *Drosophila* genome. However, no further evidence of any other NOS isoforms in the *Drosophila* genome has since been published.

In 1996 the cloning of a NOS cDNA from the salivary glands of the blood-sucking insect *Rhodnius prolixus* was described (Yuda *et al.*, 1996). This sequence encodes a protein of 1174 amino acids and has putative binding sites for haem, tetrahydrobiopterin, calmodulin, FMN, FAD and NADPH. The sequence of the *R. prolixus* NOS had 48%, 46% and 44% identity to human eNOS, nNOS and iNOS respectively. It didn't contain the internal deletion seen for iNOS isoforms between the CaM and FMN-binding sites nor did it contain either a PDZ domain or a myristolation motif as is characteristic for the nNOS and eNOS isoforms. When expressed in a baculovirus insect cell expression system, the recombinant *R. prolixus* NOS was dependent on Ca^{2+}, CaM, NADPH, FAD, FMN and BH_4. For these reasons, the group described the *R. prolixus* NOS, which they named salivary-gland NOS, as a structurally and functionally unique NOS isoform. Multiple NOS transcripts were also shown to be produced in *R. prolixus*.

In 1998 the first molecular evidence for the presence of a NOS gene in the mollusc *Lymnae stagnalis* was presented (Korneev *et al.*, 1998). The *L. stagnalis* NOS sequence (Lym-nNOS) was identified as a NOS1 isoform and showed the occurrence of alternative splicing. In addition, using a 5′ probe, to a well-conserved region of all NOS isoforms, this group identified two major NOS transcripts in all the tissues examined including the CNS. In comparison, a probe made from the 3′ untranslated region (UTR) only identified a single transcript in the CNS. The explanations offered for this were, (i) the presence of multiple messages was due to alternative splicing, a well known phenomenon for mammalian NOS1s, or, (ii) that there are different messages transcribed from different genes and that as such there is more than one NOS isoform present in *L. stagnalis*. As for *Drosophila*, Korneev's group report preliminary data suggesting that there is more than one NOS isoform present. Although Lym-nNOS has been identified

as a NOS1 isoform, in part due to its higher amino acid similarity overall and in particular in the CaM binding domain, it does not have a PDZ domain. The Lym-nNOS amino acid sequence is truncated at its 5′ end and does not start until after the Zn binding motif that has been identified in all other NOSs including the *Drosophila* sequence. No further evidence of the existence of any other NOS isoforms in the *L. stagnalis* genome has since been published. Database and literature searches failed to identify any other *L. stagnalis* NOS isoforms.

In 1998 Luckhart and colleagues described the inducible synthesis of nitric oxide by the mosquito *Anopheles stephensi* as a means of limiting the development of the malaria parasite *Plasmodium.* They isolated ~30 kb of genomic DNA which encompassed the mosquito NOS gene AsNOS. AsNOS encodes an open reading frame of 1247 amino acids that contain all of the characteristic NOS binding domains. AsNOS, like dNOS and Lym-nNOS, shows the highest degree of similarity to the vertebrate NOS1 sequences. Further work on the gene structure of AsNOS identified 19 exons and repetitive elements within some of the larger introns (Luckhart *et al.*, 1999). In addition, multiple transcription start sites were identified within ~250 bp of the initiation methionine and transcription factor binding sites were identified for lipopolysaccharide and inflammatory cytokine-responsive elements very similar to the murine NOS2 promoter. The presence of any additional NOS genes within the *A. stephensi* genome is unknown to date.

Luckhart *et al.* (1999) took the data from the *A. stephensi* NOS gene and the NOS sequences isolated from *Drosophila*, *R. prolixus* and *L. stagnalis* and speculated that a single insect NOS gene type is associated with diverse physiologies. Therefore, together the invertebrate NOS data supports the idea of a single NOS gene being able to act as all

three and that an ancestral NOS1 gene evolved first, followed by the gene duplication and diversification of the NOS2 and NOS3 genes present in mammals.

6.3. Expression, activity and purification of *F*NOS.

To enable functional studies of *F*NOS to be performed, recombinant *F*NOS was generated in a baculovirus-insect cell expression system. This system is well established for the expression of mammalian NOS enzymes and was chosen because of its ability to express high levels of protein and perform post-translational modifications. As previously described, the 5′ and 3′ ends of *f*NOS were engineered to incorporate *Not*I restriction endonuclease sites to permit cloning into the *Not*I site of the baculovirus transfer vector pVL1393. Restriction endonuclease digestion and sequencing confirmed both the correct orientation and reading-frame of *f*NOS in the vector. The creation of *f*NOS-containing recombinant baculovirus generated NO_2^- in culture supernatants of *Sf*21 insect cells infected with these viruses and thus provided the first evidence that the *F*NOS protein translated from *f*NOS was functional.

No obvious *F*NOS expression band was noticeable upon staining of SDS-PAGE gels of infected *Sf*21 cells as compared to uninfected cells over a 96 hour time-course. A number of immunoreactive protein species were detected by western blot analysis of the time-course with a monoclonal anti-nNOS antibody. Of these proteins the two largest represent the most reasonable *F*NOS candidates, the presence of the immunoreactive low molecular weight proteins may represent proteolytic breakdown of *F*NOS.

Further western blot analysis revealed that recombinant *F*NOS was only detected by anti-nNOS antibodies and that the monoclonal antibody was more sensitive than the

polyclonal antibody. This difference in sensitivity could reflect post-translational modifications carried out by the insect cells such that in recombinant *F*NOS epitopes may not be available for recognition by the respective antibodies.

In addition, western blot analysis and the spectrophotometric measurement of NOS activity demonstrated the maximal expression and activity of *F*NOS at 24 hours. After 24 hours, activity and immunoreactive protein levels began to decline. Immunocytochemical detection of *F*NOS expression in pVL*f*NOS infected-*Sf*21 cells over a 96 hour time-course confirmed these findings.

Activity of recombinant *F*NOS was assayed by measuring the conversion of oxyhaemoglobin to methaemoglobin by NO on a dual-wavelength spectrophotometer. In order to assess the L-arginine requirements of recombinant *F*NOS, L-arginine-free TC100 medium was used for the culturing of *Sf*21 cells followed by the use of Dowex anion-exchange resin to remove any residual L-arginine from the cell cytosol. NADPH, FAD, BH_4 and oxyhaemoglobin were included in the reaction mix and activity initiated by the addition of L-arginine. For this reason it is not possible to assess the exact requirements of recombinant *F*NOS for any of these co-factors. In a crude cell lysate, maximum enzyme activity of ~3.75 pmol NO/min/mg of protein was seen on the addition of 10 μM L-arginine with a half-maximal rate at ~0.79 μM L-arginine.

The addition of exogenous calmodulin had no effect on the kinetics of recombinant *F*NOS activity. This was surprising because this behaviour is more indicative of the iNOS isoforms. However, to assess the Ca^{2+} dependence of recombinant *F*NOS the calcium chelator EGTA was added to the reaction mix. At 1 mM EGTA, *F*NOS activity was completely abolished. This phenomenom is characteristic of all constitutive NOS isoforms. Charles *et al.* (1996) expressed all three human NOS isoforms in the

baculovirus-insect cell expression system and found that the addition of EGTA had no effect on recombinant human iNOS activity but reduced the activity of recombinant human nNOS and eNOS to approximately 10% and 5% of their maximums respectively. The Ca^{2+}-dependent/CaM-independent, constitutive/inducible behaviour of recombinant *F*NOS is contradictory. One possible explanation is that recombinant *F*NOS has bound endogenous calmodulin produced by the *Sf*21 insect cells, and that this has remained bound throughout the preparation of the cell cytosol for the assay of *F*NOS activity. If this is case, it suggests that recombinant *F*NOS is Ca^{2+}/CaM-dependent and similar in activity to the mammalian constitutive NOS enzymes.

Recombinant *F*NOS was inhibited by all three of the NOS inhibitors used in these studies. The most potent inhibitor of recombinant *F*NOS activity was the non-specific NOS inhibitor L-NMMA, IC_{50} = 0.63 μM, which has a similar potency for all three NOS isoforms (Southan and Szabó, 1996; Griffith and Kilbourn, 1996; Vallance, 1996). 1400W is a highly selective mechanism based inhibitor of iNOS isoforms whose IC_{50} for recombinant *F*NOS was ~1.5 μM. 1400W binds to all NOS isoforms but only irreversibly inhibits certain isoforms, the selectivity ratio of 1400W for iNOS is 5000 as compared to 25 for nNOS and 1 for eNOS isoforms (Garvey *et al.*, 1997). L-Thiocitrulline has a K_i for nNOS of 0.06 μM and 3.6 μM for iNOS (Frey *et al.*, 1994), the IC_{50} of L-Thiocitrulline for recombinant *F*NOS was ~15 μM. The exact action of each NOS inhibitor is dependent on the system in which is being used. From the IC_{50} values obtained it can be concluded that *F*NOS was most strongly inhibited by a general NOS inhibitor followed by an iNOS selective inhibitor and then an nNOS inhibitor. However to be able to interpret what this means the same inhibitors would have to be

Table 6.1. IC_{50} values for the NOS inhibitors L-NMMA, L-Thiocitrulline and 1400W.

Isoform	L-NMMA	L-thiocitrulline	1400W
nNOS	0.65 μm	0.26 μm*	8.6 μm*
iNOS	3.9 μm	15.6 μm*	0.03 μm*
eNOS	0.70 μm	-	216 μm*
*F*NOS	0.08 μm•	15 μm•	0.2 μm•

IC_{50} and K_i values from Rees *et al.*, 1998; Frey *et al.*, 1994 & Garvey *et al.*, 1997. *Values calculated with the assumption that the inhibitor follows Michaelis Menten kinetics. • IC_{50} values for *F*NOS are estimates taken from the inhibitor analysis curves seen in Figure 5.4.

facing P180

used on a mammalian recombinant nNOS, iNOS and eNOS expressed in exactly the same system. In addition, the results would be more meaningful if they could be compared to the IC_{50}s of these inhibitors on a recombinant *Fugu* iNOS or eNOS.

Initially, purification of recombinant *F*NOS was attempted using affinity chromatography on 2′,5′ ADP sepharose, a technique that relies on the similarity of the molecular structure of ADP and NADPH. Using this method, proteins in cell cytosols that showed an affinity for NADPH were adsorbed onto 2′,5′ ADP sepharose columns and retained through an interaction between the NADPH-binding domain of the protein and the ADP moiety attached to the sepharose. Proteins that bound through a non-specific mechanism, for example by hydrophobic interactions, were removed from the column by increasing the ionic strength of the wash buffer. This was achieved by raising concentrations of NaCl to 0.5 M. Elution of the column with relatively high concentrations of NADPH (10 mM) could then be used to remove specifically bound proteins through competition for the NADPH binding-site on the protein. The fact that proteins eluted from the column using this technique demonstrated that, (i) proteins were bound to the column in the first instance, (ii) some of the proteins were refractory to elution by buffers of high ionic strength and (iii) these particular proteins were adsorbed via their affinity for NADPH and that this affinity was higher than that displayed for the structural analogue, 2′5′ ADP. SDS-PAGE and western analysis with the monoclonal nNOS antibody confirmed that one of the partially purified proteins was NOS immunoreactive.

Further purification of recombinant *F*NOS was attempted using affinity chromatography on CaM sepharose. This technique relies on the predicted affinity of recombinant *F*NOS being most like an nNOS isoform for calmodulin. NOS immunoreactive eluates from the

2′,5′ ADP sepharose column were adsorbed onto the CaM sepharose column and retained through an interaction between the CaM-binding domain of the protein and the CaM moiety attached to the sepharose. Non-specifically bound proteins were again removed from the column by the addition of wash buffer containing NaCl. Elution of the column was then attempted by the addition of increasing concentrations of the calcium chelator EGTA. When this failed to elute any specifically-bound protein, elution of the column was attempted with 1 M NaCl followed by 6 M urea, yet neither succeeded. From this it can be concluded that a very strong bond formed between the recombinant *F*NOS and the CaM on the CaM sepharose column. This was confirmed by SDS-PAGE and western analysis on a sample of the CaM sepharose column itself. From the activity studies it was suggested that recombinant *F*NOS was bound to endogenous CaM from the *Sf*21 insect cells. If this is the case then *F*NOS must have a higher affinity for the mammalian CaM of the CaM sepharose column than the *Sf*21 CaM and an irreversible exchange taken place on the column.

6.4 Conclusions.

At the outset, the primary aims of this work were to isolate, sequence and analyse NOS genes from the genome of the Japanese pufferfish *Fugu rubripes*, to isolate *Fugu* NOS cDNA and to purify and characterise any recombinant *Fugu* NOS protein.

Initial experiments, detailed in chapter 3, identified a *Fugu* NOS1 gene. A significant part of this thesis was therefore devoted to the molecular cloning, sequencing and genomic characterisation of the *Fugu* NOS1 gene, which in its entirety spanned 22,203 bp. In comparison to the ~160 kb encompassed by the human NOS1 gene this

represents a reduction in size of ~7.2 times which is in line with the overall reduction in size of the *Fugu* genome. There was conservation of exon length and intron-exon boundaries between the two genes, all predicted NOS co-factor binding domains were conserved and a NOS1 specific PDZ domain identified.

To obtain information on the structure, regulation and biochemistry of the enzmye encoded by the *Fugu* NOS1 gene a recombinant version, *F*NOS, had to be produced. First, a full-length *Fugu* NOS1 cDNA (*f*NOS) was cloned from the brain mRNA of the Japanese pufferfish *T.poecilonotus*. The full-length cDNA was itself characterised by sequencing and phylogenetic analysis. Sequence information from *f*NOS identifed an open reading frame of 4257 bp, which translated into 1419 amino acids, and confirmed the genomic organisation deduced for the *Fugu* NOS1 gene. Phylogenetic analysis grouped *f*NOS with the vertebrate nNOS isoforms as expected. *f*NOS was then cloned into the baculovirus transfer vector pVL1393 and transfected into *Sf*21 insect cells for expression.

The baculovirus-directed expression of *F*NOS permitted a preliminary enzymological study of the recombinant enzyme. *F*NOS was first shown to be functional by nitrite production as measured by the chemiluminescence assay. The NOS activity of *F*NOS was confirmed by the measurement of oxyhaemoglobin to methaemoglobin conversion by NO and its inhibition by NOS inhibitors. In addition, *F*NOS was partially purified by 2′,5′ ADP sepharose affinity chromatography.

6.5 Future research.

A major piece of future research would have to be attempting to isolate NOS2 and NOS3 genes from the *Fugu* genome. In addition to the brain material supplied from the Japanese pufferfish *T. poecilonotus*, hearts were also provided. This would be a good starting point for the possible isolation of a eNOS isoform. Use of the TRIZOL Reagent, which was used for the isolation of total RNA from the brain (as described in section 2.12), would be ideal because it not only allows the isolation of total RNA but also of DNA and protein. This would allow a comprehensive study into the possible existence of an eNOS isoform. PCR primers could be designed from the *Fugu* NOS1 gene sequence to amplify a region, such as the CaM inhibitory domain, that can be defined as nNOS, iNOS or eNOS by sequencing. For the isolation of a *Fugu* iNOS mRNA the best option would be the isolation of macrophages and RT-PCR following induction with a bacterial pathogen. The detection of an iNOS mRNA was described previously by Grabowski *et al* (1996) following the *in vivo* challenge of rainbow trout with a genetically attenuated ($AroA^-$) fish bacterial pathogen (*Aeromonas salmonicida*). Two days post-challenge, gills were isolated and snap-frozen and head kidneys dissected and enriched for macrophages. RNA was isolated from both preparations and degenerate primers used to successfully amplify iNOS mRNA by RT-PCR. Additionally, RT-PCR was used successfully to detect an iNOS mRNA in a goldfish macrophage cell line both before and after challenge with *E. coli* LPS (Laing *et al.*, 1997). In addition to the heart and brain supplied from the Japanese pufferfish *T. poecilonotus*, gills were also supplied. Unlike the rainbow trout used for iNOS mRNA isolation, the *T. poecilonotus* from which the gills were isolated had not been maintained in a laboratory and had not

been intentionally exposed to an immunological challenge. It is therefore not clear how likely it would be that these gills would be expressing iNOS mRNA. As an alternative, the use of *Fugu* cell lines would be possible. *Fugu* cell lines were set up to provide easier access to biological material from the pufferfish (Bradford *et al.*, 1997). Cell cultures were established from *T. niphobles* (kusafugu) fry and *F. rubripes* (torafugu) eye, brain, liver, spleen, fin, kidney, swimbladder and muscle. Unfortunately the corresponding author of this paper did not respond to requests made for the availability of these cell lines.

Northern blot analysis of RNA isolated from the brain, heart, gill and any other available pufferfish material, including the *Fugu* cell lines, would allow a more meaningful investigation into the structural and allelic diversity of the *f*NOS isolated through these studies. Using a probe specific to the nNOS isoform could help to answer the question of whether there are multiple transcripts and alternative splicing of *f*NOS as seen for other nNOS isoforms. Equally using a probe general to all three NOS isoforms may identify different sized transcripts that could indicate the presence of more than one isoform.

Investigation into the promoter sequence of the *Fugu* NOS1 gene would also be of interest. As described previously, *Fugu* has been used to identify conserved enhancer elements for both the *Hoxb-1* and *Hoxb-4* genes (Marshall *et al.*, 1994; Aparicio *et al.*, 1995).

In addition, more work on the *Fugu ksr* gene would be required to confirm its position in the *Fugu* genome, upstream of the *Fugu* NOS1 gene. This may be useful in establishing the evolution of this region

Future research could also include a more detailed biochemical characterisation of the recombinant *Fugu* NOS1 protein (*F*NOS). This would include a study of the exact co-factor requirements of recombinant *F*NOS necessary for enzyme activity. In addition, the use of a calmodulin antagonist would be assessed for its effect on *F*NOS activity to try and deduce the Ca^{2+}/CaM requirements. The effect of more NOS selective inhibitors, such as L-NIL and L-NIO, on enzyme activity would be used to produce a more detailed inhibition profile for *F*NOS.

Further purification of *F*NOS would be attempted by methods other than CaM sepharose affinity chromoatography including size selection or anion-exchange methods, which have been used successfully for mammalian NOS isoforms.

REFERENCES

Abu-Soud, S. H. and Stuehr, D. J. (1993). Nitric oxide synthases reveal a role for calmodulin in controlling electron transfer. *Proc. Natl. Acad. Sci. USA.* **90**. 10769-10772.

Adachi, H., Iida, S., Oguchi, S., Ohshima, H., Suzuki, H., Nagasaki, K., Kawasaki, H., Sugimura, T. and Nathan, C. (1993). Molecular cloning of a cDNA encoding an inducible calmodulin-dependent nitric-oxide synthase from rat liver and its expression in COS 1 cells. *Eur J. Biochem.* **217**. 37-43.

Adler, H., Frech, B., Thony, M., Pfister, H., Peterhans, E. and Jungi, T. W. (1995). Inducible nitric oxide synthase in cattle. Differential cytokine regulation of nitric oxide synthase in bovine and murine macrophages *J. Immunol.* **154**. 4710-4718.

Afework, M., Tomlinson, A. and Burnstock, G. (1994). Distribution and colocalization of nitric oxide synthase and NADPH-diaphorase in adrenal gland of developing, adult and aging Sprague-Dawley rats. *Cell Tissue Res.* **276**. 133-141.

Ali, J. B. M., Clark, M., Furlong, R. and Green, A. J. (1997). Comparative mapping of the TSC1 region of chromosome 9q34 in *Fugu rubripes* using the human ATS-V gene. *Ann. Hum. Genet.* **61**. 207.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**. 403-410.

Alwen, J. (1990). United Kingdom genome mapping project: background, development, components, coordination and management, and international links of the project. *Genomics.* **7**. 386-388.

Andrews, H. E., Bruckdorfer, K. R., Dunn, R. C. and Jacobs, M. (1987). Low-density lipoproteins inhibit endothelium-dependent relaxation in rabbit aorta. *Nature.* **327**. 237-239.

Aparicio, S., Morrison, A., Gould, A., Gilthorpe, J., Chaudhuri, C., Rigby, P., Krumlauf, R. and Brenner, S. (1995). Detecting conserved regulatory elements with the model genome of the Japanese puffer fish, *Fugu Rubripes*. *Proc. Nat. Acad. Sci. USA*. **92**. 1684-1688.

Armes, N., Gilley, J. and Fried, M. (1997). The comparative genomic structure and sequence of the surfeit gene homologs in the puffer fish *Fugu rubripes* and their association with CpG-rich islands. *Genome Res.* **7**. 1138-1152.

Arnold, W. P., Mittal, C. K., Katsuki, S. and Murad, F. (1977). Nitric oxide activates guanylate cyclase and increases guanosine 3′:5′ -cyclic monophosphate levels in various tissue preparations. *Proc. Natl. Acad. Sci. USA*. **74**. 3203-3207.

Baek, K. J., Thiel, B. A., Lucas, S. and Stuehr, D. J. (1993). Macrophage nitric oxide synthase subunits. Purification, characterization and role of prosthetic groups and substrate in regulating their association into a dimeric enzyme. *J. Biol. Chem.* **268**. 21120-21129.

Baxendale, S., Abdulla, S., Elgar, G., Buck, D., Berks, M., Micklem, G., Durbin, R., Bates, G., Brenner, S., Beck, S., Lehrach, H. (1995). Comparative sequence analysis of the human and pufferfish Huntington's disease genes. *Nat. Genet.* **10**. 67-76.

Beasley, D. and Eldridge, M. (1994). Interleukin-1 β and tumor necrosis factor-α synergistically induce NO synthase in rat vascular smooth muscle cells. *Am. J. Physiol.* **266**. R1197-R1203.

Beckman, J. S., Chen, J., Crow, J. P. and Ye, Y. Z. (1994). Reactions of nitric oxide, superoxide and peroxynitrite with superoxide dismutase in neurodegeneration. *Prog. Brain. Res.* **103**. 371-380.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**. 1513-1522.

Black, S. M. and Ortiz de Montellano, P. R. (1995). Characterization of rat neuronal nitric oxide synthase expressed in *Saccharomyces cerevisiae*. *DNA Cell Biol.* **14**. 789-794.

Boeddrich, A., Burgtorf, C., Crollius, H. R., Hennig, S., Bernot, A., Clark, M., Reinhardt, R., Lehrach, H. and Francis, F. (1999). Analysis of spermine synthase gene region in *Fugu rubripes*, *Tetraodon fluviatilis*, and *Danio rerio*. *Genomics*. **57**. 164-164.

Bohme, G. A., Bon, C., Stutzmann, J. M., Doble, A. and Blanchard, J. C. (1991). Possible involvement of nitric oxide in long-term potentiation. *Eur. J. Pharmacol.* **199**. 379-381.

Bosca, L. and Lazo, P. A. (1994). Induction of nitric oxide release by MRC OX-44 (anti-CD53) through protein kinase C-dependent pathway in rat macrophages. *J. Exp. Med.* **179**. 1119-1126.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the prunciple of protein-dye binding. *Anal. Biochem.* **72**. 248-254.

Bradford, C. S., Miller, A. E., Toumadje, A, Nishiyama, K., Shirahata, S. and Barnes, D. W. (1997). Characterization of cell cultures derived from *Fugu*, the Japanese pufferfish. *Mol. Mar. Biol. Biotechnol.* **6**. 279-288.

Bredt, D. S. and Snyder, S. H. (1990). Isolation of nitric oxide synthetase, a calmodulin-requiring enzyme. *Proc. Natl. Acad. Sci. USA*. **87**. 682-685.

Bredt, D. S., Hwang, P. M. and Snyder, S. H. (1990). Localization of nitric oxide synthase indicating a neuronal role for nitric oxide. *Nature*. **347**. 768-770.

Bredt, D. S., Glatt, C. E., Hwang, P. M., Fotuhi, M., Dawson, T. M. and Snyder, S. H. (1991). Nitric oxide synthase protein and mRNA are discretely localized in neuronal populations of the mammalian CNS together with NADPH diaphorase. *Neuron.* **7**. 615-624.

Bredt, D. S., Hwang, P. M., Glatt, C. E., Lowenstein, C., Reed, R. R. and Snyder, S. H. (1991). Cloned and expressed nitric oxide synthase structurally resembles cytochrome P-450 reductase. *Nature.* **351**. 714-718.

Bredt, D. S. and Snyder, S. H. (1994). Nitric Oxide - A physiologic messenger molecule. *Annu. Rev. Biochem.* **63**. 175-195.

Brenman, J. E., Chao, D. S., Xia, H., Aldape, K. and Bredt, D. S. (1995). Nitric oxide synthase complexed with dystrophin and absent from skeletal muscle sarcolemma in Duchenne muscular dystrophy. *Cell.* **82**. 743-752.

Brenman, J. E., Chao., D. S., Gee, S. H., McGee, A. W., Craven, S. E., Santillano, D. R., Wu, Z., Huang, F., Xia, H., Peters, M. F., Froehner, S. C. and Bredt, D. S. (1996). Interaction of nitric oxide synthase with the postsynaptic density protein PSD-95 and α1-syntrophin mediated by PDZ domains. *Cell.* **84**. 757-767.

Brenman, J. E., Xia, H., Chao, D. S., Black, S. M. and Bredt, D. S. (1997). Regulation of neuronal nitric oxide synthase through alternative transcripts. *Dev. Neurosci.* **19**. 224-231.

Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B. and Aparicio, S. (1993). Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature.* **366**. 265-268.

Bryant, P. J. and Woods, D. F. (1992). A major palmitoylated membrane protein of human erythrocytes shows homology to yeast guanylate kinase and to the product of a *Drosophila* tumor suppressor gene. *Cell.* **68**. 621-622.

Bult, H., Boeckxstaens, G. E., Pelckmans, P. A., Jordaens, F. H., Van Mearcke, Y. M. and Herman, A. B. (1990). Nitric oxide as an inhibitory non-adrenergic non-cholinergic neurotransmitter. *Nature*. **345**. 346-347.

Burnstock, G. (1981). Neurotransmitters and trophic factors in the autonomic nervous-system. *J. Phys.(Lond.)* **313**. 1-35

Busconi, L. and Michel, T. (1993). Endothelial nitric oxide synthase. N-terminal myristoylation determines subcellular localization. *J. Biol. Chem.* **268**. 8410-8413

Busconi, L. and Michel, T. (1995). Recombinant endothelial nitric oxide synthase: Post-translational modifications in a baculovirus expression system. *Mol. Pharmacol.* **47**. 655-659.

Cecconi, F., Crosio, C., Mariottini, P., Cesareni, G., Giorgi, M., Brenner, S. and Amaldi, F. (1996). A functional role for some *Fugu* introns larger than the typical short ones: the example of the gene coding for ribosomal protein S7 and snoRNA U17. *Nucleic Acids Res.* **24**. 3167-3172.

Chang, M. S., Leu, J. H., Yao, C. W., Huang, F. L. and Huang, C. J. (1997). Genomic structure and promoter region of c-fos gene of round-spotted pufferfish, *Tetraodon nigroviridis* (Syn *T-fluviatilis*). *Zool. Studies*. **36**. 159-169.

Charles, I. G., Chubb, A., Gill, R., Clare, J., Lowe, P. N., Holmes, L. S., Page, M., Keeling, J. G., Moncada, S. and Riveros-Moreno, V. (1993). Cloning and expression of a rat neuronal nitric oxide synthase coding sequence in a baculovirus/insect cell system. *Biochem. Biophys. Res. Commun.* **196**. 1581-1489.

Charles, I. G., Palmer, R. M. J., Hickery, M. S., Bayliss, M. T., Chubb, A. P., Hall, V., Moss, D. W. and Moncada, S. (1993). Cloning, characterization, and expression of a cDNA encoding an inducible nitric oxide synthase from the human chondrocyte. *Proc. Natl. Acad. Sci. USA*. **90**. 11419-11423.

Charles, I. G., Scorer, C. A., Moro, M. A., Fernandez, C., Chubb, A., Dawson, J., Foxwell, N., Knowles, R. G. and Baylis, S. A. (1996). Expression of human nitric oxide synthase isozymes. *Meth. Enzymol.* **268**. 449-460.

Chartrain, N. A., Geller, D. A., Koty, P. P., Sitrin, N. F., Nussler, A. K., Hoffman, E. P., Billiar, T. R., Hutchinson, N. I. and Mudgett, J. S. (1994). Molecular cloning, structure, and chromosomal localization of the human inducible nitric oxide synthase gene. *J.Biol. Chem.* **269**. 6765-6772.

Chen, P. F., Tsai, A. L. and Wu, K. K (1994). Cysteine 184 of endothelial nitricoxide synthase is involved in heme coordination and catalytic activity. *J. Biol. Chem.* **269**. 25062-25066.

Chen, P. F., Tsai, A. L. and Wu, K. K. (1995). Cysteine 99 of endothelial nitric oxide synthase (NOS-III) is critical for tetrahydrobiopterin-dependent NOS-III stability and activity. *Biochem. Biophys. Res. Commun.* **215**. 1119-1129.

Chen, P. F., Tsai, A. L., Berka, V. and Wu, K. K. (1996). Endothelial nitric-oxide synthase. Evidence for bidomain structure and successful reconstitution of catalytic activity from two separate domains generated by a baculovirus expression system. *J. Biol. Chem.* **271**. 14631-14635.

Chin, J. H., Azhar, S. and Hoffman, B. B. (1992). Inactivation of endothelial derived relaxing factor by oxidized lipoproteins. *J. Clin. Invest.* **89**. 10-18.

Cho, K-O., Hunt, C. A. and Kennedy, M. B. (1992). The rat brain postsynaptic density fraction contains a homolog of the *Drosophila* discs-large tumor suppressor protein. *Neuron.* **9**. 929-942.

Cho, H. J., Martin, E., Xie, Q-W., Sassa, S. and Nathan, C. (1995). Inducible nitric oxide synthase: Identification of amino acid residues essential for dimerization and binding of tetrahydrobiopterin. *Proc. Natl. Acad. Sci USA*. **92**. 11514-11518.

Colasanti, M., Lauro, G. M. and Venturini, G. (1995). NO in Hydra feeding response. *Nature*. **374**. 505.

Colbert, E. H. and Morales, M. (1991). Evolution of the Vertebrates. (Wiley-Liss, New York).

Crane, B. R., Arvai, A. S., Gachhui, R., Wu, C., Ghosh, D. K., Getzoff, E. D., Stuehr, D. J. and Tainer, J. A. (1997). The structure of nitric oxide synthase oxygenase domain and inhibitor complexes. *Science*. **278**. 425-431.

Crane, B. R., Arvai, A. S., Ghosh, D. K., Wu, C., Getzoff, E. D., Stuehr, D. J. and Tainer, J. A. (1998). Structure of nitric oxide synthase oxygenase dimer with pterin and substrate. *Science*. **279**. 2121-2125.

Creager, M. A., Cooke, J. P., Mendelsohn, M. E., Gallagher, S. J., Coleman, S. M., Loscalzo, J. and Dzau, V. J. (1990). Impaired vasodilation of forearm resistance vessels in hypercholesterolemic humans. *J. Clin. Invest*. **86**. 228-234.

Creager, M. A., Gallagher, S. J., Girerd, X. J., Coleman, S. M., Dzau, V. J. and Cooke, J. P. (1992). L-arginine improves endothelium-dependent vasodilation in hypercholesterolemic humans. *J. Clin. Invest*. **90**. 1248-1253.

Crnogorac-Jurcevic, T., Brown, J. R., Lehrach, H. and Schalkwyk, L. C. (1997). *Tetraodon fluviatilis*, a new puffer fish model for genome studies. *Genomics*. **41**. 177-184.

D'alessio, G., Didonato, A., Jaffe, K., Maldonado, H. and Zabala, N. A. (1982). Arginine and memory consolidation in praying mantis. *J. Comp. Phys.* **147**. 231-235.

Dawson, T. M., Bredt, D. S., Fotuhi, M., Hwang, P. M. and Snyder, S. H. (1991). Nitric oxide synthase and neuronal NADPH diaphorase are identical in brain and peripheral tissues. *Proc. Natl. Acad. Sci. USA*. **88**. 7797-7801.

Dawson, T. M., Dawson, V. L. and Snyder, S. H. (1992). A novel neuronal messenger molecule in brain: the free radical, nitric oxide. *Annal. Neurol.* **32**. 297-311.

Dawson, V. L., Dawson, T. M., Bartley, D. A., Uhl, G. R and Snyder, S. H. (1993). Mechanisms of nitric oxide mediated neurotoxicity in primary brain cultures. *J. Neurosci.* **13**. 2651-2661.

Dawson, T. M. and Dawson, V. L. (1996). Nitric oxide synthase: role as a transmitter/mediator in the brain and endocrine system. *Annu. Rev. Med.* **47**. 219-227.

DeLisi, C. (1988). The Human Genome Project. *Am. Scientist.* **76**. 488-498.

Desai, K. M., Sessa, W. C. and Vane, J. R. (1991). Involvement of nitric oxide in the reflex relaxation of the stomach to accomodate food or fluid. *Nature*. **351**. 477-479.

Downes, H., Edwards, M. W., Elsey, T. S. and Walters, C. L. (1976). Determination of a non-volatile nitrosamine by using denitrosation and a chemiluminescence analyser. *Analyst*. **101**. 742-748.

Dulbecco, R. (1986). A turning point in cancer research: sequencing the human genome. *Science*. **231**. 1055-1056.

Dulbecco, R. (1987). The Italian genome program. *Genomics*. **7**. 294-297.

Dun, N. J., Dun, S. L., Wu, S. Y. and Förstermann, U. (1993). Nitric oxide synthase immunoreactivity in rat superior cervical ganglia and adrenal glands. *Neurosci. Lett.* **158**. 51-54.

Dunham, I., Shimizu, N., Roe, B. A., *et al.* (1999). The DNA sequence of human chromosome 22. *Nature*. **402**. 489-495.

Eck, R. V. and Dayhoff, M. O. (1966). Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Silver Spring, Maryland.

Edwards, Y. J. K., Elgar, G., Clark, M. S. and Bishop, M. J. (1998). The identification and characterization of microsatellites in the compact genome of the Japanese pufferfish, *Fugu rubripes*: Perspectives in functional and comparative genomic analyses. *J. Mol. Biol.* **278**. 843-854.

Elgar, G., Rattray, F., Greystrong, J. and Brenner, S. (1995). Genomic structure and nucleotide sequence of the p55 gene of the puffer fish *Fugu rubripes*. *Genomics*. **27**. 442-446.

Elgar, G., Sandford, R., Aparicio, S., Macrae, A., Venkatesh, B. and Brenner, S. (1996). Small is beautiful: comparative genomics with the pufferfish (*Fugu rubripes*). *Trends Genet.* **12**. 145-150.

Elgar, G., Clark, M. S., Meek, S., Smith, S., Warner, S., Edwards, Y. J. K., Bouchireb, N., Cottage, A., Yeo, G. S. H., Umrania, Y., Williams, G and Brenner, S. (1999). Generation and analysis of 25 Mb of genomic DNA from the pufferfish *Fugu rubripes* by sequence scanning. *Genome Res.* **9**. 960-971.

Eliasson, M. J. L., Blackshaw, S., Schell, M. J. and Snyder, S. H. (1997). Neuronal nitric oxide synthase alternatively spliced forms: prominent functional localizations in the brain. *Proc. Natl. Acad. Sci. USA*. **94**. 3396-3401.

Feelisch, M. and Noack, E. A. (1987). Correlation between nitric oxide formation during degradation of organic nitrates and activation of guanylate cyclase. *Eur. J. Pharmacol.* **139**. 19-30.

Feldman, P. L., Griffith, O. W., Hong, H. and Stuehr, D. J. (1993). Irreversible inactivation of macrophage and brain nitric oxide synthase by L-NG-methylarginine requires NADPH-dependent hydroxylation. *J. Med. Chem.* **36**. 491-496.

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specified tree topology. *Systematic Zoology*. **20**: 406-416.

Förstermann, U., Mulsch, A. and Bohme, E. (1986). Stimulation of soluble guanylate cyclase by an acetycholine-induced endothelium-derived factor from rabbit and canine arteries. *Circ. Res.* **58**. 531-538.

Fujisawa, H., Ogura, T., Kurashima, Y., Yokoyama, T., Yamahita, J. and Esumi, H. (1994) Expression of two types of nitric oxide synthase mRNA in human neuroblastoma cell lines. *J. Neurochem.* **63**. 140-145.

Furchgott, R. F., and Zawadzki, J. (1980). The obligatory role of endothelial cells in the relaxation of arterial smooth muscle by acetylcholine. *Nature*. **288**. 373-376.

Furfine, E. S., Harmon, M. F., Paith, J. E., Knowles, R. G., Salter, M., Kiff, R. J., Duffy, C., Hazelwood, R., Oplinger, J. A. and Garvey, E. P. (1994). Potent and selective inhibition of human nitric oxide synthases. Selective inhibition of neuronal nitric oxide synthase by S-methyl-L-thiocitrulline and S-ethyl-L-thiocitrulline. *J. Biol. Chem.* **269**. 26677-26683.

Gachhui, R., Ghosh, D. K., Wu, C., Parkinson, J., Crane, B. R. and Stuehr, D. J. (1997). Mutagenesis of acidic residues in the oxygenase domain of inducible nitric oxide synthase identifies a glutamate involved in arginine binding. *Biochemistry*. **36**. 5097-5103.

Galea, E., Reis, D. J. and Feinstein, D. L. (1994). Cloning and expression of inducible nitric oxide synthase from rat astrocytes. *J. Neurosci. Res.* **37**. 406-414.

Garcia-Cardena, G., Fan, R., Stern, D. F., Liu, J. W. and Sessa, W. C. (1996). Endothelial nitric oxide synthase is regulated by tyrosine phosphorylation and interacts with caveolin-1. *J. Biol. Chem.* **271**. 27237-27240.

Garcia-Cardena, G., Martasek, P., Masters, B. S. S., Skidd, P. M., Couet, J., Li, S. W., Lisanti, M. P. and Sessa, W. C. (1997). Dissecting the interaction between nitric oxide synthase (NOS) and caveolin. Functional significance of the NOS caveolin binding domain in vivo. *J. Biol. Chem.* **272**. 25437-25440.

Garthwaite, J., Charles, S. L. and Chess-Williams, R. (1988). Endothelium-derived relaxing factor release on activation of NMDA receptor activation suggests role as intracellular messenger in the brain. *Nature*. **336**. 385-388.

Garvey, E. P., Oplinger, J. A., Furfine, E. S., Kiff, R. J., Laszlo, F., Whittle, B. J. and Knowles, R. G. (1997). 1400W is a slow, tight binding, and highly selective inhibitor of inducible nitric-oxide synthase in vitro and in vivo. *J. Biol. Chem.* **272**. 4959-4963.

Geller, D. A., Lowenstein, C. J., Shapiro, R. A., Nussler, A. K., Disilvio, M., Wang, S. C., Nakayama, D. K., Simmons, R. L., Snyder, S. H. and Billiar, T. R. (1993). Molecular cloning and expression of inducible nitric oxide synthase from human hepatocytes. *Proc. Natl. Acad. Sci. USA*. **90**. 3491-3495.

Gellner, K. and Brenner, S. (1999). Analysis of 148 kb of genomic DNA around the wnt1 locus of *Fugu rubripes*. *Genome Res.* **9**. 251-258.

Geng, Y. J., Almqvist, M. and Hansson, G. K. (1994). cDNA cloning and expression of inducible nitric-oxide synthase from rat vascular smooth muscle cells. *Biochim. Biophys. Acta*. **1218**. 421-424.

Gerber, N. C. and Ortiz De Montellano, P. R. (1995). Neuronal nitric oxide synthase - Expression in Escherichia coli, irreversible inhibition by phenyldiazene, and active site topology. *J. Biol. Chem.* **270**. 17791-17796.

Gilley, J. and Fried, M. (1999). Extensive gene order differences within regions of conserved synteny between *Fugu* and human genomes: implications for chromosomal evolution and the cloning of disease genes. *Hum. Mol. Genet.* **8**. 1313-1320.

Gnanapandithen, K., Chen, Z., Kau, C-L., Gorczynski, R. M. and Marsden, P. A. (1996). Cloning and characterization of murine endothelial constitutive nitric oxide synthase. *Biochim. Biophys. Acta Gene Struct. Expr.* **1308**. 103-106.

Göttgens, B., Gilbert, J. G. R., Barton, L. M., Aparicio, S., Hawker, K., Mistry, S., Vaudin, M., King, A., Bentley, D., Elgar, G. and Green, A. R. (1998). The pufferfish SLP-1 gene, a new member of the SCL/TAL-1 family of transcription factors. *Genomics*. **48**. 52-62.

Grabowski, P. S., Laing, K. J., McGuigan, F. E., Hardie, L. J., Ralston, S. H. and Secombes, C. J. (1995). Detection of mRNA for a nitric oxide synthase in macrophages and gill of rainbow trout challenged with an attenuated pathogen. In: *'The Biology of Nitric Oxide'*. Editors; Moncada, S., Stamler, J., Gross, S. and Higgs, E. A. Portland Press. Page 48.

Green, L. C., Tannenbaum, S. R. and Goldman, P. (1981a). Nitrate synthesis in the germfree and conventional rat. *Science*. **212**. 56-58.

Green, L. C., Ruiz de Luzuriaga, K., Wagner, D. A., Rand, W., Istfan, N., Young, V. R. and Tannenbaum, S. R. (1981b). Nitrate biosynthesis in man. *Proc. Natl. Acad. Sci. USA*. **78**. 7764-7768.

Grewal, P. K., Todd, L. C., Van der Maarel, S., Frants, R. R. and Hewitt, J. E. (1998). FRG1, a gene in the FSH muscular dystrophy region on human chromosome 4q35, is highly conserved in vertebrates and invertebrates. *Gene*. **216**. 13-19.

Griffith, O. W. and Kilbourn, R. G. (1996). Nitric oxide synthase inhibitors: amino acids. *Methods Enzymol.* **268**: 375-92

Hall., A. V., Antoniou, H., Wang, Y., Cheung, A. H., Arbus, A. M., Olson, S. L., Lu, W. C., Kau, C-L. and Marsden, P. A. (1994). Structural organization of the human neuronal nitric oxide synthase gene (*NOS1*). *J. Biol. Chem.* **269**. 33082-33090.

Hantraye, P., Brouillet, E., Ferrante, R., Palfi, S., Dolan, R., Matthews, R. T. and Beal, M. F. (1996). Inhibition of neuronal nitric oxide synthase prevents MPTP-induced parkinsonism in baboons. *Nat. Med.* **2**. 1017-1021.

Hardman, J. G. and Sutherland, E. W. (1969). Guanyl cyclase, an enzyme catalyzing the formation of guanosine 3′,5′-monophosphate from guanosine triphosphate. *J. Biol. Chem.* **244**. 6363-6370.

Harteneck, C., Klatt, P., Schmidt, K. and Mayer, B. (1994). Expression of rat brain nitric oxide synthase in baculovirus-infected insect cells and characterization of the purified enzyme. *Biochem. J.* **304**. 683-686.

Hendricks, W. (1995). Neuronal nitric oxide synthase contains a discs-large homologous region (DHR) sequence motif. *Biochem. J.* **305**. 687-688.

Henry, Y., Ducrocq, C., Drapier, J. C., Servent, D., Pellat, C. and Guissani, A. (1991). Nitric-oxide, a biological effector. Electron paramagnetic resonance detection of nitrosyl-iron-protein complexes in whole cells. *Eur. Biophys. J.* **20**. 1-15.

Hibbs, J. B. Jr., Vavrin, Z. and Taintor, R. A. (1987). L-arginine is required for expression of the activated macrophage effector mechanism causing selective metabolic inhibition in target cells. *Biochem. J.* **260**. 243-248.

Hibbs, J. B. Jr., Taintor, R. R., Vavrin, Z. and Rachlin, E. M. (1988). Nitric oxide: a cytotoxic activated macrophage effector molecule. *Biochem. Biophys. Res. Commun.* **157**. 87-94.

Hillier, B. J.; Christopherson, K. S., Prehoda, K. E., Bredt, D. S. and Lim, W. A. (1999). Unexpected modes of PDZ domain scaffolding revealed by structure of nNOS-syntrophin complex. *Science*. **284**. 812-815.

Hinegardner, R. (1968). Evolution of cellular DNA content in teleost fishes. *The American Naturalist.* **102**. 517-523.

Hinegardner, R. and Rosen, D. E. (1972). Cellular DNA content and the evolution of teleostean fishes. *The American Naturalist.* **106**. 621-644.

Holmqvist, B. I., Östholm, T., Alm, P. and Ekström, P. (1994). Nitric oxide synthase in the brain of a teleost. *Neurosci. Lett.* **171**. 205-208.

How, G. F., Venkatesh, B. and Brenner, S. (1996). Conserved linkage between the puffer fish (*Fugu rubripes*) and human genes for platelet-derived growth factor receptor and macrophage colony-stimulating factor receptor. *Genome Res.* **6**. 1185-1191.

Huang, P. L., Dawson. T. M., Bredt, D. S., Snyder, S. H. and Fishman, M. C. (1993). Targeted Disruption of the Neuronal Nitric Oxide Synthase Gene. *Cell.* **75**. 1273-1286.

Hylland, P. and Nilsson, G. E. (1995). Evidence that acetylcholine mediates increased cerebral blood flow velocity in Crucian Carp through a nitric oxide-dependent mechanism. *J. Cereb. Blood Flow Metab.* **15**. 519-524.

Ignarro, L. J., Harbison, R. G., Wood, K. S. and Kadowitz, P. J. (1986). Activation of purified soluble guanylate cyclase by endothelium-derived relaxing factor from intrapulmonary artery and vein: stimulation by acetylcholine, bradykinin and archidonic acid. *J. Pharmacol. Exp. Ther.* **237**. 893-900.

Ignarro, L. J., Buga, G. M., Wood, K. S., Byrns, R. E. and Chaudhuri, G. (1987). Endothelium-derived relaxing factor produced and released from artery and vein is nitric oxide. *Proc. Natl. Acad. Sci. USA.* **84**. 9265-9269.

Ignarro, L. J., Fukuto, J. M., Griscavage, J. M., Rogers, N. E. and Byrns, R. E. (1993). Oxidation of nitric oxide in aqueous solution to nitrite but not nitrate: comparison with enzymatically formed nitric oxide from L-arginine. *Proc Natl. Acad. Sci. USA.* **90**. 8103-8107.

Ishikawa, E., Ishikawa, S., Davis, J. W., and Sutherland, E. W. (1969). Determination of guanosine 3′,5′-monophosphate in tissues and of guanyl cyclase in rat intestine. *J. Biol. Chem.* **244**. 6371-6376.

Janssens, S. P., Shimouchi, A., Quertermous, T., Bloch, D. B. and Bloch, K. D. (1992). Cloning and Expression of a cDNA encoding human endothelium-derived relaxing factor/nitric-oxide synthase. *J. Biol. Chem.* **267**. 14519-14522.

Jeong,Y. and Yim, J. (1997). Molecular cloning of a cDNA encoding a constitutive nitric oxide synthase from rabbit brain. Direct submission to EMBL/GenBank/DDBJ databases. Accession Number U91584.

Jones, S. J. M. (1995). An Update And Lessons From Whole-Genome Sequencing Projects. *Curr. Opin. Genet. Dev.* **5**. 349-353.

Karlsen, A. E, Andersen, H. U., Vissing, H., Larsen, P. M., Fey, S. J., Cuartero, B. G., Madsen, O. D., Petersen, J. S., Mortensen, S. B., Mandruppoulsen, T. , Boel, E. and Nerup, J. (1995). Cloning and expression of cytokine-inducible nitric oxide synthase cDNA from rat islets of langerhans. *Diabetes*. **44**. 753-758.

Kehrer-Sawatzki, H., Maier, C., Moschgath, E. Elgar, G. and Krone, W. (1998). Genomic characterization of the Neurofibromatosis Type 1 gene of *Fugu rubripes*. *Gene*. **222**. 145-153.

Kehrer-Sawatzki, H., Maier, C., Moschgath, E., Elgar, G. and Krone, W. (1999). Characterization of three genes, AKAP84, BAW and WSB1, located 3′ to the Neurofibromatosis type 1 locus in *Fugu rubripes*. *Gene*. **235**. 1-11.

Khan, M. T. and Furchgott, R. F. (1987). Additional evidence that endothelium-derived relaxing factor is nitric oxide. In Pharmacology (edited by Rand, M.J. and Raper, C.). 341-344.

Kishimoto, J., Spurr, N., Liao, M., Lizhi, L., Emson, P. and Xu, W. (1992). Localization of brain nitric oxide synthase (NOS) to human chromosome 12. *Genomics*. **14**. 802-804.

Koide, M., Kawahara, Y., Tsuda, T., Nakayama, I. and Yokoyama, M. (1994). Expression of nitric oxide synthase by cytokines in vascular smooth muscle cells. *Hypertension*. **23**. 145-148.

Kone, B. C., Schwobel, J., Turner, P., Mohaupt, M. G. and Cangro, C. B. (1995). Role of NF- kappa B in the regulation of inducible nitric oxide synthase in an MTAL cell line. *Am. J. Physiol. Renal Fluid Electrolyte Physiol*. **269**. 38-45.

Koop, B. F. and Nadeau, J. H. (1996). Pufferfish and a new paradigm for comparative genome analysis. *Proc. Natl. Acad. Sci. USA*. **93**. 1363-1365.

Koprowski, H., Zheng, Y. M., Heber, K. E., Fraser, N., Rorke, L., Fu, Z. F., Hanlon, C. and Dietzschold, B. (1993). In vivo expression of inducible nitric oxide synthase in experimentally induced neurologic diseases. *Proc. Natl. Acad. Sci. USA*. **90**. 3024-3027.

Korneev, S. A., Piper, M. R., Picot, J., Phillips, R., Korneeva, E. I. and O'Shea, M. (1998) Molecular characterization of NOS in a mollusc: expression in a giant modulatory neuron. *J. Neurobiol.* **35**. 65-76.

Kornfeld, K., Hom, D. B. and Horvitz, H. R. (1995). The *ksr-1* gene encodes a novel protein kinase involved in RAS-mediated signaling in *C. elegans*. *Cell*. **83**. 903-913.

Kosuga, K., Yui, Y., Hattori, R., Sase, K., Eizawa, H., Aoyama, T., Inoue, R. and Sasayama, S. (1994). Cloning of an inducible nitric oxide synthase from rat polymorphonuclear neutrophils. *Endothelium J. Endothelial Cell Res.* **2**. 217-221.

Kunz, D., Muhl, H., Walker, G. and Pfeilschifter, J. (1994). Two distinct signaling pathways trigger the expression of inducible nitric oxide synthase in rat renal mesangial cells. *Proc. Natl. Acad. Sci. USA*. **91**. 5387-5391.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of bacteriophage T4. *Nature*. **227**. 680-685.

Laing, K. J., Grabowski, P. S., Belosevic, M. and Secombes, C. J. (1996). A partial sequence for nitric oxide synthase from a goldfish (*Carassius auratus*) macrophage cell line. *Immunol. Cell. Biol.* **74**. 374-379.

Lamas, S., Marsden, P. A., Li, G. K., Tempst, P. and Michel, T. (1992). Endothelial nitric oxide synthase: Molecular cloning and characterization of a distinct constitutive enzyme isoform. *Proc. Natl. Acad. Sci. USA*. **89**. 6348-6352.

Larsson, B. and Phillips, S. C. (1998). Isolation and cloning of a novel, human neuronal nitric oxide synthase cDNA. *Biochem. Biophys. Res. Commun.* **251**. 898-902.

Lee, M. A., Cai, L., Hubner, N., Lee, Y. A. and Lindpaintner, K. (1997). Tissue and development-specific expression of multiple alternatively spliced transcripts of rat neuronal nitric oxide synthase. *J. Clin. Invest.* **100**. 1507-1512.

Lin, A. W., Chang, C. C. and McCormick, C. C. (1996). Molecular cloning and expression of an avian macrophage nitric-oxide synthase cDNA and the analysis of the genomic 5′-flanking region. *J. Biol. Chem.* **271**. 11911-11919.

List, B. M., Klatt, P., Werner, E. R., Schmidt, K. and Mayer, B. (1996). Overexpression of neuronal nitric oxide synthase in insect cells reveals requirement of haem for tetrahydrobiopterin binding. *Biochem. J.* **315**. 57-63.

Liu, J. W., Garcia-Cardena, G. and Sessa, W. C. (1995). Biosynthesis and palmitoylation of endothelial nitric oxide synthase: Mutagenesis of palmitoylation sites, cysteine-15 and/or cysteine-26, argues against depalmitoylation-induced translocation of the enzyme. *Biochemistry*. **34**. 12333-12340.

Llevadot, R., Estivill, X., Scambler, P. and Pritchard, M. (1998). Isolation and genomic characterization of the *TUPLE1/HIRA* gene of the pufferfish *Fugu rubripes*. *Gene*. **208**. 279-283.

Lorsbach, R. B., Murphy, W. J., Lowenstein, C. J., Snyder, S. H. and Russell, S. W. (1993). Expression of the nitric oxide synthase gene in mouse macrophages activated for tumor cell killing. Molecular basis for the synergy between interefon-γ and lipopolysaccharide. *J. Biol. Chem.* **268**. 1908-1913.

Lowe, P. N., Smith, D., Stammers, D. K., Riveros-Moreno, V., Moncada, S., Charles, I. and Boyhan, A. (1996). Identification of the domains of neuronal nitric oxide synthase by limited proteolysis. *Biochem. J.* **314**. 55-62.

Lowenstein, C. J., Glatt, C. S., Bredt, D. S. and Snyder, S. H. (1992). Cloned and expressed macrophage nitric oxide synthase contrasts with the brain enzyme. *Proc. Natl. Acad. Sci. USA.* **89**. 6711-6715.

Luckhart, S., Vodovotz, Y., Cui, L. and Rosenberg, R. (1998). The mosquito *Anopheles stephensi* limits malaria parasite development with inducible synthesis of nitric oxide. *Proc. Natl. Acad. Sci. USA.* **95**. 5700-5705.

Luckhart, S. and Rosenberg, R. (1999). Gene structure and polymorphism of an invertebrate nitric oxide synthase gene. *Gene.* **232**. 25-34.

Luscher, T. F., Richard, V. and Tanner, F. C. (1991). Endothelium-derived vasoactive factors and their role in the coronary circulation. *Trends Cardiovasc. Med.* **1**. 179-185.

Lustig, H. S., Ahern, K. V. and Greenberg, D. A. (1992). Anti-Parkinsonian drugs and in vitro excitotoxicity. *Brain Res.* **597**. 148-150.

Lyons, C. R., Orloff, G. J. and Cunningham, J. M. (1992). Molecular cloning and functional expression of an inducible nitric oxide synthase from a murine macrophage cell-line. *J. Biol. Chem.* **267**. 6370-6374.

MacNaul, K. L. and Hutchinson, N. I. (1993). Differential expression of iNOS and cNOS mRNA in human vascular smooth muscle cells and endothelial cells under normal and inflammatory conditions. *Biochem. Biophys. Res. Commun.* **196**. 1330-1334.

Maier, R., Bilbe, G., Rediske, J. and Lotz, M. (1994). Inducible nitric oxide synthase from human articular chondrocytes: cDNA cloning and analysis of mRNA expression. *Biochim. Biophys. Acta.* **1208**. 145-150.

Magee, T., Fuentes, A. M., Garban, H., Rajavashisth, T., Marquez, D., Rodriguez, J. A., Rajfer, J. and Gonzalez-Cadavid., N. F. (1996). Cloning of a novel neuronal nitric oxide synthase expressed in penis and lower urinary tract. *Biochem. Biophys. Res. Commun.* **226**. 145-151.

Markewitz, B. A., Michael, J. R. and Kohan, D. E. (1993). Cytokine-induced expression of a nitric oxide synthase in rat renal tubule cells. *J. Clin. Invest.* **91**. 2138-2143.

Marletta, M. A., Yoon, P. S., Iyengar, R., Leaf, C. D. and Wishnok, J. S. (1988). Macrophage oxidation of L-arginine to nitrite and nitrate: Nitric oxide is an intermediate. *Biochemistry.* **27**. 8706-8711.

Marletta, M. A. (1993). Nitric oxide synthase structure and mechanism. *J. Biol. Chem.* **268**. 12231-12234.

Marletta, M. A. (1994). Nitric oxide synthase: Aspects concerning structure and catalysis. *Cell.* **78**. 927-930.

Marsden, P. A., Schappert, K. T., Chen, H. S., Flowers, M., Sundell, C. L., Wilcox, J. N., Lamas, S. and Michel, T. (1992). Molecular cloning and characterization of human endothelial nitric oxide synthase. *FEBS Lett.* **307**. 287-293.

Marsden, P. A., Antoniou, H., Cheung, A., Arbus, A. and Hall, A. V. (1993). Structure and chromosomal localization of the human constitutive endothelial nitric oxide synthase gene. *J. Biol. Chem.* **268**. 17478-17488.

Marsden, P. A., Antoniou, H., Cheung, A., Arbus, A. and Hall, A. V. (1994a). Characterization of the human neuronal nitric oxide synthase (NOS) gene: Novel allelic and structural mRNA diversity. *Proc. Int. Conf. Biochem. Mol. Biol. Nitric Oxide, 1st* (abstr.).

Marsden, P A, Heng, H. H. Q., Duff, C. L., Shi, X. M., Tsui, L. C. and Hall, A. V. (1994b). Localization of the human gene for inducible nitric-oxide synthase (NOS2) to chromosome 17q11.2-q12. *Genomics.* **19**. 183-185.

Marshall, H., Studer, M., Popperl, H., Aparicio, S., Kuroiwa, A., Brenner, S. and Krumlauf, R. (1994). A conserved retinoic acid response element required for early expression of the homeobox gene hoxb-1. *Nature.* **370**. 567-571.

Martásek, P., Liu, Q., Liu, J., Roman, L., Gross, S. S., Sessa, W. C. and Masters, B. S. S. (1996). Characterization of bovine endothelial nitric oxide synthase (eNOS) expressed in E. coli. *Biochem. Biophys. Res. Commun.* **219**. 359-365.

Martásek, P., Miller, R. T., Liu, Q., Liu, J., Roman, L., Salerno, J. C., Migita, C. T., Raman, C. S., Gross, S. S., Ikeda-Saito, M. and Masters, B. S. S. (1998). The C331A mutant of neuronal nitric oxide synthase is defective in arginine binding. *J. Biol. Chem.* **273**. 34799-34805.

Martinez, A., Riveros-Moreno, V., Polak, J. M., Moncada, S. and Sesma, P. (1994). Nitric oxide (NO) synthase immunoreactivity in the starfish Marthasterias glacialis. *Cell Tissue Res.* **275**. 599-603.

Mason, P. J., Stevens, D. J., Luzzatto, L., Brenner, S. and Aparicio, S. (1995). Genomic Structure and Sequence of the *Fugu Rubripes* Glucose-6-Phosphate-Dehydrogenase Gene (G6PD). *Genomics.* **26**. 587-591.

McMillan, K., Bredt, D. S,, Hirsch, D. J., Snyder, S. H., Clark, J. E. and Masters, B. S. S. (1992). Cloned, expressed rat cerebellar nitric oxide synthase contains stoichiometric amounts of heme, which binds carbon monoxide. *Proc. Natl. Acad. Sci. USA.* **89**. 11141-11145.

McMillan, K. and Masters, B. S. S. (1995). Prokaryotic expression of the heme-binding and flavin-binding domains of rat neuronal nitric oxide synthase as distinct polypeptides: Identification of the heme-binding proximal thiolate ligand as cysteine-415. *Biochemistry.* **34**. 3686-3693.

Michaud, N. R., Thierren, M., Cacace, A., Edsall, L. C., Spiegel, S., Rubin, G. M. and Morrison, D. K. (1997). KSR stimulates Raf-1 activity in a kinase-independent manner. *Proc. Natl. Acad. Sci. USA.* **94**. 12792-12796.

Miles, C., Elgar, G., Coles, E., Kleinjan, D. J., Van Heyningen, V. and Hastie, N. (1998). Complete sequencing of the *Fugu* WAGR region from WT1 to PAX6: Dramatic compaction and conservation of synteny with human chromosome 11p13. *Proc. Natl. Acad. Sci. USA.* **95**. 13068-13072.

Miller, V. M. and Vanhoutte, P. M. (1986). Endothelium-dependent responses in isolated blood vessels of lower vertebrates. *Blood Vessels.* 23. 225-235.

Miller, R. T., Martásek, P., Raman, C. S. and Masters, B. S. S. (1999). Zinc content of *Escherichia coli*-expressed constitutive isoforms of nitric oxide synthase. Enzymatic activity and effect of pterin. *J. Biol. Chem.* **274**. 14537-14540.

Moncada, S. and Higgs, A. (1993). Mechanisms of disease - the L-arginine-nitric oxide pathway. *N. Engl. J. Med.* **329**. 2002-2012.

Moncada, S. and Palmer, R. M. J. (1990). The L-arginine:nitric oxide pathway in the vessel wall. In: *'Nitric oxide from L-arginine: a bioregulatory system'*. Editors; Moncada, S. and Higgs, E. A. Portland Press. Pages 19-33.

Moss, D. W., Wei, X., Liew, F. Y., Moncada, S. and Charles, I. G. (1995). Enzymatic characterisation of recombinant murine inducible nitric oxide synthase. *Eur. J. Pharmacol.* **289**. 41-48.

Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**. 459-472.

Mulsch, A. Bohme, E. and Busse., R. (1987). Stimulation of soluble guanylate cyclase by endothelium-derived relaxing factor from cultured endothelial cells. *Eur. J. Pharmacol.* **135**. 247-250.

Mundel, P., Bachmann, S., Bader, M., Fischer, A., Kummer, W., Mayer, B. and Kriz, W. (1992). Expression of nitric oxide in kidney macula densa cells. *Kidney Int.* **42**. 1017-1019.

Nadaud, S., Bonnardeaux, A., Lathrop, M. and Soubrier, F. (1994). Gene structure, polymorphism and mapping of the human endothelial nitric oxide synthase gene. *Biochem. Biophys. Res. Commun.* **198**. 1027-1033.

Naito, T., Saito, Y., Yamamoto, J., Nozaki, Y., Tomura, K., Hazama, M., Nakanishi, S. and Brenner, S. (1998). Putative pheromone receptors related to the Ca(2+)-sensing receptor in *Fugu. Proc. Natl. Acad. Sci. USA.* **95**. 5178-5181.

Nakane, M., Schmidt, H. H. H. W., Pollock, J. S., Förstermann, U. and Murad, F. (1993). Cloned human brain nitric oxide synthase is highly expressed in skeletal muscle. *FEBS Lett.* **316**. 175-180.

Nakane, M., Pollock, J. S., Klinghofer, V., Basha, F., Marsden, P. A., Hokari, A., Ogura, T., Esumi, H. and Carter, G. W. (1995). Functional expression of 3 isoforms of human nitric oxide synthase in baculovirus-infected insect cells. *Biochem. Biophys. Res. Commun.* **206**. 511-517.

Narhi, L. O. and Fulco, A. J. (1986). Characterization of a catalytically self-sufficient 119,000-dalton cytochrome-P-450 monooxygenase induced by barbiturates in *Bacillus megaterium*. *J. Biol. Chem.* **261**. 7160-7169.

Needleman and Wunsch. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**. 443-453.

Nehls, M., Luno, K., Schorpp, M., Pfeifer, D., Krause, S., Matysiak-Scholze, U., Dierbach, H. and Boehm, T. (1995). YAC/P1 contigs defining the location of 56 microsatellite markers and several genes across a 3.4-cM interval on mouse chromosome 11. *Mamm. Genome.* **6**. 321-31.

Nelson, D. R., Kamataki, T., Waxman, D. J., Guengerich, F. P., Estabrook, R. W., Feyereisen, R., Gonzalez, F. J., Coon, M. J., Gunsalus, I. C., Gotoh, O., Okuda, K. and Nebert, D. W. (1993). The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA and Cell Biol.* **12**. 1-51.

Neumann, N. F., Fagan, D. and Belosevic, M. (1995). Macrophage activating factor(s) secreted by mitogen stimulated goldfish kidney leukocytes synergize with bacterial lipopolysaccharide to induce nitric oxide production in teleost macrophages. *Dev. Comp. Immunol.* **19**. 473-482.

Nishida, K., Harrison, D. G., Navas, J. P., Fisher, A. A., Dockery, S. P., Uematsu, M., Nerem, R. M., Alexander, R. W. and Murphy, T. J. (1992). Molecular cloning and characterization of the constitutive bovine aortic endothelial cell nitric oxide synthase. *J. Clin. Invest.* **90**. 2092-2096.

Nishimura, J. S., Martasek, P., Mcmillan, K., Salerno, J. C., Liu, Q., Gross, S. S and, Masters, B. S. S. (1995). Modular structure of neuronal nitric oxide synthase: Localization of the arginine binding site and modulation by pterin. *Biochem. Biophys. Res. Commun.* **210**. 288-294.

Norris, P. J., Charles, I. G., Scorer, C. A. and Emson, P, C. (1995). Studies on the localisation and expression of nitric oxide synthase using histochemical techniques. *Histochem. J.* **27**. 745-756.

Oberbäumer, I., Moser, D. and Bachmann, S. (1998). Nitric Oxide Synthase 1 mRNA: Tissue-Specific Veriants from Rat with Alternative First Exons. *Biol. Chem.* **379**. 913-919.

Ogura, T., Yokoyama, T., Fujisawa, H., Kurashima, Y. and Esumi, H. (1993). Structural Diversity of Neuronal Nitric Oxide Synthase mRNA in the Nervous System. *Biochem. Biophys. Res. Commun.* **193**. 1014-1022.

Östholm, T., Holmqvist, B. I., Alm, P. and Ekström, P. (1994). Nitric oxide synthase in the CNS of the atlantic salmon. *Neurosci. Lett.* **168**. 233-237.

Oyan A. M., Goksoyr, A. and Holmqvist, B. (1998). Partial cloning of neuronal nitric oxide synthase of Atlantic salmon and expression of mRNAs in brain tissue. Direct submission to EMBL/GenBank/DDBJ data banks. Accession Number AJ006209.

Palmer, R. M. J., Ferrige., A. G. and Moncada, S. (1987). Nitric oxide release accounts for the biological activity of endothelium derived relaxing factor. *Nature*. **327**. 524-526.

Palmer, R. M. J., Ashton, D. S. and Moncada, S. (1988a). Vascular endothelial cells synthesize nitric oxide from L-arginine. *Nature*. **333**. 664-666.

Palmer, R. M. J., Rees, D. D., Ashton, D. S. and Moncada, S. (1988b). L-Arginine is the physiological precursor for the formation of nitric oxide in endothelium-dependent relaxation. *Biochem. Biophys. Res. Commun.* **153**. 1251-1256.

Palmer, R. M. J. and Moncada, S. (1989). A novel citrulline-forming enzyme implicated in the formation of nitric oxide by vascular endothelial cells. *Biochem. Biophys. Res. Commun.* **158**. 348-352.

Palmer, R. M., Hickery, M. S., Charles, I. G., Moncada, S. and Bayliss, M. T. (1993). Induction of nitric oxide synthase in human chondrocytes. *Biochem. Biophys. Res. Commun.* **193**. 398-405.

Pearson,W. R. and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci, USA*. **85**. 2444-2448.

Perrella, M. A., Yoshizumi, M., Fen, Z., Tsai, J. C., Hseih, C. M., Kourembanas, S. and Lee, M. E. (1994). Transforming growth factor-β_1, but not dexamethasone, down-regulates nitric-oxide synthase mRNA after its induction by interleukin-1 β in rat smooth muscle cells. *J. Biol. Chem.* **269**. 14595-14600.

Perry, J. M., Moon, N., Zhao, Y., Dunham, W. R. and Marletta, M. A. (1998). The high-potential flavin and heme of nitric oxide synthase are not magnetically linked: implications for electron transfer. *Chem. Biol.* **5**. 355-364.

Pizon, V., Cuny, G. and Bernerdi, G. (1984). Nucleotide sequence organization in the very small genome of a tetraodontid fish, *Arothton diadematus. Eur. J. Biochem.* **140**. 25-30.

Ponting, C. P. and Phillips, C. (1995). DHR domains in syntrophins, neuronal NO synthases and other intracellular proteins. *Trends Biochem. Sci.* **20**. 102-103.

Poulter, R. and Butler, M. (1998). A retrotransposon family from the pufferfish (*fugu*) *Fugu rubripes*. *Gene*. **215**. 241-249.

Poulter, R., Butler, M. and Ormandy, J. (1999). A LINE element from the pufferfish (*Fugu*) *Fugu rubripes* which shows similarity to the CR1 family of non-LTR retrotransposons. *Gene*. **227**. 169-179.

Powers, D. A. (1991). Evolutionary genetics of fish. *Adv. Genet.* **29**. 119-228.

Radi, R., Cosgrove, T. P., Beckman, J. S, Freeman, B. A. (1993). Peroxynitrite-induced luminol chemiluminescence. *Biochem. J.* **290**.51-57.

Radomski, M. W., Martin, J. F and Moncada, S. (1991). Synthesis of nitric-oxide by the hemocytes of the American horseshoe-crab (*Limulus polyphemus*). *Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences*. **334**. 129-133.

Raman, C. S., Li, H., Martásek, P., Král, V., Masters, B. S. S. and Pouios, T. L. (1998). Crystal structure of constitutive endothelial nitric oxide synthase: A paradigm for pterin function involving a novel metal center. *Cell*. **95**. 939-950.

Rand, M. J. (1992). Nitrergic transmission: Nitric oxide as a mediator of nonadrenergic, noncholinergic neuroeffector transmission. *Clin. Exp. Pharmacol. Physiol.* **19**. 147-169.

Rast, J. P., Haire, R. N., Litman, R. T., Pross, S. and Litman, G. W. (1995). Identification and characterization of T-cell antigen receptor-related genes in phylogenetically diverse vertebrate species. *Immunogenet.* **42**. 204-212.

Reboul, J., Gardiner, K., Monneron, D., Uze, G. and Lutfalla, G. (1999). Comparative genomic analysis of the interferon/interleukin-10 receptor gene cluster. *Genome Res.* **9**. 242-250.

Rees, D. D., Palmer, R. M. J., Hodson, H. F. and Moncada, S. (1989b). A specific inhibitor of nitric oxide formation from L-arginine attenuates endothelium-dependent relaxation. *Br. J. Pharmacol.* **96**. 418-424.

Regulski, M. and Tully, T. (1995). Molecular and biochemical characterization of dNOS: A Drosophila Ca^{2+}/calmodulin-dependent nitric oxide synthase. *Proc. Natl. Acad. Sci. USA*. **92**. 9072-9076.

Richards, M. K. and Marletta, M. A. (1994). Characterization of neuronal nitric oxide synthase and a C415H mutant, purified from a baculovirus overexpression system. *Biochemistry*. **33**. 14723-14732.

Riveros-Moreno, V., Heffernan, B., Torres, B., Chubb, A., Charles, I. and Moncada, S. (1995). Purification to homogeneity and characterization of rat brain recombinant nitric oxide synthase. *Eur. J. Biochem.* **230**. 52-57.

Robinson, L. J., Weremowicz, S., Morton, C. C. and Michel, T. (1994). Isolation and chromosomal localization of the human endothelial nitric oxide synthase (NOS3) gene. *Genomics*. **19**. 350-357.

Rodriguez-Crespo, I. and Ortiz de Montellano, P. R. (1996). Human endothelial nitric oxide synthase: Expression in Escherichia coli, coexpression with calmodulin, and characterization. *Arch. Biochem. Biophys.* **336**. 151-156.

Rodríguez-Crespo, I., Moënne-Loccoz, P., Loehr, T. M. and Ortiz de Montellano, P. R. (1997). Endothelial nitric oxide synthase: Modulations of the distal heme site produced by progressive N-terminal deletions. *Biochemistry*. **36**. 8530-8538.

Roman, L. J., Sheta, E. A., Martasek, P., Gross, S. S., Liu, Q. and Masters, B. S. S. (1995). High-level expression of functional rat neuronal nitric oxide synthase in Escherichia coli. *Proc. Natl. Acad. Sci. USA.* **92**. 8428-8432.

Ruff, P., Speicher, D. W. and Husain-Chishti, A. (1991). Molecular identification of a major palmitoylated erythrocyte membrane protein containing the src homology 3 motif. *Proc. Natl. Acad. Sci. USA.* **88**. 6595-6599.

Saeij, J.P. (1999). Cyprinus carpio mRNA for inducible nitric oxide synthase (iNOS gene). Direct submission to EMBL/GenBank/DDBJ databases. Accession Number AJ242906.

Sagami, I. and Shimizu, T. (1998). The crucial roles of Asp-314 and Thr-315 in the catalytic activation of molecular oxygen by neuronal nitric-oxide synthase: A site-directed mutagenesis study. *J. Biol. Chem.* **273**. 2105-2108.

Sakuma, I., Stuehr, D. J., Gross, S. S., Nathan, C. and Levi, R. (1988). Identification of arginine as a precursor of endothelium-derived relaxing factor (EDRF). *Proc. Natl. Acad. Sci. USA.* **85**. 8664-8667.

Salerno, J. C., Harris, D. E., Irizarry, K., Patel, B., Morales, A. J., Smith, S. M. S., Martásek, P., Roman, L. J., Masters, B. S. S., Jones, C. L., Weissman, B. A., Lane, P., Liu, Q. and Gross, S. S. (1997). An autoinhibitory control element defines calcium-regulated isoforms of nitric oxide synthase. *J. Biol. Chem.* **272**. 29769-29777.

Sandford, R., Burn, T., Vaudin, M., Elgar, G. and Brenner, S. (1996). The *Fugu* genome: A novel tool for transcription mapping. *Cytogenet. Cell. Genet.* **72**. 20.

Sari, M. A., Booker, S., Jaouen, M., Vadon, S., Boucher, J. L., Pompon, D. and Mansuy, D. (1996). Expression in yeast and purification of functional macrophage nitric oxide synthase. Evidence for cysteine-194 as iron proximal ligand. *Biochemistry.* **35**. 7204-7213.

Sathasivam, K., Baxendale, S., Mangiarini, L., Bertaux, F., Hetherington, C., Kanazawa, I., Lehrach, H. and Bates, G. P. (1997). Aberrant processing of the *Fugu* HD (FrHD) mRNA in mouse cells and in transgenic mice. *Hum. Mol. Genet.* **6**. 2141-2149.

Scheinker, V., Stasiv, Y. and Enikolopv, G. (1999). Neuronal nitric oxide synthase cDNA from *Xenopus laevis*. Direct submission to EMBL/GenBank/DDBJ databases. Accession Number AAD55136.

Schepens, J., Cuppen, E., Wieringa, B. and Hendriks, W. (1997). The neuronal nitric oxide synthase PDZ motif binds to -G(D,E)XV* carboxyterminal sequences. *FEBS Lett.* **409**. 53-56.

Schmidt, H. H. H. W., Nau, H., Wittfoht, W., Gerlach, J., Prescher, K. E., Klein, M. M., Niroomand, F. and Bohme, E. (1988). Arginine is a physiological precursor of endothelium-derived nitric oxide. *Eur. J. Pharmacol.* **154**. 213-216.

Schmidt, H. H. W. and Murad, F. (1991). Purification and characterization of a human NO synthase. *Biochem. Biophys. Res. Commun.* **181**. 1372-1377.

Schmidt, H. H., Warner, T. D., Ishii, K., Sheng, H. and Murad, F. (1992). Insulin secretion from pancreatic B cells caused by L-arginine-derived nitrogen oxides. *Science.* **255**. 721-723.

Schmidt, H. H. H. W. and Walter, U. (1994). NO at work. *Cell.* **78**. 919-925.

Schober, A., Malz, C. R. and Meyer, D. L. (1993). Enzymehistochemical demonstration of nitric oxide synthase in the diencephalon of the rainbow trout (*Oncorhynchus myckiss*). *Neurosci. Lett.* **151**. 67-79.

Schofield, J. P., Elgar, G., Greystrong, J., Lye, G., Deadman, R., Micklem, G., King, A., Brenner, S. and Vaudin, M. (1997). Regions of human chromosome 2 (2q32-q35) and mouse chromosome 1 show synteny with the pufferfish genome (*Fugu rubripes*). *Genomics*. **45**. 158-167.

Schoor, W. P. and plumb, J. A. (1994). Induction of nitric oxide synthase in channel catfish *Ictalurus punctatus* by *Edwardsiella ictaluri*. *Dis. Aquat. Org.* **19**. 153-155.

Schüddekopf, K., Schorpp, M. and Boehm, T. (1996). The whn transcription factor encoded by the nude locus containsan evolutionarily conserved and functionally indispensable activation domain. *Proc. Natl. Acad. Sci. USA*. **93**. 9661-9664.

Schultz, G., Bohme, E. and Munske, K. (1969). Guanyl cyclase. Determination of enzyme activity. *Life Sci.* **8**. 1323-1332.

Schuman, E. M. and Madison, D. V. (1991). A requirement for the intercellular messenger nitric oxide in long-term potentiation. *Science*. **254**. 1503-1506.

Seo, H. G., Fujii, J., Soejima, H., Niikawa, N. and Taniguchi, N. (1995). Heme requirement for production of active endothelial nitric oxide synthase in baculovirus-infected insect cells. *Biochem. Biophys. Res. Commun.* **208**. 10-18.

Sessa, W. C., Harrison, J. K., Barber, C. M., Zeng, D., Durieux, M. E., Dangelo, D. D., Lynch K. R. and Peach, M. J. (1992). Molecular cloning and expression of a cDNA encoding endothelial cell nitric-oxide synthase. *J. Biol. Chem.* **267**. 15274-15276.

Sessa, W. C., Barber, C. M. and Lynch, K. R. (1993). Mutation of N-myristoylation site converts endothelial cell nitric oxide synthase from a membrane to a cytosolic protein. *Circ. Res.* **72**. 921-924.

Sherman, P. A., Laubach, V. E., Reep, B. R. and Wood, E. R.(1993). Purification and cDNA sequence of an inducible nitric oxide synthase from a human tumor cell line. *Biochemistry*. 32. 11600-11605.

Silvagno, F., Xia, H. and Bredt, D. S. (1996). Neuronal nitric-oxide synthase-μ, an alternatively spliced isoform expressed in differentiated skeletal muscle. *J. Biol. Chem.* **271**. 11204-11208.

Sinsheimer, R. (1989). The Santa Cruz Workshop, May 1985. *Genomics*. **5**. 954-965.

Sirsjo, A., Soderkvist, P., Soderkvist, T., Carlsson, M., Ost, M. and Gidlof, A. (1994). Different induction mechanisms of mRNA for inducible nitric oxide synthase in rat smooth muscle cells in culture and in aortic strips. *FEBS Lett.* **338**. 191-196.

Söderström, V., Hylland, P. and Nilsson, G. E. (1995). Nitric oxide synthase inhibitor blocks acetylcholine induced increase in brain blood flow in rainbow trout. *Neurosci. Lett.* **197**. 191-194.

Southan, G. J. and Szabó, C. (1996). Selective pharmacological inhibition of distinct nitric oxide synthase isoforms. *Biochem. Pharmacol.* **51**. 383-394.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**. 503-517.

Stuehr, D. J. and Marletta, M. A. (1985). Mammalian nitrate biosynthesis: mouse macrophages produce nitrite and nitrate in response to *Escherichia coli* lipopolysaccharide. *Proc. Natl. Acad. Sci. USA*. **82**. 7738-7742.

Stuehr, D. J, Gross, S., Sakuma, I., Levi, R. and Nathan, C. (1989). Activated murine macrophages secrete a metabolite of arginine with the bioactivity of endothelium-derived relaxing factor and the chemical reactivity of nitric oxide. *J. Exp. Med.* **169**. 1011-1120.

Stuehr, D. J., Cho, H. J, Kwon, N. S., Weise, M. F. and Nathan, C. F. (1991). Purification and characterization of the cytokine-induced macrophage nitric oxide synthase: An FAD- and FMN-containing flavoprotein. *Proc. Natl. Acad. Sci. USA.* **88**. 7773-7777.

Sundaram, M. and Han, M. (1995). The C. elegans *ksr-1* Gene Encodes a Novel Raf-Related Kinase Involved in Ras-Mediated Signal Transduction. *Cell.* **83**. 889-901.

Tabor, S. and Richardson, C. C. (1995). A single residue in DNA polymerases of the *Escherichia coli* DNA polymerase I family is critical for distinguishing between deoxy- and dideoxyribonucleotides. *Proc. Natl. Acad. Sci. USA.* **92**. 6339-6343.

Tassone, F., Villard, L., Clancy, K. and Gardiner, K. (1999). Structures, sequence characteristics, and synteny relationships of the transcription factor E4TF1, the splicing factor U2AF35 and the cystathionine beta synthetase genes from *Fugu rubripes*. *Gene.* **226**. 211-223.

Terada, Y., Tomita, K., Nonoguchi, H. and Marumo, F. (1992). Polymerase chain reaction localization of constitutive nitric oxide synthase and soluble guanylate cyclase messenger RNAs in microdissected rat nephron segments. *J. Clin. Invest.* **90**. 659-665.

Thierren, M., Chang, H. C., Solomon, N. M., Karim, F. D., Wassarman, D. A. and Rubin, G. M. (1995). KSR, a Novel Protein Kinase Required for RAS Signal Transduction. *Cell.* **83**. 879-888.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**. 4673-4680.

Trower, M. K., Orton, S. M., Purvis, I. J., Sanseau, P., Riley, J., Christodoulou, C., Burt, D., See, C. G., Elgar, G., Sherrington, R., Rogaev, E. I., StGeorgeHyslop, P., Brenner, S., Dykes, C. W. (1996). Conservation of synteny between the genome of the pufferfish (*Fugu rubripes*) and the region on human chromosome 14 (14q24.3) associated with familial Alzheimer disease (AD3 locus). *Proc. Natl. Acad. Sci. USA*. **93**. 1366-1369.

Tugendreich, S., Bassett Jr., D. E., McKusick, V. A., Boguski, M. S. and Hieter, P. (1994). Genes conserved in yeast and humans. *Hum. Mol. Genet.* **3**. 1509-1517.

US Department of Health and Human Services and US Department of Energy (1990). Understanding our genetic inheritance: the first five years, FY 1991-1995. DOE/ER-0452P. National Technical Information Service, Springfield, Virginia.

Vallance, P. (1996). Use of L-arginine and its analogs to study nitric oxide pathway in humans. *MethodsEnzymol.* **269**. 453-459.

Venema, R. C., Rong Zou, H. J., Ryan, J. W. and Venema, V. V. (1997). Subunit interactions of endothelial nitric oxide synthase. *J. Biol. Chem.* **272**. 1276-1282.

Verma-Kurvari, S. and Johnson, J. E. (1997). Identification of an achaete-schuate homolog, Fashl, from *Fugu rubripes*. *Gene*. **200**. 145-148.

Villard, L., Tassone, F., Crnogorac-Jurcevic, T., Clancy, K. and Gardiner, K. (1998). Analysis of pufferfish homologues of the AT-rich human APP gene. *Gene*. **210**. 17-24.

Vorherr, T., Knopfel, L., Hofmann, F., Mollner, S., Pfeuffer, T. and Carafoli, E. (1993). The calmodulin binding domain of nitric-oxide synthase and adenylyl cyclase. *Biochemistry*. **32**. 6081-6088.

Wang, Y. and Marsden, P. A. (1995). Nitric oxide synthases: gene structure and regulation. *Adv. Pharmacol.* **34**. 71-90.

Wang, Y., Goligorsky, M. S., Lin, M., Wilcox, J. N. and Marsden, P. A. (1997). A novel, testis specific mRNA transcript encoding an NH_2-terminal truncated nitric oxide synthase. *J. Biol. Chem.* **272**. 11392-11401.

Wang. Y., Newton, D. C., Robb, G. B., Kau, C-L., Miller, T. L., Cheung, A. H., Hall, A. V., Vandamme, S., Wilcox, J. N. and Marsden, P. A. (1999). RNA diversity has profound effects on the translation of neuronal nitric oxide synthase. *Proc. Natl. Acad. Sci. USA*. **96**. 12150-12155.

Weiner, C. P., Lizasoain, I., Baylis, S. A., Knowles, R. G., Charles, I. G. and Moncada, S. (1994). Induction of calcium-dependent nitric oxide synthases by sex hormones. *Proc. Natl. Acad. Sci. USA*. **91**. 5212-5216.

Wentworth, J. M., Schoenfeld, V., Meek, S., Elgar, G., Brenner, S. and Chatterjee, V. K. K. (1999). Isolation and characterisation of the retinoic acid receptor-α gene in the Japanese pufferfish, *F. rubripes*. *Gene*. **236**. 315-323.

Werner-Felmayer, G., Golderer, G., Werner, E. R., Grobner, P and Wachter, H. (1994). Purification and characterization of nitric oxide synthase from the low eukaryote Physarum polycephalum. *First Int. Cong. Biochem. Mol. Biol. Nitric Oxide*. **73**. A30.

White, A.A. and Aurbach, G.D. (1969). Detection of guanyl cyclase in mammalian tissues. *Biochem. Biophys. Acta*. **191**. 686-697.

Wilcox, C. S., Welch, W. J., Murad, F., Gross, S. S., Taylor, G., Levi, R. and Schmidt, H. H. (1992). Nitric oxide synthase in macula densa regulates glomerular capillary pressure. *Proc. Natl. Acad. Sci. USA*. **89**. 11993-11997.

Wolff, D. J. and Datto, G. A. (1992). Identification and characterization of a calmodulin-dependent nitric oxide synthase from GH-3 pituitary cells. *Biochem. J.* **285**. 201-206.

Woods, D. F. and Bryant, P. J. (1991). The discs-large tumor suppressor gene of Drosophila encodes a guanylate kinase homolog localized at septate junctions. *Cell*. **66**. 451-464.

Woods, D. F. and Bryant, P. J. (1993). ZO-1, DlgA and PSD-95/SAP90: Homologous proteins in tight, septate and synaptic cell junctions. *Mech. Dev.* **44**. 85-89.

Wu, C. Q., Zhang, J. G., Abu-Soud, H., Ghosh, D. K. and Stuehr, D. J. (1996). High-level expression of mouse inducible nitric oxide synthase in *Escherichia coli* requires co-expression with calmodulin. *Biochem. Biophys. Res. Commun.* **222**. 439-444.

Xie, J., Roddy, P., Rife, T., Murad, F. and Young, A. (1995). Two closely but separable promoters for human neuronal nitric oxide synthase gene transcription. *Proc. Natl. Acad. Sci. USA*. **92**. 1242-1246.

Xie, Q. W., Cho, H. J., Calaycay, J., Mumford, R. A., Swiderek, K. M., Lee, T. D., Ding, A. H., Troso, T. and Nathan, C. (1992). Cloning and characterization of inducible nitric oxide synthase from mouse macrophages. *Science*. **256**. 225-228.

Xu, W., Gorman, P., Sheer, D., Bates, G., Kishimoto, J., Lizhi, L. and Emson, P. (1993). Regional localization of the gene coding for human brain nitric oxide synthase (NOS1) to 12q24.2-24.31 by fluorescent in situ hybridization. *Cytogenet. Cell Genet.* **64**. 62-63.

Xu, W., Charles, I. G., Moncada, S., Gorman, P., Sheer, D., Liu, L. and Emson, P. (1994). Mapping of the genes encoding human inducible and endothelial nitric oxide synthase (NOS2 and NOS3) to the pericentric region of chromosome 17 and to chromosome 7, respectively. *Genomics*. **21**. 419-422.

Xu, W., Charles, I. G., Lui, L., Moncada, S. and Emson, P. (1996). Molecular cloning and structural organization of the human inducible nitric oxide synthase gene (NOS2). *Biochem. Biophys. Res. Commun.* **219**. 784-788.

Yao, C. W., Leu, J. H., Chin, C., Chou, C.K. and Huang, C. J. (1998). Round-spotted pufferfish (*Tetraodon fluviatilis*) snf5 gene is oriented in a tail-to-tail manner with the set gene which encodes an inhibitor of protein phosphatase 2A. *DNA Cell Biol.* **17**. 69-82.

Yamaguchi, F., Macrae, A. D. and Brenner, S. (1996). Molecular cloning of two cannabinoid type 1-like receptor genes from the puffer fish *Fugu rubripes*. *Genomics*. **35**. 603-605.

Yamaguchi, F. and Brenner, S. (1997). Molecular cloning of 5-hydroxytryptamine (5-HT) type 1 receptor genes from the Japanese puffer fish, *Fugu rubripes*. *Gene*. **191**. 219-223.

Yeo, G. S. H., Elgar, G., Sandford, R., Brenner, S. (1997). Cloning and sequencing of complement component C9 and its linkage to DOC-2 in the pufferfish *Fugu rubripes*. *Gene*. **200**. 203-211.

Yuda, M., Hirai, M., Miura, K., Matsumura, H., Ando, K. and Chinzei, Y. (1996). cDNA cloning, expression and characterization of nitric-oxide synthase from the salivary glands of the blood-sucking insect *Rhodnius prolixus*. *Eur. J. Biochem.* **242**. 807-812.

Zhang, M. and Vogel, H. J. (1994). Characterization of the calmodulin-binding domain of rat cerebellar nitric oxide synthase. *J. Biol. Chem.* **269**. 981-985.

APPENDIX

Table A.1. Primers used for the sequencing of the *Fugu* NOS1 gene

Primer	Sequence 5' → 3'	Position
ge1f2	CAAGCGTGCCTCGTCTCTAGAGAGGCAG	43 - 70
gi1f2	CGCACTACGATTGTTTGACTCCT	218 - 240
gi1f	GTCGGTTAAAGGACAAAAGAGGGG	339 - 362
gi1f3	AAGCCGAACATCCAGCGGGATGCG	734 - 757
gi1f4	CATACGCGTCCGTGTACGTGCGCG	1226 - 1249
gi1f5	AAGAGTCCGAGCCTTCCGTGTGCC	1587 - 1610
nNOS2FORA	GCCCTGGAGATGTTGAAGAATGTGCTGC	1820 - 1847
e2f	GAGATTGAACCAGTGCTGCGC	2171 - 2191
gi3b	TCCTCGAACTACCTGACCTAACCG	2485 - 2508
gi3a	CTCAGTTGCTGACACTTCCAGCTG	2643 - 2666
g3D1	CCACAGTCTCGTTGACTCAGTCTC	2985 - 3008
g3D2	CTCATTTGGAGTACGTGGATGAGG	3109 - 3132
g4D	GGTGTAGGAATCGACACATCTCTG	3458 - 3481
g4D1	AATCCACGACTGTCTCTGCCCAGC	3668 - 3691
n42	CCTCAGACCTGCAGAATACAGAGT	4086 - 4109
n48	ATGTGTGAGACTGTTCCAG	4274 - 4292
n55	TTGTGAGTCGGTAGTTCCTGTGTC	4708 - 4731
n65	CGCCTTTGTGTGTTCCCACATTAC	5158 - 5181
n66	ATGCTGTGCTCTTAGCATTGGCTG	5565 - 5588
n68fg	GACCTTCTTCCCTCCTGCATTTCC	5770 - 5793
n67fg	CAAATTGTCCCGTCAGTGTGGC	5954 - 5975
gi4f	CTCGGCATCACTGCGAGTCAAACA	6516 - 6539
g4D	AGAATGATGGAGAGGACTCAGAGC	6821 - 6844
n41rga	CGTGCATTAGACGGCTTTCGCTTC	7121 - 7144
n41rgb	TGCCCTCCGCTTCCACGTCAGTGC	7491 - 7514
n41rg	GTCGTGTCTGGAGTTACTGTCCAAAACCG	7822 - 7850
n5	GGTGACTTATCAAATGGGCTCTGC	8268 - 8291
n8	AGGTCTGAGTCAGTGGTGAGAAGAG	8617 - 8641
n14	TCATGTGTCACTGGCCACACG	8866 - 8886
n18	CACAGTGTTTGCATGCACATCGGT	8987 - 9010
n21	CCTTATGATGCCCAACCTGCACGC	9099 - 9122
n20	GGAGCGCAGGTATGTCACATTTGT	9323 - 9346
n26	GACCCTCATCTTACGCTCAGCTGCA	9691 - 9716
n28	GCTAACCGCTCCAACCAAACTGGC	9964 - 9987
n38	TTGATGCCAGAGACTGCACAACAG	10368 - 10391
n37f	GAGCCTCAGCCTTGCTGTGGGTAC	11226 - 11249
n29f	TGCCAATAATACACCCAAAGTGAG	11756 - 11779
n23	GAAGAAGGCATTTGAACTGCTCCG	12167 - 12190
835nNOS	ATGTACTGCGCGTACCCCTGAAGG	12434 - 12457
n19	TGTGGGATCCCACATTACCTTCTG	12776 - 12799
n25	TAACGTCCAGCCCGTTTGGACTAC	13112 - 13135
e12f	TCATGCAAAGTGACCATAGTGGAC	13377 - 13400

Primer	Sequence 5' → 3'	Position
n30	GCTCAATTACCGCCTCACGCCTTC	13535 - 13559
n36	GCTTGCCAAGTGAGCGCGCAAACA	13849 - 13871
n35f	CTGAGGATACGACGGTCTGGGTTC	14211 - 14234
n27f	AGATGCCAGCTGAGAACACTGACG	14687 - 14710
n11	CGCACCATGTTATCATCTCGTTAG	15303 - 15326
n1	TGCCAGGTTAAATTGATG	15653 - 15670
n12	CCTCGCTCGGTTTGTCAGGACTTT	16015 - 16038
n16	GTCTATGTTTGGGCTTTGGCT	16291 - 16311
n15f	ACTGTGTGCAGTTACCGTCATTGT	16718 - 16741
n7	CTAATCATGGTCTTGTAGGAATG	17004 - 17026
Ge22f1	GGAGTGAGCTGGAGCATTTAAGCC	18713 - 18736
Gi28f1	GGCAAAGCGGCACAGGCAGCGGGC	21333 - 21356
gi1r4	AGACAAGCATGAGATGGAGGTGCG	498 - 475
gi1r3	TATCGCATCCCGCTGGATGTTCGG	760 - 737
gi1r	ACGCTCCGCTTCGACTGTCCTCCG	1183 - 1160
gi1r2	CGCGCACGTACACGGACGCGTATG	1249 - 1226
5RACE3	TAGGCACACGGAAGGCTCGGACTC	1612 - 1589
gi3ar1	CAGCTGGAAGTGTCAGCAACTGAG	2666 - 2643
gi3ar2	GGAATGCGTGGAGGGTATGCAGGG	3053 - 3030
g4u	CAGAGATGTGTCGATTCCTACACC	3481 - 3458
n42rga	CAGTGACTCTCCGACCTGTGAGC	3900 - 3978
n48r	TGCCCTACAAGAGGTACTGGAACAG	4308 - 4284
n55r	CAGACACAGGAACTACCGACTCAC	4733 - 4710
n65r	GTAATGTGGGAACACACAAAGGCG	5181 - 5158
n66r	GACAGCCAATGCTAAGAGCACAGC	5590 - 5567
g3u	TTGAGGTCCTGTGCCACACTGACGG	5987 - 5963
gi4rb	GAGTAATCAATCCCCATGGAAGC	6212 - 6190
gi4ra	CCCAACTAATGCCGTTCCATTTCC	6318 - 6295
e4r	GTCTCCCAGTTCTTGATCTTCAG	6709 - 6686
n62rga	CTGAGTCCTCTCCATCATTCTTCC	6841 - 6818
n62	AAAGTCTCGCTGCTCTGTGAAGAC	6922 - 6899
n49	TCAACCAGGAGACTGGAAGATATG	7019 - 6986
n56	TGTGCACACTTATGTGCTTCCAGT	7032 - 7018
n41	CTCAGCGAATGACCCTGTGAGTTG	7307 - 7285
n9	ACAAAGCCGAGGCAGAAAGACACG	8538 - 8515
n17	GAGGAAATTTGACAACGAGCCTGCC	8767 - 8744
n20r	ACAAATGTGACATACCTGCGCTCC	9346 - 9323
n26r	GTGCAGCTGAGCGTAAGATGAGGGTC	9716 - 9691
n28r	GCCAGTTTGGTTGGAGCGGTTAGC	9987 - 9964
gi7r	TCATCTACTGCGGTGATTTTGGCG	10934 - 10911
n37	GTACCCACAGCAAGGCTGAGGCTC	11249 - 11226

Table A.1. continued.

Primer	Sequence 5' → 3'	Position
n29	CTAATGGCGACAGCATGTTAAGGG	11634 - 11611
n22	GGAGTCATGCTTCTTGGCAGAAAG	11914 - 11937
828nNOS	CGGAGCAGTTCAAATGCCTTCTTC	12190 - 12167
n24	CCTTCAGGGGTACGCGCAGTACATT	12451 - 12433
n25r	GTAGTCCAAACGGGCTGGACGTTAC	13135 - 13111
e12r	CTGGTACTCATAGGAAGGCGTGAG	13571 - 13548
n36r	GAACCCAGACCGTCGTATCCTCAG	14234 - 14219
n35	GCTTTGACTCTCTTGGCCATTGCG	14458 - 14435
n27	CACCGCGCCGCTCAACGCTTCAGT	14824 - 14801
n10	CCACACACTTGCATGTACGATTCTC	15090 - 15066
n2	ATGGTGTTATTATCACTC	15451 - 15434
n13	GGGACGGAACAATCACTGATGTAG	15838 - 15815
n15	GTCCGAGCTAATGAAACC	16553 - 16536
n6	TCCTCATGATTGTTGCTACTGGGCG	16848 - 16824
n3	TCTCTTTGGAGGGGACTG	17213 - 17196
Ge22r1	CCCTCTCGCATTAGCATAAGCTGG	19070 - 19048
Gi28r1	CGAGCGTGCGTGAAGTGTCTAATC	21919 - 21896

Table A.1. continued.

```
GTTGCTCTGTATTTATAAAGCGCCACGGTGCGCGGAACGGCACAAGCGTGCCTCGTCTCT
AGAGAGGCAGAGCGCATCAGTTTCCACTTGGAGACGGGCTTTTGGGGAGAAGATCGACCT
GTGCAGCCGCCAAAGGAGAGTCGCGGCAGAGTAGGACTTTTGGACCAACTTGGGTGCGAG
CTTGGTGTCCCCGgtgagttgtttttggatgttttaacgcactacgattgtttgactcct
caatcggttgctgttgaaaattcctcagatggattgcagagattaattaatcatcattaa
tcactgattaatcccgatgaattagtttgaaataatgtgtcggttaaaggacaaaagagg
ggttgaacatttattataatagcaattatatgaataataatcatattttatgcgcaaaat
aaacttgttgcttttgtaaatcggcctacaatatgcaccaatttagttcttagtcgcacc
tccatctcatgcttgtcttgttgacctgtctcgcattaaagagatgctgagcgcggaaga
cgcacctgcttctggttcttattgtgccttttttttggtgtttttttttgtgtttttggtg
cattttgtgtgcagatgcaatttttaaagttgctctttgatttcagattttatttccttt
ctgtgctgtgtccggaatgtcttttgtttggtttttaatgtccctttgggggaggttgtg
agaccgagtgttgaagccgaacatccagcgggatgcgataaacaggtgacgcttctcact
gctccagaactttgggccatttaaagcatgtgccaccgggtctgccccttccatcacac
ttaaccgtatcctgaagctgaaagcggcgtgcggtagatatttatgaatattacaagcgg
ctgcctttcaagaaagtttccaatttctatattctgagcacattctcatcaatattcaca
gcggttattggtaatatcgcatcggcggaggattaaagggccgcctgctgaaatgattcc
catgctccaggatggtttggtctgtatttgtgaaaccccccccgcgctcccacttcaca
tcccccggcgccttcacccaaaccggttcttcggataatggaatatacatttaatttttc
cttatatcaagagcaatatcggaggacagtcgaagcggagcgtgtgtgcgtgtgcgtgtg
cgtgggaggggggcacgaaaagtggcatacgcgtccgtgtacgtgcgcgctataacttgt
catctcttcatttaatcattttatgtattcatttccttttttctgttgtcttttgctta
taaccacaataaaaaaaacactgacacattgtccaacaagaattaattacctgcataata
tttcatttaaagacaccttgacccccccccccccccttgacagactgtaaacacatcat
gcagtgaagattatttactctcaaattaatagaactcccatttcccctcatctgcctccc
tctatttcttttcagAAACAGCCCTGGGGAGACATTATACTGCTTTCAATAGTCACTAAC

ACACACTCACAACGCAGGCACCATGCAAGAGTCCGAGCCTTCCGTGTGCCTACTGCAGCC
                      M  Q  E  S  E  P  S  V  C  L  L  Q  P

CAACATCATCTCTGTCCGCCTTTTTAAGAGAAAAGTTGGTGGTCTTGGTTTTTTGGTAAA
 N  I  I  S  V  R  L  F  K  R  K  V  G  G  L  G  F  L  V  K

ACAAAGGGTGTCCAAGCCCCCTGTCATTGTGTCTGACATCATCCGCGGCGGCGCCGCCGA
 Q  R  V  S  K  P  P  V  I  V  S  D  I  I  R  G  G  A  A  E

GGAGTGCGGTCTGGTGCAAGTGGGCGACATCGTGTTAGCGGTCAACAACAAGTCCCTGGT
 E  C  G  L  V  Q  V  G  D  I  V  L  A  V  N  N  K  S  L  V

GGATCTGTCCTACGAAAGGGCCCTGGAGATGTTGAAGAATGTGCTGCCAGAGAGCCACGC
 D  L  S  Y  E  R  A  L  E  M  L  K  N  V  L  P  E  S  H  A

TGTGCTGATTCTCCGTGGACCAGAGGGTTTCACCACGCACCTGGAAACAACCATATCCGG
 V  L  I  L  R  G  P  E  G  F  T  T  H  L  E  T  T  I  S  G

AGATGGCCGCCAACGGACAGTTCGGGTCACGCGTCCCATCTTCCCGGCCTCAAAGTCTTA
 D  G  R  Q  R  T  V  R  V  T  R  P  I  F  P  A  S  K  S  Y
```

Figure A.1. Complete nucleotide sequence and predicted amino acid sequence of the *Fugu* NOS1 gene. Lower case letters represent introns, upper case letters represent exons. Exon-intron boundaries are underlined. The ATG start and TAG stop codons are highlighted in bold. Amino acids are shown in single-letter symbols.

```
CGAGAATTGCTCCCCGCTCGGTCCATTCGGGCCAGGGCAGCAGGTCAACAAGGAGTCCCA
 E  N  C  S  P  L  G  P  F  G  P  G  Q  Q  V  N  K  E  S  Q

GCTCAGGGCCATTGAGAACCTGTCCTCTCCACTGCAAAAAGGAAGCGTGCAGGCTCAGGA
 L  R  A  I  E  N  L  S  S  P  L  Q  K  G  S  V  Q  A  Q  D

CCCCCTGCTGTTGAGGGACGGGGGCCGGGGACTGTGTAACGGGCTGGAGGACAACAATGA
 P  L  L  L  R  D  G  G  R  G  L  C  N  G  L  E  D  N  N  E

GTTGATGAAGGAGATTGAACCAGTGCTGCGCCTCGTCAAAAACAGCAAGAAGGAGATCAA
 L  M  K  E  I  E  P  V  L  R  L  V  K  N  S  K  K  E  I  N

TGGAGAGGGCCAGAGGCACGTAGGGAGAAGAGATGCTGAGATTCAAGTGACCTGgtaatt
 G  E  G  Q  R  H  V  G  R  R  D  A  E  I  Q  V  T  W

tgtgtcgtgatattgatgtcgtcgttgtagtgttctttgtgattgtttccctgctcggtg
tggttggctatagagccccgcgagaattaaggccgagatttgggtcaagagcccttcaac
gaggtgattattagctgacaccagagtctataacttgtcttctgattgaaagcgtgtcct
tttatttgctttctcaccttgaaatcctcgaactacctgacctaaccgtgttagttatca
aagacacggtctaaaattatagcgggcccctccaagaacacttaatatctgcatactgcg
agataaaaaacgaaaacacaaatcagagaagcaaaagtgaagcatttaatgttgaaatta
acctcagttgctgacacttccagctgtgggtggggagggaaagcaagactcgaagggttt
aaaggcttcccatgaaaaagaacgtattaactggcttctttgtattgaggagtgacactt
ttgacacagcattccaagtgttgaggaagcaaaagggtttgatggagactcaaaagagtt
ggctttacaggtttttctgtccttttgactgttcactctgtggtatactttcatcaggc
ttcacagtccgagcctgtattccattccaaactctggatgaaccatatcggcatcccggt
tatatgcttttgtactcccacagccaattgaaatcagattgaggccacagtctcgttgac
tcagtctcaacacttcattcctgaccaatccctgcataccctccacgcattcctgaatac
atttcgattgattttgttgatatgagagtctcaaagcattagcaatagctcatttggagt
acgtggatgaggatgaaaacctgctaacacatgtcaacctttgatcaaggtgcttttgg
gcaacatcatgtcacaatttttcgcatgtcatgcaaacaacatgaaagagagctgccaga
gagccacagcatgaggttgcttaagccgcacgacagcgctagctctaatgacatgtaaca
tactgtatgtcactttgctagaagatgccgtacactgggactcacgtgcgcctgcaaagt
ttggcacctggcgctaacatttagaacgcgagttgctttgcatcgggtgttagcgggtac
aaacctaacctacaactgtttgactcccagGGGCGCTGGTGTAGGAATCGACACATCTCT
                              G  A  G  V  G  I  D  T  S  L

GCAGTTGGACTCCTGTAAGAACAAAATGCCGGAAAAAGAGCCCGGTGTGCCGCAGAACGC
 Q  L  D  S  C  K  N  K  M  P  E  K  E  P  G  V  P  Q  N  A

CGACAATGACAAGgtaaaaaactgtctgcctccattaaacatgattgctttaattcccat
 D  N  D  K

cattcccacatgtctctgttttcatgtctttatctaaatcaagctgtgtttttgtccaac
acgttggaatccacgactgtctctgcccagctgtgcaataatccaaacgtccaacgcgtt
ccgtgcctctagaaacccagtgcctacatctgtgttggttctttgcagagagtagcaaag
cttttccctaagatcccatcacatagggcattagctctttccggggaaagctggcgcagc
attaaatccagcttatgcctctcccacagaagggtccgctcacaggtcggagagtcactg
acaggggatagtgtcgtaatctgcacatgatctgcattatccacccacttgtgacgtcgc
ttgagcccgtctcaatgggaccggaagaggtgaaattactccagggtgcataatgaatgt
gtgcagctgctgatccccccccccatcccctccaacaaatgtgacccctgctgaaaagatt
cagcccctcagacctgcagaatacagagttaatctgaatctttgggagattatttgttgt
gtgggcggagggctggacctggaagaatgcacacttaagagtggatgaagaggacgtgtg
tgtgtgtttgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgccacacactcagga
aatatacttgtgcatgtgtgagactgttccagtacctcttgtagggcaccggttgcacaa
ctggtgtctgcgaaacccagaaaccaagatttttttttttctttttaactctctgggatt
aattacagagattttaagcgacgtgtcagtagcctttagacaaactctcagttgggcgag
ggatcgctcatgagactgcagtgtgagatcctgcctgtcgtttcatgcttaataatgctg
ctgccaacagctgaatttgctatcatccatcaggaatcatccatcatcaggaatcagttt
tcacgtcagatgattcatcttttcttttgcattagttccaatatctttttttcacccctgc
```

Figure A.1. continued.

ttttgatcctttgatctggtgtcttgatgattctttatacgtgttgaaatatctgatcc
tgttcatttctatgcagttacgttgttttgtgagtcggtagttcctgtgtctgcctagcc
tcacctgtgcctcggcgcctcctcccacctatatgcgctttctttagctggtgaatcttg
ttgcttcatgtgtcaaaatggggtgataatgtaaccagtacggcatcccttaaagaatgc
tattattctgttattcctgagatcagttccagtcattatccagttcctctgatactggaa
accaggacagatcattgatacgtgatggtatttaaacatgtcagctgataagtcttctgg
caacattcacaacacaacagcaattataagaaggctttgctgttactaaatgtgctttgt
tacttagcaaccacagacaaaggatcagacatatctgttaaacttggttcattcagtgtt
gttgagttgtattattattattgttgttatgtttattgttattattattatttatgcgc
ctttgtgtgttcccacattacttaaacattacagtgataaatgaagatggacagtgcttt
atgtgacacttttactttacttttcctgtgttctgttttccaggatgctgtataaacatg
ttagaaacgtactagcagtgatctgtatatatcagttaaaacaagccgaaatgctaatgt
tatcgtccatcatatgagaatggacttctgaaaaaacatgagcaaaaatgctcatgtttt
cttccagcccattcaagggtcggataagagctgaagtgtttacatgcgctgcagaggtgg
aacaatctggtttccctttttcagtatatttattagatttctttcagactatgttccta
gtaagattcaaccactctgctatttcagtccagcaggatttatcatgctgtgctcttagc
attggctgtccaaatgtgggagctaatagaacctgaaaagagagctacacgctgcttctg
ctgctgtcctttgtttatggtcatatgaaccttctcagtgatgcagaatttacctttcag
gagtggcacctaggggttcttcgctgaaagtccaaattagaatctcttttttaaccgaaat
tctattgaagaccttcttccctcctgcatttccactctgactttaaaaataagagataaa
atcacccgtttaggtgtttttgatgttggagaatcattctctgaagtgttttttctcagc
ttcatttcaaatcctaagattgttgcgttttcttttatgtgtttcctctttctttcctca
attttggtggaaacaaattgtcccgtcagtgtggcacaggacctcaaactttctttttgc
ttctctccttgtgcaaaatcaaacgtaatttagtccacctcactcctctccaactgtttg
aagacagcaaatcacagtgaagtttggtattcccacttttcatcggccccacgagtcgag
ggattgtgtcgctgttgttcagaggcttttctgctgtgcgttacagatgctgggggatat
tatgaaagagcttccatggggattgattactctgatcttaggcaatcaaactgtgctttt
ctctagagccagtgctcaggttaaattgcaggcattcatcagtcaataggagttggaaat
ggaacggcattagttggggagaatctaggtggttgctgtctgtgagagagactcaagcat
aatagagaaaagcaactctcaaacgaaagtccaccaagctgtgacggctggcattgcttt
gtataactgctcaaattgactttaccagtcggtaaccaatcagggcatttatccagcgcc
gcaccgtgaggagatcgaatgtctcttcagctttcctcggcatcactgcgagtcaaacag
ctggtattaaatgcatcccagaggttgcacacggttgaaaagcacgcaataatgactcat
ttgtgcccttgcag**CCTCCTGCCGAGGCCAGGACCTCACCTACTAAATCCCTGCAGAACG**
P P A E A R T S P T K S L Q N G

GAAGCCCCTCCAAATGCCCTCGCTTCCTGAAGATCAAGAACTGGGAGACCGGCGCCATCC
S P S K C P R F L K I K N W E T G A I Q

AAAATGACACATTACACAACAGCTCCACCAAGgtgagaggcatggaaatgtgtttcatta
N D T L H N S S T K

atggttgtaataagctgaaggaaccatctggtagaggggaagaatgatggagaggactca
gagctgagatttgcaaaggtgaatgcgggagacgactgtcggctgccaacaggctcacgt
cttcacagagcagcgagactttggtttgagggcgccaacgtgcatctttccatctctggg
tcaatgctttataagccctgatgtgcatatcttccagtctcctggttgactggaagcaca
taagtgtgcacactcatcaacgcatcatttgcaagtcaagtagattgaagaataatagca
catcgcagcacgtctcctctccgcttcacactctgccactcgtgcattagacggctttcg
cttcccaaattttggcgattaagctcattacaatcaaaaatgaaaaatggggtgggctgg
gggggtgttgggggtgttaagcggttagcgcctgctgaacaatgaaccactgggagagaa
aatacggacttaccttatggggacaactcacagggtcattcgctgagtcacattattacg
cagcattttcatgaaagtgagaaatccgacgtgacgtttggcagatatttttatactgcc
tcaatttatgccaggaggggctcaagaaagcagaaaaagctattcatcaatccggttatg
ctcttattgctcttgaaaataacaactgctgtttgctaattgtcactttatgccctccgc
ttccacgtcagtgcaaaacaacacaagaaaacaacattatttacttttatatatgtgggg
tccatgttgctaagcatgtcctcagctgtcacccattagtcctgatctgaccgacattgc
taacgctaagtgaaacaacaaaatcattacaagacaagcacgcaaagtttattcttcagt
ctataatcagtttcatgttaagtcacatctttccttattcttatcgcatgttatttatat
atcacatgtgctgtgaaaactactgcagttatagcctataaatgggcatatcatgtttct
catttatctaattttatcattgtcgtgtctggagttactgtccaaaaccgttgttaattg

Figure A.1. continued.

```
ttagcagatttaaccggctattgatcggccttaaacgttcatgaattcagaagcttctac
aatgacaaactcagctcacctttgcgggaccgtggtcatactaatcacgcaccccagcca
gcctttgtcactttaacagacatttcagcctcagccaatagcagcctataaagataatac
gagatgacccaaagctttgtttaaagtgaaaatggctgttgaaaaaaatcaacagagaga
attggttccttatcccaaatcatctccccctcatctcgggcttggtggtgaaatttcaat
caggtttcctcttgtgccacctgttcccgtcgagcacaaggtgaaccgcgttagcattca
ccccccgaaatgccgcagccatttgtaaacagcagccaaatgactttggtgacttatcaa
atgggctctgcatcccaatgaaaggcctgttgtggcatcttcagatttattatatatcac
agatgctgccatattggagcctgatttgtctgtttcaaaaaggacagctggctggcatga
atgcccggagcagtatgatgagcggacatgctgcagcagaagtgattcctttgttttatt
tttaacattccctgtgctgctcatcacgtctttttcacttatcttacgtcggccgtgtc
tttctgcctcggctttgttccagcctcgtgttgttatacacccttatccttttttaaa
ccccttctggctgcgatgggtatattttaataaattaggtctgagtcagtggtgagaaga
ggaggaagtttatttggatggtcctcatctagagagagtgagagaggaagagagagaga
gagcaccaatgtgtgcgtctgcattaaaagagcgactttaacggcaggctcgttgtcaaa
tttcctcgctgacccgttccatctatcgtttctgctgcccgctcactcgtttgagctggt
gaaaactttaagcaaaccaaagtgacagccgacttcatctctcatcatgtgtcactggc
cacacgaagctgccttcagccacgaacagctccagtttcatattttcatatttgctttt
cgctcttttaggttctgtttaatgaaggaagagagtcggtgactcacacagtgtttgcat
gcacatcggtcctgtgtgagccttgaatgtgatgatgtgatgggcagaatggtttgttta
cagACGCCAATGTGCCCCGAGAACGTGTGCTACGGCTCCCTTATGATGCCCAACCTGCAC
   T  P  M  C  P  E  N  V  C  Y  G  S  L  M  M  P  N  L  H

GCCCGCAAACCGGAGGAGGTCAGGAGCAAAGAGGAGCTTCTGAAACTGGCCACCGACTTC
A  R  K  P  E  E  V  R  S  K  E  E  L  L  K  L  A  T  D  F

ATTGACCAGTATTACACCTCCATCAAAAGgtaagcccagtccgttccgtccgttccgtta
I  D  Q  Y  Y  T  S  I  K  R

tcaaatgtaattctgagcgcttaaagtcaaagctgggtgaggagcagagaaggggattct
ttttaacacattggtggcgtttggagcgcaggtatgtcacatttgtgagaaataaagcac
ctaacgtgtaaacctttccacatcacagtgaaggcgggggcaaggatttggaaaatgggt
aagacagtgaccagatgaaggtttgagtgactggaaaacatgtgcattcaaaacaagact
tttcttgagctaatcagacccactaaagtctgaggtaacggttttctaaactcatggatt
tcaggaggtcattgaaaatgattttccttgatccggaaggtttgaaacaacctttctttt
ctccggatacacctgaagtttagaacttctgtgcaccgctttcctcaaactttcatattg
attcagcttttctttaattaaaagccttttgaccctcatcttacgctcagctgcacgtgg
ttcacaatcaatatcagattttggattttctgcgtcatgttagcagctctgggtgtaaga
tgatgaccccacccccccgcacacggatctcccggtgcatgggggggcagcatcggttta
ctgtcgctctactttctgtgaaacatcgtcgcatcacaggacagtctgtgagtcctcgct
gagaaaaactggttaccacgatggaacaacagcgagtggccgagcccttcttgaactctt
tatgctaaccgctccaaccaaactggctagcttgccgagcctaaggcagccattctaact
acagttggcttttctggtttggcttgaagctctgctgtggagtctaattaaggcctgtct
gtaatgtgactgcagGTACGGCTCCAAGGCCCATACGGACAGACTGGAGGAGGTGACCAA
               Y  G  S  K  A  H  T  D  R  L  E  E  V  T  K

GGAGATTGAAGCGACTGGAACCTACCAGCTGAAAGACACAGAACTGATTTATGGAGCCAA
 E  I  E  A  T  G  T  Y  Q  L  K  D  T  E  L  I  Y  G  A  K

ACATGCCTGGAGGAATGCTGCCCGTTGTGTGGGAAGAATCCAGTGGTCTAAACTACAGgt
 H  A  W  R  N  A  A  R  C  V  G  R  I  Q  W  S  K  L  Q

gaaagacgcctataatcttgttgcacggtgctctaattatatcttttatatcgagtctgg

tctttttacaaacaatggaaaaactcttgcgctcacatttcagGTTTTTGATGCCAGAGA
                                          V  F  D  A  R  D

CTGCACAACAGCTCATGGAATGTACAACTACATCTGTAACCACATCAAGTACGCTACCAA
 C  T  T  A  H  G  M  Y  N  Y  I  C  N  H  I  K  Y  A  T  N
```

Figure A.1. continued.

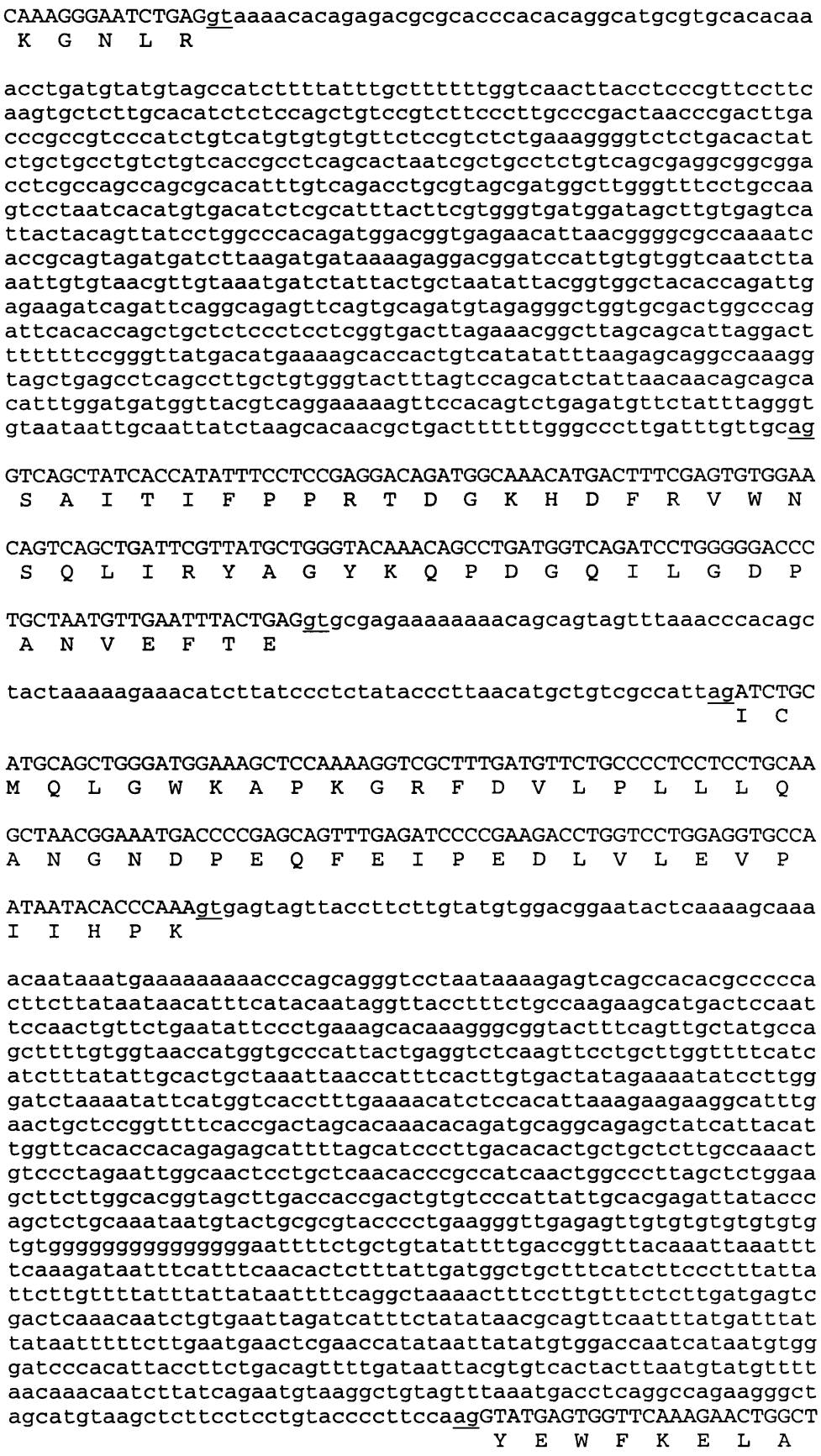


Figure A.1. continued.

```
CTCAAGTGGTACGCCCTCCCTGCTGTCTCAAACATGATGCTGGAGATCGGAGGCCTGGAG
L  K  W  Y  A  L  P  A  V  S  N  M  M  L  E  I  G  G  L  E

TTCACTGCCTGTCCCTTCAGTGGCTGGTACATGGGCACAGAGATCGGCGTGAGGGACTTC
F  T  A  C  P  F  S  G  W  Y  M  G  T  E  I  G  V  R  D  F

TGCGACACGTCCCGCTACAACATGCTGGAGgtaacgtccagcccgtttggactacagaga
C  D  T  S  R  Y  N  M  L  E

gatgcagactgagactagcaggtagttaaatgcaatatttccctgtttactgcttaggag

GTTGCAAACAAGATGGGCTTGGACACCAGAAAGACCTCCTCCCTCTGGAAAGATCAGGCT
V  A  N  K  M  G  L  D  T  R  K  T  S  S  L  W  K  D  Q  A

TTGGTGGAGGTCAATATCGCCGTCCTTCACAGCTTCCAGgtaaacattcccaccttgagc
L  V  E  V  N  I  A  V  L  H  S  F  Q

tctcccgtcgaactccctcgagaagaaaacacatcagttctgtgcttcattctcagTCAT
                                                          S  C

GCAAAGTGACCATAGTGGACCATCACTCGGCGACCGAGTCCTTCATGAAGCACATGGAGA
   K  V  T  I  V  D  H  H  S  A  T  E  S  F  M  K  H  M  E  N

ACGAGTACCGGGTGCGAGGCGGCTGCCCCGGAGACTGGGTGTGGATTGTGCCTCCCATGT
   E  Y  R  V  R  G  G  C  P  G  D  W  V  W  I  V  P  P  M  S

CCGGAAGTATCACGCCGGTTTTCCACCAAGAAATGCTCAATTACCGCCTCACGCCTTCCT
   G  S  I  T  P  V  F  H  Q  E  M  L  N  Y  R  L  T  P  S  Y

ATGAGTACCAGgtgaccgccgccctcagcacagctgcacatacacactcgtgcttgataA
   E  Y  Q

accaggcgcacagagagcaccatgaaagtcatttcctgttagaaacaaagcaaactgagc
accttagagaaggaaagggatgcatggaattgtgaattacagatacgagactgcatctaa
actcttaatttcttctcccatatgcaaaagCTTGATCCCTGGCATACCCATGTGTGGAAA
                               L  D  P  W  H  T  H  V  W  K

GGAGTCAACGGGACGCCCACAAAGAAACGAGCCATTGGATTCAAAAAGCTTGCCAAgtga
G  V  N  G  T  P  T  K  K  R  A  I  G  F  K  K  L  A  K

gcgcgcaaacagaaatattagacaaggcggacattcggtgttcccgaagtggctatttgt
ttttctcatcttcctctttctgcatccgtaatctcattgttacgttaaatccctaatccc
cggacttgatatgtaatatagagccggatgtgataatccagggccaaacgaagtcatcga
cttcttggaaaaaaaaaaaaaaaagtcagacacatgggatgcaagctcgagcactcagca
atgtgatttattcaatcagcccctgggggagtaatcaaagcttttccagggttttatcag
actatgcgtgcgctcatgcatgcatttgtcctgctattaatatagtttccctgaggatac
gacggtctgggttctaatcgtcatctatctgctgctttgacataaaaataacaaaaatga
tccccttgaaatctgttatcagctctggttttatgctctatgtaatttaatgaacctctt
tgaatttgaatgcccattttagctctaactgttcattttcttaacccccccctccttgca
gAGCGGTGAAGTTTTCAGCCAAACTCATGGGTCACGCAATGGCCAAGAGAGTCAAAGCCA
   A  V  K  F  S  A  K  L  M  G  H  A  M  A  K  R  V  K  A  T

CCATATTGTTCGCCACCGAGACGGGAAAATCGCAGGATTACGCCAAAACTCTCTGCGAAA
   I  L  F  A  T  E  T  G  K  S  Q  D  Y  A  K  T  L  C  E  I

TCTTCAAGCACGCATTCGACCCAAAGgtaccaggttgcatcactttgataagcagtcggg
   F  K  H  A  F  D  P  K

cacagaaaggacgaaaaatttgggggggggggggttcaaaaggatcacgctgcggccgg
tcttcggccctccgctctccctcttctctggctgaaaaagacgagcagatgccagctgag
```

Figure A.1. continued.

```
aacactgacgagagggtagaagaacaaaatgtgtcatttgtcacatttgtcttcgtctct
catgcagcgaagcgccgccggtggaaggcaaaatttaaccactgaagcgttgagcggcgc
ggtggtttaagaccagatccggaggcacctctggctgtgcaaatcagtttagcggaagtg
tgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgcgtgagtgtgtgtgtg
tgttgggggtagagagagaagaagatgaacgtcagtcaaccactttgtgcatttcctcca
tgtcataactcattccccgtcacccccccgccccccccttccatgactatgatcaatctcta
aacaagagaatcgtacatgcaagtgtgtggatgtgtgagtgggtgttgaggcacagatga
atgcccgtggctctttctttgctttttttcccccccgcttttctttttcttgtcgtggg
gctaaagcctgcgtctttttcacgcctttgctcatgcgcgtgcgcgcgagcaatgtcaga
tgcatgccacggtagagaaataccacgctgaaaggtgtgattcacgggcggttgttgtgc
gccgcaccatgttatcatctcgttagaggcgaggagagggtggtcattacaatttcttcc
tgctagctttctgtgtgttcaacctttcatttacagtctcaagccatctgtgaagagtgt
gtgtctagaagaagagtgataataacaccattacacaagtggaggcttacgactcccccct
ccgtgagccgtctaattctttttttttttttgtcgtctctgtagGTCATGTCTATGGACG
                                              V  M  S  M  D  D

ATTATGATGTGGTGGATCTGGAGCATGAGACGCTGGTGTTAGTGGTGACCAGCACGTTTG
  Y  D  V  V  D  L  E  H  E  T  L  V  L  V  V  T  S  T  F  G

GCAACGGCGACCCCCCTGAGAACGGAGAGgtatgcgcgcgaggaagcagcgttgccaggt
  N  G  D  P  P  E  N  G  E

TAaattgatgttaggcacagttcactgagagtcatgcccgtctctgtagAAATTTGGAGC
                                                    K  F  G  A

CGCCTTAATGGAGATGCGGGAGCCGACGTCCAACACAGAAGACAGAAAgtgagtccagtc
 A  L  M  E  M  R  E  P  T  S  N  T  E  D  R  K

ttggttcccccctctccctcctccctccagtgtaactacatcagtgattgttccgtccctt

cagGAGCTACAAGGTCCGTTTCAACAGCGTGTCCTCCCACTCCGACACTCGCAAGTCCTC
    S  Y  K  V  R  F  N  S  V  S  S  H  S  D  T  R  K  S  S

AAGTGACGAACCGGACGCCAAGATTCACTTTGAAAGCACCGGACCTCTGGCCAATGTCAG
 S  D  E  P  D  A  K  I  H  F  E  S  T  G  P  L  A  N  V  R

gttaataaactctcattctctttttttgaggggggaatagattgaattttaaacctcgc
tcggtttgtcaggactttcttgcttctgaagcaatttggggctgacggatgttatcagac
agtttgctgacctgttaaagcaccagaggagaccgaacttgatgtgggctgactgattca
cgaaggaaggatgacttatccgctccagcaggagcgtttgcattcatctaaaaccagacg
gtgcatctattcagaaacacgtctgctgaatttccacttttcaaggcagaaggaggagat
agttgctgacgtggggaaggccgcctccatgtctatgtttgggctttggctgcttctctc
gctccatctcttccctcactcgtgtgtctctgtttagGTTCTCTGTGTTTGGCCTTGGAT
                                              F  S  V  F  G  L  G  S

CCAGAGCCTACCCACACTTCTGCGCCTTTGCCCACGCCGTGGACACGCTGTTTGAAGAGC
  R  A  Y  P  H  F  C  A  F  A  H  A  V  D  T  L  F  E  E  L

TCGGGGGGGAGCGCATCCTACGCATGGGAGAAGGGGATGAGCTGTGTGGACAAGAGGAGG
  G  G  E  R  I  L  R  M  G  E  G  D  E  L  C  G  Q  E  E  A

CTTTCAGAACCTGGGCGAAAAAGGTTTTTAAGgtcggtttcattagctcggacactttgt
  F  R  T  W  A  K  K  V  F  K

atgtggtacgtggatgttggacgtcaaaggtttactgttgcatagttttgataatcccca
aaggtgatgtgtaaaatcaacactactactatttttgtgattttgtaggtttttgaatc
actataattttagcctgcgagtaatatttttgagagaactgtgtgcagttaccgtcattg
tatttttgatagcagcaataggcaccaaacattgctcagtgacacctgagacttatgatc
cgtgttaccagagcccaaacactcgcccagtagcaacaatcatgaggaagaacgtccata
tttgaacagttttaatggggagggtctggttcagaaccgcttcagaggtgttgacctcat
```

Figure A.1. continued.

```
caagagactccatgtgatcactttagactttagcctctggtaaaaggggaaaaaccttt
tctaatcagacctaatagaaactctaatcatggtcttgtaggaatgagaaagaaagtttc
agatataaatagaaaaaactctcaaaatgaaatagttgctctgatgtatgtttgagatct
agtaatttagatttgctcagatgtattccacctggtgagcaacggatcgggacctttttg
gactcaaaactattggattaatattgctttcctatcagtcccctccaaagagagacgcgg
ggagactctgctggcccctgttgagtcgatgaagggttgtgttctttcctgtgacgcagG
                                                            A

CCGCCTGCGACGTGTTCTGTGTCGGCGATGACGTGAACATCGAGAAGGCCAACAACTCCC
  A  C  D  V  F  C  V  G  D  D  V  N  I  E  K  A  N  N  S  L
TGATCAGCAACGACCGCAGCTGGAAGAAAAACAAGTTCCGTCTGACATACACGGCCGAGG
  I  S  N  D  R  S  W  K  K  N  K  F  R  L  T  Y  T  A  E  A

CGCCGAGCCTCACAAAAGgtaaatgatcatgggggggggcagatgaaataagttttaca
  P  S  L  T  K  A

ggaggggtcgataaaggcaaatatgcactgcagtgtgaaatgaccacaatgcatctgga
tgagcgatggtaatgttattgattaaagccattcactctttcattacactgtttggtgcg
cagcacattggctccactctggcatgaagaatagtaaagcaattgagcttaaagaacccc
tctacaaaaggggtcgatcgacccaagaacaaggtctgaaactgactgatggctcctgtc
cctgatttgcattatgtaattccaagtctttcatagcgctctggggctttaaatgagatt
ggaacatcaggaacaacagctgctagaatacagacgagcgtcataaatcagaaaatattg
tatgttgaactttaattgtgaggagagacggttctttttagcctctaaaagcatccgtgt
gatgttctgccaacagCTTTGTACGGCGTGCACAAGAAGAAGGTCCATGCAGCAAAGATG
                  L  Y  G  V  H  K  K  K  V  H  A  A  K  M

CTCGATTCTCAGAATTTACAGAGTCCGAAATCCAAgtaagttggagctgttgcgatctgc
L  D  S  Q  N  L  Q  S  P  K  S  K

tctgttcgatgtcaacatggtgcgaacttaacggcaacgggcacaaaatggcgtcgcttc
tctaaagggttcacttgtttctgcagTCGTTCCACCATTCTCGTACGGCTGGACACAAAT
                          R  S  T  I  L  V  R  L  D  T  N

AACCAAGACAGCCTGAAATACAAGCCAGGAGACCATCTGGGCATCTTCCCTGGCAACCAC
N  Q  D  S  L  K  Y  K  P  G  D  H  L  G  I  F  P  G  N  H

GAGGACCTGGTGTCGGCTCTCATAGATAAGCTGGAGGATGCGCCGCCTGTCAATCAGATT
E  D  L  V  S  A  L  I  D  K  L  E  D  A  P  P  V  N  Q  I

GTAAAAGTGGAGTTCTTAGAGGAGAGGAACACTGCCCTGGgtgggtgacacaatggcacc
V  K  V  E  F  L  E  E  R  N  T  A  L  G

gcatggaataagtgtgcatcctgagaatttatggtgtaaaatctgcgtctcactactcca
atccatgcctgtcgtgtgcaaccttcttcccctcagGTGTGATAAGTAACTGGACCAATG
                                    V  I  S  N  W  T  N  E

AGACTCGGGTCCCTCCCTGCACCATCAACCAGGCCTTCCAGTACTTCCTGGACATCACCA
  T  R  V  P  P  C  T  I  N  Q  A  F  Q  Y  F  L  D  I  T  T

CCCCGCCCAGCCCCGTACTGCTGCAGCAGTTCGCTGCTCTGGCCACTAATGAGAAAGAGA
  P  P  S  P  V  L  L  Q  Q  F  A  A  L  A  T  N  E  K  E  K

AACGGAAACTCGAGGTCCTCAGCAAGgcaagtcctcacacattcgctcttcagaaataca
  R  K  L  E  V  L  S  K

ggctaaaagacagtgttcacacaattagcatagccacttctaaaacaacaggccaacagg
ccactctcttttgcctgtaacgtgcgttgcgttgatccaaatatgtacatttggagtgag
ctggagcatttaagccagcaacccccccctgctgtgcctcccacagttcactcttgatttc
tgtcagtgtgtggaagaggctgactgaggccgtgcgctgactcatgcttatatatggaat
cgctgtcaccaaaccaccagacaatgatggaggaaaaaagtagtgcccccgacaacagaa
```

Figure A.1. continued.

```
gcgtttttaccagcgcgctaacgttcaggcctggtcggacaagaacttctacttatccgc
ccagaaacttgttgtcatgtgtaagacggaagcaactttaacaatatgcagaaaaacacc
gtctttgtcaaggctaatagctgtagccagcttatgctaatgcgagaggggtcaggaggc
agttttaagatgaccattactcatcactggttcgagcctgtcatcgctgacacaaacaca
cccagcaaatggggatttttgtgttttcctgccaatggctgctcacgtaattccgtcctt
tccttggcttccgtccacagGGCTTGCAGGAGTATGAGGAGTGGAAGTGGTACAACAACC
                    G  L  Q  E  Y  E  E  W  K  W  Y  N  N  P

CCACCCTTGTGGAGGTACTAGAGGAATTCCCCTCCATCCAGATGCCTTCTACGCTGCTGC
  T  L  V  E  V  L  E  E  F  P  S  I  Q  M  P  S  T  L  L  L

TCTCCCAGCTTCCCCTGCTGCAGCCTCGTTACTACTCCATCAGCTCCTCTCCAGACCTGC
  S  Q  L  P  L  L  Q  P  R  Y  Y  S  I  S  S  S  P  D  L  H

ACCCGGGAGAGATCCACCTCACGGTTGCTGTGGTCTCCTACCGTACCAGAGgtgaggaac
  P  G  E  I  H  L  T  V  A  V  V  S  Y  R  T  R  D

tcctgcgtccatcacacatgtaatgtctaatataatgtgaataggtcttatttcaacatc
caataacccaaataaaaggaaactgcccagtctgtccagcagcatttggggatttggtgt
tagcttcactgctagcgctgaagtacccgtcacatttttccagtatatcctcacagctac
gttcatctataatatgcatttacatgtaggtttgattgtatagttgggctgaattatagc
aatttctgttgagtgagagagcgagagaattggacctttgctctgaaattcaggttcggg
gccatttttttttgccacttaaagagcggaagcgaagagaaacaggtaaaatcttgcttg
gcaaccatccttttcagATGGAGCAGGGTCGATCCACCATGGAGTGTGTTCGTCGTGGCT
                  G  A  G  S  I  H  H  G  V  C  S  S  W  L

CAGCAGGATAGAGAAGGGGGAGATGGTGCCGTGTTTTGTCCGAAGgtaaatatcagtggg
 S  R  I  E  K  G  E  M  V  P  C  F  V  R  R

gagtttttccaggttgtactgacaaacaagggtgggatcgtcccgcatggctgacatttt

actaacgagatttttattgggttttttttgttttgaagTGCTCCATCCTTCCAGCTTCCCA
                                       A  P  S  F  Q  L  P  K

AAAACAACCAAACACCTTGCATCCTGGTGGGTCCGGGAACCGGAATCGCCCCATTCCGGA
  N  N  Q  T  P  C  I  L  V  G  P  G  T  G  I  A  P  F  R  S

GCTTTTGGCAACAGCGACTGTATGACCTTGAACACAACGgtgagatcccctcatcagctg
  F  W  Q  Q  R  L  Y  D  L  E  H  N  G

catttccatcgtatttagagcaaatgtacagtataacaactctttatactcttcctttac

ccctctgtagGCATCGAGTCATGCCCAATGATCCTGGTGTTTGGCTGTCGACAGTCTGAG
          I  E  S  C  P  M  I  L  V  F  G  C  R  Q  S  E

ATTGACCACATCTACAATGAGGAGACCATCCAAGCCAAGAACAAGAACGTGTTCAAGGAG
I  D  H  I  Y  N  E  E  T  I  Q  A  K  N  K  N  V  F  K  E

CTGTACACGGCCTATTCCAGAGAGCCCGGCAAACCAAAGgtactaacagcttttaaatgt
L  Y  T  A  Y  S  R  E  P  G  K  P  K

ccataatttaacacagcatgtacaaaattgacaaatggcacttttttcattttccctgcc
tctcagtggtctcctggatcttctgctctcttcatttccatgtgtgactttgtcatttttc
cctctgtctccttgtcagAAATATGTGCAGGATGCACTGCGTGAGCAGCTGTCGGAGCGG
                  K  Y  V  Q  D  A  L  R  E  Q  L  S  E  R

GTGTACCAGTGCCTGAGGGAGGAGGGAGGACACATCTACGTGTGCGGGGATGTTACGATG
V  Y  Q  C  L  R  E  E  G  G  H  I  Y  V  C  G  D  V  T  M
```

Figure A.1. continued.

```
GCCGGGGATGTTCTCAAGAACGTCCAGCAGATCATCAAGCAAGAGGGCAACATGAGCCTG
A  G  D  V  L  K  N  V  Q  Q  I  I  K  Q  E  G  N  M  S  L

GAGGAAGCCGGCTTGTTCATCAGCAAGCTTCGGgtacgacagatggtagcggtgctgttc
E  E  A  G  L  F  I  S  K  L  R

atcgctgaagcagcagcgccaacacgacagtgtgaatgtatggcaagaaaatatagcgag
tatttcctcagaggaaggatgacaaagcaccgattagcttttggctaatcatgctaatgt
ttacacctgcccttcccttggtcgtagtttagagagcaaccatagaagaagacttgttgt
aaagtaggtgtgggagataatctcggcattagcattgacccttttctcgcctttactacc
taacacagGATGAGAACCGCTACCATGAGGACATCTTTGGAGTCACCCTGCGCACCTACG
       D  E  N  R  Y  H  E  D  I  F  G  V  T  L  R  T  Y  E

AGGTCACCAACAGGATTCGGTCAGAGTCCATTGCCTACATCGAAGAGAATAAAAAGGACT
  V  T  N  R  I  R  S  E  S  I  A  Y  I  E  E  N  K  K  D  S

CCGATGAgtaagtgattcacacctatttgtgtatttgcatacattgggtccaggttaaat
  D  E

ggggttagcgtagtgttagctttccttctggagtttgcatgttgtccctgtccctatgta
ggctttattgagtgcttggtagtatctataaccaaatctccatgcagaatttacacagct
ttataaatgtggtgagatatttgaaacataatggcaaagcggcacaggcagcgggcgttg
tgttgccaaccatagctatagctagctaaagctcactcccgtctagcacatatatatgga
gaattctaccactttatccattttctactggacactgacgctgtggtgactcagttctga
gtgcattaaccatcttccaggcttctttgatgataatattccttttctgcgtgcagcatt
taacacctgaatgcctgaatagtggaggaagctaaagaacgtgcggccatatttgttcat
cttttcccataaggaacctttctgtcacatttcttcaaaatcttgggatttttatccaaa
ccggccagattatcagagcgcctccgtctccgaagagcgtctgtgtaaatccgactgcaa
tctgcaacctgtagttgctggcgacggcgccgtgcgatcgctgccgcgcttcagatgaaa
acagagcgggtaaccatggaaacctgccgacagtggcattctctaagccagcaggattac
agcgaactgaagtggctctgtgtttcctgtttttgtttccactgcacatatttaagatta
gacacttcacgcacgctcgttgtgagagccaaacaagagccctacatcggcgccggggct
catcgagcgccagcaaaatgcatcactgatgacacttcctgtgtgaattgtcacttgaat
tttacatttggccatcaaaccccccccctctgttccccgagtgagtattggctgcgactg
tttccctgcctttatctctctgtcgtctcccctcagGGTGTTCTGCTCATAGGTCTCTGA
                                      V  F  C  S

ATCCGGGCCTTTCTGACCACGACCAGGCTGGACGAGGCGGCGGCGCTGGCTGAAGAATAC
TGT
```

Figure A.1. continued.

```
                                                          GTTGCTCTGTATTTATAAAG
CGCCACGGTGCGCGGAACGGCACAAGCGTGCCTCGTCTCTAGAGAGGCAGAGCGCATCAG
TTTCCACTTGGAGACGGGCTTTTGGGGAGAAGATCGACCTGTGCAGCCGCCAAAGGAGAG
TCGCGGCAGAGTAGGACTTTTGGACCAACTTGGGTGCGAGCTTGGTGTCCCCGAAACAGC
CCTGGGGAGACATTATACTGCTTTCAATAGTCACTAACACACACTCACAACGCAGGCACC

ATGCAAGAGTCCGAGCCTTCCGTGTGCCTACTGCAGCCCAACATCATATCTGTCCGCCTT
M  Q  E  S  E  P  S  V  C  L  L  Q  P  N  I  I  S  V  R  L

TTTAAGAGAAAAGTTGGTGGTCTTGGTTTTTTGGTAAAACAAAGGGTGTCCAAGCCCCCT
F  K  R  K  V  G  G  L  G  F  L  V  K  Q  R  V  S  K  P  P

GTCATTGTGTCAGACATCATCCGCGGCGGCGCAGCCGAGGAGTGCGGTCTGGTGCAAGTG
V  I  V  S  D  I  I  R  G  G  A  A  E  E  C  G  L  V  Q  V

GGCGACATCGTGTTAGCGGTCAACAACAAGTCCCTGGTGGATCTGTCCTACGAAAGGGCC
G  D  I  V  L  A  V  N  N  K  S  L  V  D  L  S  Y  E  R  A

CTGGAGATGTTGAAGAATGTGCTGCCAGAGAGCCACGCTGTGCTGATTCTCCGTGGACCA
L  E  M  L  K  N  V  L  P  E  S  H  A  V  L  I  L  R  G  P

GAGGGTTTCACCACGCACCTGGAAACAACCATATCCGGAGATGGCCGCCAACGGACAGTT
E  G  F  T  T  H  L  E  T  T  I  S  G  D  G  R  Q  R  T  V

CGGGTCACGCGTCCCATCTTCCCGGCCTCAAAGTCTTACGAGAATTGCTCCCCGCTCGGT
R  V  T  R  P  I  F  P  A  S  K  S  Y  E  N  C  S  P  L  G

CCATTCGGGCCAGGGCAGCAGGTCAACAAGGAGTCCCAGCTCAGGGCCATTGAGAACCTG
P  F  G  P  G  Q  Q  V  N  K  E  S  Q  L  R  A  I  E  N  L

TCCTCTCCACTGCAAAAAGGAAGCGTGCAGGTTCAGGACCCCCTGCTGTTGAGAGACGGG
S  S  P  L  Q  K  G  S  V  Q  V  Q  D  P  L  L  L  R  D  G

GGCCGGGGACTGTGTAACGGGCTGGAGGACAACAATGAGTTGATGAAGGAGATTGAACCA
G  R  G  L  C  N  G  L  E  D  N  N  E  L  M  K  E  I  E  P

GTGCTGCGCCTCATCAAAAACAGCAAGAAGGAGATCAATGGAGAGGGCCAGAGGCATGTG
V  L  R  L  I  K  N  S  K  K  E  I  N  G  E  G  Q  R  H  V

GGGAGAAGAGATGCTGAGATTCAAGTGACCTGGGGCGCTGGTGTAGGAATCGACACATCT
G  R  R  D  A  E  I  Q  V  T  W  G  A  G  V  G  I  D  T  S

CTGCAGTTGGACTCCTGTAAGAACAAAATGCCGGAAAAAGAGCCCGGTGTGCCGCAGAAC
L  Q  L  D  S  C  K  N  K  M  P  E  K  E  P  G  V  P  Q  N

GCCGACAATGACAAGCCTCCTGCCGAGGCCAGGACCTCACCTACTAAATCCCTGCAGAAC
A  D  N  D  K  P  P  A  E  A  R  T  S  P  T  K  S  L  Q  N

GGAAGCCCCTCCAAATGCCCTCGCTTCCTGAAGATCAAGAACTGGGAGACCGGCGCCGTC
G  S  P  S  K  C  P  R  F  L  K  I  K  N  W  E  T  G  A  V

CAAAATGACACATTACACAACAGCTCCACCAAGACGCCAATGTGCCCCGAGAACGTGTGC
Q  N  D  T  L  H  N  S  S  T  K  T  P  M  C  P  E  N  V  C
```

Figure A.2. Complete nucleotide and predicted amino acid sequence of full-length *Fugu* NOS1 cDNA (*f*NOS). ATG start and TAG stop codons are highlighted in bold. Amino acids are shown in the single-letter code.

```
TACGGCTCCCTTATGATGCCCAACCTGCACGCCCGCAAACCGGAGGAGGTCAGGAGCAAA
Y  G  S  L  M  M  P  N  L  H  A  R  K  P  E  E  V  R  S  K

GAGGAGCTTCTGAAACTGGCCACTGACTTCATTGACCAGTATTACACCTCCATCAAAAGG
E  E  L  L  K  L  A  T  D  F  I  D  Q  Y  Y  T  S  I  K  R

TACGGCTCCAAGGCCCATACGGACAGACTGGAGGAGGTGACCAAGGAGATTGAAGCGACT
Y  G  S  K  A  H  T  D  R  L  E  E  V  T  K  E  I  E  A  T

GGAACCTACCAGCTGAAAGACACAGAACTGATTTATGGAGCCAAACATGCCTGGAGGAAT
G  T  Y  Q  L  K  D  T  E  L  I  Y  G  A  K  H  A  W  R  N

GCTGCCCGTTGTGTGGGAAGAATCCAGTGGTCTAAACTACGGGTTTTTGATGCTAGAGAC
A  A  R  C  V  G  R  I  Q  W  S  K  L  R  V  F  D  A  R  D

TGCACAACAGCTCATGGAATGTACAACTACATCTGTAACCACATCAAGTACGCTACCAAC
C  T  T  A  H  G  M  Y  N  Y  I  C  N  H  I  K  Y  A  T  N

AAAGGGAATCTGAGGTCAGCTATCACCATATTTCCTCCGAGGACAGATGGCAAACATGAC
K  G  N  L  R  S  A  I  T  I  F  P  P  R  T  D  G  K  H  D

TTTCGAGTGTGGAACAGTCAGCTGATTCGTTATGCTGGGTACAAACAGCCTGATGGTCAG
F  R  V  W  N  S  Q  L  I  R  Y  A  G  Y  K  Q  P  D  G  Q

ATCCTGGGGGACCCTGCTAATGTTGAATTTACTGAGATCTGCATACAGCTGGGATGGAAA
I  L  G  D  P  A  N  V  E  F  T  E  I  C  I  Q  L  G  W  K

GCTCCAAAAGGTCGCTTTGATGTTCTGCCCCTCCTCCTGCAAGCTAACGGAAATGACCCC
A  P  K  G  R  F  D  V  L  P  L  L  L  Q  A  N  G  N  D  P

GAGCTGTTTGAGATCCCCGAAGACCTGGTTCTGGAGGTGCCAATAATACACCCAAAGTAT
E  L  F  E  I  P  E  D  L  V  L  E  V  P  I  I  H  P  K  Y

GAGTGGTTCAAAGAACTGGCTCTCAAGTGGTACGCCCTCCCTGCCGTCTCAAACATGATG
E  W  F  K  E  L  A  L  K  W  Y  A  L  P  A  V  S  N  M  M

CTGGAGATCGGAGGCCTGGAGTTCACTGCCTGTCCCTTCAGTGGCTGGTACATGGGTACA
L  E  I  G  G  L  E  F  T  A  C  P  F  S  G  W  Y  M  G  T

GAGATCGGCGTGAGGGACTTCTGCGACACGTCCCGCTACAACATGCTGGAGGAGGTTGCA
E  I  G  V  R  D  F  C  D  T  S  R  Y  N  M  L  E  E  V  A

AACAAGATGGGCTTGGACACCAGAAAGACCTCCTCCCTCTGGAAAGATCAGGCTTTGGTG
N  K  M  G  L  D  T  R  K  T  S  S  L  W  K  D  Q  A  L  V

GAGGTCAACATCGCCGTCCTTCACAGCTTCCAGTCATGCAAAGTGACCATAGTGGACCAT
E  V  N  I  A  V  L  H  S  F  Q  S  C  K  V  T  I  V  D  H

CACTCGGCAACCGAGTCCTTCATGAAGCACATGGAGAACGAGTACCGGGTGCGAGGCGGC
H  S  A  T  E  S  F  M  K  H  M  E  N  E  Y  R  V  R  G  G

TGCCCCGGAGACTGGGTGTGGATTGTGCCTCCCATGTCCGGAAGTATCACGCCGGTTTTC
C  P  G  D  W  V  W  I  V  P  P  M  S  G  S  I  T  P  V  F

CACCAAGAAATGCTCAATTACCGCCTCACGCCTTCCTATGAGTACCAGCTTGATCCCTGG
H  Q  E  M  L  N  Y  R  L  T  P  S  Y  E  Y  Q  L  D  P  W

CATACCCATGTGTGGAAAGGAGTCAACGGGACGCCCACAAAGAAACGAGCCATTGGATTC
H  T  H  V  W  K  G  V  N  G  T  P  T  K  K  R  A  I  G  F
```

Figure A.2. continued.

```
AAAAAGCTTGCCAAAGCGGTGAAGTTTTCAGCCAAACTCATGGGTCACGCAATGGCCAAG
K  K  L  A  K  A  V  K  F  S  A  K  L  M  G  H  A  M  A  K

AGAGTCAAAGCCACCATATTGTTCGCCACCGAGACGGGAAAATCGCAGGATTACGCCAAA
R  V  K  A  T  I  L  F  A  T  E  T  G  K  S  Q  D  Y  A  K

ACTCTCTGCGAAATCTTCAAGCACGCATTCGACCCAAAGGTCATGTCTATGGACGATTAT
T  L  C  E  I  F  K  H  A  F  D  P  K  V  M  S  M  D  D  Y

GATGTGGTGGATCTGGAGCATGAGACGCTGGTGTTAGTGGTGACCAGCACGTTTGGCAAC
D  V  V  D  L  E  H  E  T  L  V  L  V  V  T  S  T  F  G  N

GGCGACCCCCCTGAGAACGGAGAGAAATTTGGAGCCGCCTTAATGGAGATGCGGGAGCCG
G  D  P  P  E  N  G  E  K  F  G  A  A  L  M  E  M  R  E  P

ACATCCAACGCAGAAGACAGAAAGAGCTACAAGGTCCGTTTCAACAGCGTGTCCTCCCAC
T  S  N  A  E  D  R  K  S  Y  K  V  R  F  N  S  V  S  S  H

TCCGACACTCGCAAGTCCTCAAGTGACGAACCGGACGCCAAGATTCACTTTGAAAGCACC
S  D  T  R  K  S  S  S  D  E  P  D  A  K  I  H  F  E  S  T

GGACCTCTGGCCAATGTCAGGTTCTCTGTGTTTGGCCTTGGATCCAGAGCCTACCCACAC
G  P  L  A  N  V  R  F  S  V  F  G  L  G  S  R  A  Y  P  H

TTCTGCGCCTTTGCCCACGCCGTGGACACGCTGTTTGAAGAGCTCGGGGGGGGAGCGCATC
F  C  A  F  A  H  A  V  D  T  L  F  E  E  L  G  G  E  R  I

CTACGCATGGGAGAAGGGGATGAGCTGTGTGGACAAGAGGAGGCTTTCAGAACCTGGGCG
L  R  M  G  E  G  D  E  L  C  G  Q  E  E  A  F  R  T  W  A

AAAAAGGTTTTTAAGGCCGCCTGCGACGTGTTCTGTGTCGGCGATGACGTGAACATCGAG
K  K  V  F  K  A  A  C  D  V  F  C  V  G  D  D  V  N  I  E

AAGGCCAACAACTCCCTGATCAGCAACGACCGCAGCTGGAAGAAAAACAAGTTCCGTCTG
K  A  N  N  S  L  I  S  N  D  R  S  W  K  K  N  K  F  R  L

ACATACACGGCCGAGGCGCCGAGCCTCACAAAAGCTTTGTACGGCGTGCACAAGAAGAAG
T  Y  T  A  E  A  P  S  L  T  K  A  L  Y  G  V  H  K  K  K

GTCCATGCAGCAAAGATGCTCGATTCTCAGAATTTACAGAGTCCGAAATCCAATCGTTCC
V  H  A  A  K  M  L  D  S  Q  N  L  Q  S  P  K  S  N  R  S

ACCATTCTCGTACGGCTGGACACAAATAACCAAGACAGCCTGAAATACAAGCCAGGAGAC
T  I  L  V  R  L  D  T  N  N  Q  D  S  L  K  Y  K  P  G  D

CATCTGGGCATCTTCCCTGGCAACCACGAGGACCTGGTGTCGGCTCTCATAGATAAGCTG
H  L  G  I  F  P  G  N  H  E  D  L  V  S  A  L  I  D  K  L

GAGGATGCGCCGCCTGTCAATCAGATTGTAAAAGTGGAGTTCTTAGAGGAGAGAAACACT
E  D  A  P  P  V  N  Q  I  V  K  V  E  F  L  E  E  R  N  T

GCCCTGGGTGTGATAAGTAACTGGACCAATGAGACTCGGGTCCCTCCCTGCACCATCAAC
A  L  G  V  I  S  N  W  T  N  E  T  R  V  P  P  C  T  I  N

CAGGCCTTCCAGTACTTCCTGGACATCACCACCCCGCCCAGCCCCGTACTGCTGCAGCAG
Q  A  F  Q  Y  F  L  D  I  T  T  P  P  S  P  V  L  L  Q  Q

TTCGCTGCTCTGGCCACTAATGAGAAAGAGAAACGGAAACTCGAGGTCCTCAGCAAGGGC
F  A  A  L  A  T  N  E  K  E  K  R  K  L  E  V  L  S  K  G
```

Figure A.2. continued.

```
     TTGCAGGAGTATGAGGAGTGGAAGTGGTACAACAACCCCACCCTTGTGGAGGTACTAGAG
     L  Q  E  Y  E  E  W  K  W  Y  N  N  P  T  L  V  E  V  L  E

     GAATTCCCCTCCATCCAGATGCCTTCTACGCTGCTGCTCTCCCAGCTTCCCCTGCTGCAG
     E  F  P  S  I  Q  M  P  S  T  L  L  L  S  Q  L  P  L  L  Q

CCTCGTTACTACTCCATCAGCTCCTCTCCAGACCTGCACCCGGGAGAGATCCACCTCACG
     P  R  Y  Y  S  I  S  S  S  P  D  L  H  P  G  E  I  H  L  T

     GTTGCTGTGGTCTCCTACCGTACCCGAGATGGAGCAGGGTCGATCCACCATGGAGTGTGT
     V  A  V  V  S  Y  R  T  R  D  G  A  G  S  I  H  H  G  V  C

     TCGTCGTGGCTCAGCAGGATAGAGAAGGGGGAGATGGTGCCGTGTTTTGTCCGAAGTGCT
     S  S  W  L  S  R  I  E  K  G  E  M  V  P  C  F  V  R  S  A

     CCATCCTTCCAGCTTCCCAAAAACAACCAAACACCTTGCATCCTGGTGGGTCCGGGAACC
     P  S  F  Q  L  P  K  N  N  Q  T  P  C  I  L  V  G  P  G  T

     GGAATCGCCCCATTCCGGAGCTTTTGGCAACAGCGACTGTATGACCTTGAACACAACGGC
     G  I  A  P  F  R  S  F  W  Q  Q  R  L  Y  D  L  E  H  N  G

     ATCGAGTCATGCCCAATGATCCTGGTGTTTGGCTGTCGACAGTCTGAGATTGACCACATC
     I  E  S  C  P  M  I  L  V  F  G  C  R  Q  S  E  I  D  H  I

     TACAAAGAGGAGACCATCCAAGCCAAGAACAAGAACGTGTTCAAGGAGCTGTACACGGCC
     Y  K  E  E  T  I  Q  A  K  N  K  N  V  F  K  E  L  Y  T  A

     TATTCCAGAGAGCCCGGCAAACCAAAGAAATATGTGCAGGATGCACTGCGTGAGCAGCTG
     Y  S  R  E  P  G  K  P  K  K  Y  V  Q  D  A  L  R  E  Q  L

     TCGGAGCGGGTGTACCAGTGCCTGAGGGAGGAGGGAGGACACATCTACGTGTGCGGGGAT
     S  E  R  V  Y  Q  C  L  R  E  E  G  G  H  I  Y  V  C  G  D

     GCTACGATGGCGGGGGATGTTCTCAAGAACGTCCAGCAGATCATCAAACAAGAGGGCAAC
     A  T  M  A  G  D  V  L  K  N  V  Q  Q  I  I  K  Q  E  G  N

     ATGAGCCTGGAGGAAGCCGGCTTGTTCATCAGCAAGCTTCGGGATGAGAACCGCTACCAT
     M  S  L  E  E  A  G  L  F  I  S  K  L  R  D  E  N  R  Y  H

     GAGGACATCTTTGGAGTCACCCTGCGCACCTACGAGGTCACCAGCAGGATTCGGTCAGAG
     E  D  I  F  G  V  T  L  R  T  Y  E  V  T  S  R  I  R  S  E

     TCCATTGCCTACATCGAAGAGAATAAAAAAGGATTCCGATGAGGTGTTCTGCTCATAGGTC
     S  I  A  Y  I  E  E  N  K  K  D  S  D  E  V  F  C  S  *

     TCTGAATCCGGGCCTTTCTGACCACGACCAGGCTGGACGAGGCGGCGGCGCTGGCTGAAG
     AATACTGT
```

Figure A.2. continued.

```
dnos      ------------------------------------------------------------
anonos    ------------------------------------------------------------
rpnos     ------------------------------------------------------------
orynos1   MEEHVFGVQQIQPNVISVRLFKRKVGGLGFLVKERVSKPPVIISDLIRGGAAEQSGLIQA
humnos1   MEDHMFGVQQIQPNVISVRLFKRKVGGLGFLVKERVSKPPVIISDLIRGGAAEQSGLIQA 60
ratnos1   MEENTFGVQQIQPNVISVRLFKRKVGGLGFLVKERVSKPPVIISDLIRGGAAEQSGLIQA
musnos1   MEEHTFGVQQIQPNVISVRLFKRKVGGLGFLVKERVSKPPVIISDLIRGGAAEQSGLIQA
xennos1   MEEYEFSVKQLQPNVISVRLFKRKVGGLGFLAKQRRNKPPVIISDLIRGGAAEQSGLVQV
fugnos1   MQESEPSVCLLQPNIISVRLFKRKVGGLGFLVKQRVSKPPVIVSDIIRGGAAEECGLVQV
susnos3   ------------------------------------------------------------
bovnos3   ------------------------------------------------------------
humnos3   ------------------------------------------------------------
musnos3   ------------------------------------------------------------
musnos2   ------------------------------------------------------------
ratnos2   ------------------------------------------------------------
humnos2   ------------------------------------------------------------
ccnos2    ------------------------------------------------------------
lymnos    ------------------------------------------------------------

dnos      -------------------------MSQHFTSIFENLRFVTIKRATNAQQQQQQQQQQQL
anonos    -------------------------MADTTTVVVERREVAEGRESSKANHIGEERRGYDV
rpnos     ------------------------------------------------------------
orynos1   GDIILAVNGRPLVDLSYDSALEVLRGVASETHVVLILRGPEGFTTNLETTFTGDGTPKTI
humnos1   GDIILAVNGRPLVDLSYDSALEVLRGIASETHVVLILRGPEGFTTHLETTFTGDGTPKTI 120
ratnos1   GDIILAVNDRPLVDLSYDSALEVLRGIASETHVVLILRGPEGFTTHLETTFTGDGTPKTI
musnos1   GDIILAVNDRPLVDLSYDSALEVLRGIASETHVVLILRGPEGFTTHLETTFTGDGTPKTI
xennos1   GDIILAVNDRPLVDASYESALEILRSISSETFVVLILRGPEGFTTHLETTFSGDGTPKTI
fugnos1   GDIVLAVNNKSLVDLSYERALEMLKNVLPESHAVLILRGPEGFTTHLETTISGDGRQRTV
susnos3   ------------------------------------------------------------
bovnos3   ------------------------------------------------------------
humnos3   ------------------------------------------------------------
musnos3   ------------------------------------------------------------
musnos2   ------------------------------------------------------------
ratnos2   ------------------------------------------------------------
humnos2   ------------------------------------------------------------
ccnos2    ------------------------------------------------------------
lymnos    ------------------------------------------------------------

dnos      QQQQQQLQQQKAQTQQQNSRKIKTQATPTLNGNGLLSGNPNGGGGDSSPSHEVDHPGGAQ
anonos    SRKR--------------------CSISVHG-----GGTEGGGGNMRTNYR--------
rpnos     ------------------------------------------------------------
orynos1   RVTQPLGAPTKAVDLSHQPPSAGKEQPRPVDGAAGPGSWPQPTQGHGQEAGSPSRANGLA
humnos1   RVTQPLGPPTKAVDLSHQPP-AGKEQPLAVDGASGPGNGPQHAYDDGQEAGSLPHANGLA 179
ratnos1   RVTQPLGPPTKAVDLSHQPS-ASKDQSLAVDRVTGLGNGPQHAQGHGQGAGSVSQANGVA
musnos1   RVTQPLGTPTKAVDLSRQPS-ASKDQPLAVDRVPGPSNGPQHAQGRGQGAGSVSQANGVA
xennos1   RVTRPLCSKSKSAELTSQSL---YSKDHMVDSATSSAG----SS--WQEVHQLLNLNGLD
fugnos1   RVTRPIFPASKSYENCSPLG--PFGPGQQVNKESQLRAIENLSS--PLQKGSVQAQDPLL
susnos3   ------------------------------------------------------------
bovnos3   ------------------------------------------------------------
humnos3   ------------------------------------------------------------
musnos3   ------------------------------------------------------------
musnos2   ------------------------------------------------------------
ratnos2   ------------------------------------------------------------
humnos2   ------------------------------------------------------------
ccnos2    ------------------------------------------------------------
lymnos    ------------------------------------------------------------
```

Figure A.3. Clustal W (1.74) multiple sequence alignment of full-length NOS amino acid sequences. Names abbreviated as per figure 4.6, with human NOS1 numbered .

```
dnos      GAQAAGGLPSLSGTPLRHHKRASISTASPPIRERRGTNTSIVVELDGSGSG---------
anonos    ---------ELSPASLRIHRK-S----SHDIRNTLLGPDGEVLHLHDP------------
rpnos     ----------------MVHAECVW---WLGIRILFVPPVSLEMHSVNVNNM---------
orynos1   PRTSSQDPAKKSGWAGLQGSGDKN-----ELLKEIEPVLTLLAGGSKAVDGGGPAKAETR
humnos1   PRPPGQDPAKKATRVSLQGRGENN-----ELLKEIEPVLSLLTSGSRGVKGGAPAKAEMK 234
ratnos1   -----IDPTMKSTKANLQDIGEHD-----ELLKEIEPVLSILNSGSKATNRGGPAKAEMK
musnos1   -----IDPTMKNTKANLQDSGEQD-----ELLKEIEPVLSILTGGGKAVNRGGPAKAEMK
xennos1   ----GGKGDDISISPSLNRGEKAN-----DILKEIEPIVSLLQNAGKELNGDEHRNVEQK
fugnos1   ------LRDGGRG--LCNGLEDNN-----ELMKEIEPVLRLVKNSKKEINGEGQRHVGRR
susnos3   -------------------MGNLK-----SVGQEPGPPCGLGLGLGLGLCG---K-----
bovnos3   -------------------MGNLK-----SVGQEPGPPCGLGLGLGLGLCG---K-----
humnos3   -------------------MGNLK-----SVAQEPGPPCGLGLGLGLGLCG---K-----
musnos3   -------------------MGNLK-----SVGQEPGPPCGLGLGLGLGLCG---K-----
musnos2   -------------------MACPW-----KFLFKVKSYQSDLTEEKDINNN---------
ratnos2   -------------------MACPW-----KFLFKVKSYQGDLKEEKDINNN---------
humnos2   -------------------MACPW-----KFLFKTKFHQYAMNGEKDINNN---------
ccnos2    -------------------MGNQA-----TKANKNATPHQITPNTQCENNN---------
lymnos    ------------------------------------------------------------


dnos      --------------------SGSGGGGVGVGQGAGCPPSGSCTASGKSSRELSPSPKNQQ
anonos    --------------------SGKGGDGMG-------------------------------
rpnos     --------------------SIQQQQQHQ-------------------------------
orynos1   DTGVQVDRDFDAKSHKPLPLGVENDRVFSDLWGKGSAPVVLN--------NPYSEKEQPP
humnos1   DMGIQVDRDLDGKSHKPLPLGVENDRVFNDLWGKGNVPVVLN--------NPYSEKEQPP 286
ratnos1   DTGIQVDRDLDGKSHKAPPLGGDNDRVFNDLWGKDNVPVVLN--------NPYSEKEQSP
musnos1   DTGIQVDRDLDGKLHKAPPLGGENDRVFNDLWGKGNVPVVLN--------NPYSENEQSP
xennos1   DAEVQVE------SNSQTQPSMQKDQVNGIWKNNNKMPVVLN--------NPYLECEQVT
fugnos1   DAEIQVTWG----AGVGIDTSLQLDSCKNKMPEKE--PGVPQ--------N--ADNDKPP
susnos3   ----------------------QGPATPAPEPSRAPAPATP------------HAPEHSP
bovnos3   ----------------------QGPASPAPEPSRAPAPATP------------HAPDHSP
humnos3   ----------------------QGPATPAPEPSRAPASLLP------------PAPEHSP
musnos3   ----------------------QGPASPAPEPSQAPAPPSPT--------R--AAPDHSP
musnos2   --------------------VKKTPCAVLSPTIQDDPKS-----------HQNGSPQLLT
ratnos2   --------------------VEKTPGAIPSPTTQDDPKSH---------KHQNGFPQFLT
humnos2   --------------------VEKAPCATSSPVTQDDLQYHNLS------KQQNESPQPLV
ccnos2    --------------------VILKKTTPNTQCENNNVILQP--------ITPNMCENNNV
lymnos    ------------------------------------------------------------


dnos      QPRKMSQDYRSRAGSFMHLDDEGRSLLMRKPMRLKNIEGRPEVYDTLHCKGREILSCSKA
anonos    ---KMP--------------------AVVKPIKLKSIVTKAESYDTMHGKASDVMSCSRE
rpnos     ---PQQ--------------------QLLKPIRLANVSTQAQSLDTLHYKCQQEGPCMEQ
orynos1   ASGKQSPT-K-----------NGSPSKCPRFLKVKNWETDVVLTDTLHLKSTLETGCTEH
humnos1   TSGKQSPT-K-----------NGSPSKCPRFLKVKNWETEVVLTDTLHLKSTLETGCTEY 334
ratnos1   TSGKQSPT-K-----------NGSPSRCPRFLKVKNWETDVVLTDTLHLKSTLETGCTEH
musnos1   ASGKQSPT-K-----------NGSPSRCPRFLKVKNWETDVVLTDTLHLKSTLETGCTEQ
xennos1   LTGRHSPA-KSQV--------NGSPSKCPRYLKIRNWDSNTTLNDTLHSKASMPTPCTQQ
fugnos1   AEARTSPT-KSLQ--------NGSPSKCPRFLKIKNWETGAIQNDTLHNSSTKTPMCPEN
susnos3   A--PNSPT-L------------TRPPEGPKFPRVKNWEVGSITYDTLCAQSQQDGPCTPR
bovnos3   A--PNSPT-L------------TRPPEGPKFPRVKNWELGSITYDTLCAQSQQDGPCTPR
humnos3   ---PSSP--L------------TQPPEGPKFPRVKNWEVGSITYDTLSAQAQQDGPCTPR
musnos3   -----PL---------------TRPPDGPRFPRVKNWEVGSITYDTLSAQAQQDGPCTSR
musnos2   GTAQNVPESLDKL--------HVT-STRPQYVRIKNWGSGEILHDTLHHKATSDFTCKSK
ratnos2   GTAQNVPESLDKL--------HVTPSTRPQHVRIKNWGNGEIFHDTLHHKATSDISCKSK
humnos2   ETGKKSPESLVKL--------DATPLSSPRHVRIKNWGSGMTFQDTLHHKAKGILTCRSK
ccnos2    ILQQITPNMKWKN--------KVN--RCPFSKQLKNYQDGLFHQDTLHSRAVKSQICMSN
lymnos    ------------------------------------------------------------
```

Figure A.3 continued.

```
dnos      TCTSSIMNIGN----AAVEARKSDLILEHAKDFLEQYFTSIKRTSCTAHETRWKQVRQSI
anonos    VCMGSVMTPHV----IGTETRKPEIVQQHAKDFLDQYYSSIRRLKSPAHDSRWQQVQKEV
rpnos     ACLASVIYAGVN---LKPRVRPKEELLAHAKDFLDQYFASIRRLQSPAHEARWAQVEKEV
orynos1   ICMGSIMFPSQHTRRP-EDIRTKEQLFPLAKEFIDQYYSSIKRFGSKAHMERLEEVNKEI
humnos1   ICMGSIMHPSQHARRP-EDVRTKGQLFPLAKEFIDQYYSSIKRFGSKAHMERLEEVNKEI 393
ratnos1   ICMGSIMLPSQHTRKP-EDVRTKDQLFPLAKEFLDQYYSSIKRFGSKAHMDRLEEVNKEI
musnos1   ICMGSIMLPSHHIRKS-EDVRTKDQLFPLAKEFLDQYYSSIKRFGSKAHMDRLEEVNKEI
xennos1   ICMGAVMTPPHYVRKP-EDIRTKEELLPLAKEFIDQYFSSIKRFGSKAHMDRLQEVTKEI
fugnos1   VCYGSLMMPNLHARKP-EEVRSKEELLKLATDFIDQYYTSIKRYGSKAHTDRLEEVTKEI
susnos3   RCLGSLVLPRKLQSRPSPGPPPAEQLLSQARDFINQYYSSIKRSGSQAHEERLQEVEAEV
bovnos3   CCLGSLVLPRKLQTRPSPGPPPAEQLLSQARDFINQYYSSIKRSGSQAHEERLQEVEAEV
humnos3   RCLGSLVFPRKLQGRPSPGPPAPEQLLSQARDFINQYYSSIKRSGSQAHEQRLQEVEAEV
musnos3   RCLGSLVFPRKLQSRPTQGPSPTEQLLGQARDFINQYYNSIKRSGSQAHEQRLQEVEAEV
musnos2   SCLGSIMNPKSLTRGPRDKPTPLEELLPHAIEFINQYYGSFKEAKIEEHLARLEAVTKEI
ratnos2   LCMGSIMNSKSLTRGPRDKPTPVEELLPQAIEFINQYYGSFKEAKIEEHLARLEAVTKEI
humnos2   SCLGSIMTPKSLTRGPRDKPTPPDELLPQAIEFVNQYYGSFKEAKIEEHLARVEAVTKEI
ccnos2    VCEGSVMTPKAMTRCPSSTMPGSDDILTQAVDFINQYYKSIKNSKIEEHLSRLEEVTKEI
lymnos    --MGSLSQQAHG---PPDAPRSKEELLIHAKDFINQYFTSFQMNKTRAHFHRLGEINDLI
             .::                      :   * :*::**: *::      *  *   :   :

dnos      ETTGHYQLTETELIYGAKLAWRNSSRCIGRIQWSKLQVFDCRYVTTTSGMFEAICNHIKY
anonos    EASGSYHLTETELIYGAKLAWRNSSRCIGRIQWSKLQVFDCRYVTTTSGMFEAICNHIKY
rpnos     AATGTYELTETELVYGAKLAWRNAPRCIGRIQWAKLQVFDCRQVTTTSGMFEALCNHIKY
orynos1   ESTSTYQLKDTELIYGAKHAWRNASRCVGRIQWSKLQVFDARDCTTAHGMFNYICNHIKY
humnos1   DTTSTYQLKDTELIYGAKHAWRNASRCVGRIQWSKLQVFDARDCTTAHGMFNYICNHVKY 453
ratnos1   ESTSTYQLKDTELIYGAKHAWRNASRCVGRIQWSKLQVFDARDCTTAHGMFNYICNHVKY
musnos1   ESTSTYQLKDTELIYGAKHAWRNASRCVGRIQWSKLQVFDARDCTTAHGMFNYICNHVKY
xennos1   EATDTYQLRDTELIYGAKHAWRNASRCVGRIQWSKLQLFDARDCTTAHGMFNYICNHIKY
fugnos1   EATGTYQLKDTELIYGAKHAWRNAARCVGRIQWSKLQVFDARDCTTAHGMYNYICNHIKY
susnos3   ATTGTYHLGESELVFGAKQAWRNAPRCVGRIQWGKLQVFDARDCSSAQEMFTYICNHIKY
bovnos3   ASTGTYHLRESELVFGAKQAWRNAPRCVGRIQWGKLQVFDARDCSSAQEMFTYICNHIKY
humnos3   AATGTYQLRESELVFGAKQAWRNAPRCVGRIQWGKLQVFDARDCRSAQEMFTYICNHIKY
musnos3   AATGTYQLRESELVFGAKQAWRNAPRCVGRIQWGKLQVFDARDCRTAQEMFTYICNHIKY
musnos2   ETTGTYQLTLDELIFATKMAWRNAPRCIGRIQWSNLQVFDARNCSTAQEMFQHICRHILY
ratnos2   ETTGTYQLTLDELIFATKMAWRNAPRCIGRIQWSNLQVFDARSCSTASEMFQHICRHILY
humnos2   ETTGTYQLTGDELIFATKQAWRNAPRCIGRIQWSNLQVFDARSCSTAREMFEHICRHVRY
ccnos2    EATGSYRLTTKELEFGAKQAWRNAPRCIGRIQWANLQLFDARKCRTAEDMFQMLCDHIQF
lymnos    EKSGTYDLTMAELTFGAKHAWRNAPGCIGRSQWSKLQVFDAREIGTPREMFEALCSHIRY
             :. * *   ** :.:* ****:. *:** **.:**:**.*   :.  *:  :* *: :

dnos      ATNKGNLRSAITIFPQRTDAKHDYRIWNNQLISYAGYKQADGKIIGDPMNVEFTEVCTKL
anonos    ATNKGNLRSAITIFPQRTDGKHDYRIWNNQIISYAGYKNADGKIIGDPANVEFTDFCVKL
rpnos     STNKGNIRSAITIFPHRTDGKHDFRIWNKQLISYAGHKSKDGTVIGDPACVEFTEICIKL
orynos1   ATNKGNLRSAITIFPQRTDGKHDFRVWNSQLIRYAGYKQPDGSTLGDPANVQFTEICIQQ
humnos1   ATNKGNLRSAITIFPQRTDGKHDFRVWNSQLIRYAGYKQPDGSTLGDPANVQFTEICIQQ 513
ratnos1   ATNKGNLRSAITIFPQRTDGKHDFRVWNSQLIRYAGYKQPDGSTLGDPANVQFTEICIQQ
musnos1   ATNKGNLRSAITIFPQRTDGKHDFRVWNSQLIRYAGYKQPDGSTLGDPANVEFTEICIQQ
xennos1   ATNKGNLRSAITIFPQRTDMKHDFRIWNAQLIRYAGYKQPDGSVLGDPANVELTEICIQQ
fugnos1   ATNKGNLRSAITIFPPRTDGKHDFRVWNSQLIRYAGYKQPDGQILGDPANVEFTEICMQL
susnos3   ATNRGNLRSAITVFPQRTPGRGDFRIWNSQLVRYAGYRQQDGSVRGDPANVEITELCIQH
bovnos3   ATNRGNLRSAITVFPQRAPGRGDFRIWNSQLVRYAGYRQQDGSVRGDPANVEITELCIQH
humnos3   ATNRGNLRSAITVFPQRCPGRGDFRIWNSQLVRYAGYRQQDGSVRGDPANVEITELCIQH
musnos3   ATNRGNLRSAITVFPQRCPGRGDFRIWNSQLIRYAGYRQQDGSVRGDPANVEITELCIQH
musnos2   ATNNGNIRSAITVFPQRSDGKHDFRLWNSQLIRYAGYQMPDGTIRGDAATLEFTQLCIDL
ratnos2   ATNSGNIRSAITVFPQRSDGKHDFRIWNSQLIRYAGYQMPDGTIRGDPATLEFTQLCIDL
humnos2   STNNGNIRSAITVFPQRSDGKHDFRVWNAQLIRYAGYQMPDGSIRGDPANVEFTQLCIDL
ccnos2    ATNGGNVRSAITVFPQRTDGQHDFRVWNSQLIRYAGYKMTDGTIIGDPASVDFTEICIEL
lymnos    ATNEGKIRSTITIFPQRKEGRPDFRVWNTQLISYAGYKLGDGKVIGDPANVEFTEMCVEM
          :** *::**:**:** *   : *:*:** *:: ***::  **   **.  :::*:.* .
```

Figure A.3. contiunued.

```
dnos      GWKSKGSEWDILPLVVSANGHDPDYF-DYPPELILEVPLTHPKFEWFSDLGLRWYALPAV
anonos    GWKSKRTEWDILPLVVSANGHDPDYF-DYPPELILEVPLSHPQFKWFAELNLRWYAVPMV
rpnos     GWKGKGTMFDVLPLVLSANGEDPDYF-DLPPELVFEVPLSHPKYKWFSELGLKWFALPAV
orynos1   GWKPPRSRFDVLPLLLQANGNDPELF-QIPPELVLEVPIRHPKFEWFKDLGLKWYGLPAV
humnos1   GWKPPRGRFDVLPLLLQANGNDPELF-QIPPELVLEVPIRHPKFEWFKDLGLKWYGLPAV 572
ratnos1   GWKAPRGRFDVLPLLLQANGNDPELF-QIPPELVLEVPIRHPKFDWFKDLGLKWYGLPAV
musnos1   GWKPPRGRFDVLPLLLQANGNDPELF-QIPPELVLEVPIRHPKFDWFKDLGLKWYGLPAV
xennos1   GWKALHGRFDVLPLMLQADGNDPELF-EIPPELIMEVNIRHPRLEWFKDLGLKWYGLPAV
fugnos1   GWKAPKGRFDVLPLLLQANGNDPEQF-EIPEDLVLEVPIIHPKYEWFKELALKWYALPAV
susnos3   GWTPGNGRFDVLPLLLQAPDEPPELF-ALPPELVLEVPLEHPTLEWFAALGLRWYALPAV
bovnos3   GWTPGNGRFDVLPLLLQAPDEAPELF-VLPPELVLEVPLEHPTLEWFAALGLRWYALPAV
humnos3   GWTPGNGRFDVLPLLLQAPDEPPELF-LLPPELVLEVPLEHPTLEWFAALGLRWYALPAV
musnos3   GWTPGNGRFDVLPLLLQAPDESPELF-TLPPEMVLEVPLEHPTLEWFAALGLRWYALPAV
musnos2   GWKPRYGRFDVLPLVLQADGQDPEVF-EIPPDLVLEVTMEHPKYEWFQELGLKWYALPAV
ratnos2   GWKPRYGRFDVLPLVLQAHGQDPEVF-EIPPDLVLEVTMEHPKYEWFQELGLKWYALPAV
humnos2   GWKPKYGRFDVVPLVLQANGRDPELF-EIPPDLVLEVAMEHPKYEWFRELELKWYALPAV
ccnos2    GWTPRYGQFDVLPLVLQATEEDPSVFLKFPQHLILEVPMKHQQYKWFKDLNLRWFALPAV
lymnos    GWKPKHGMFDLLPLVLSAAENSPEYF-ELPTELVLEVTLKHPEYPWFAEMGLKWYALPTD
          **.     :*::**::.*  . *. *   * .:::** : *    **  : *:*:.:*

dnos      SSMLFDVGGIQFTATTFSGWYMSTEIGSRNLCDTNRRNMLETVALKMQLDTRTPTSLWKD
anonos    SSMLFDCGGIQFTATAFSGWYMSTEIGCRNLCDANRRNLLEPIAIKMGLDTRNPTSLWKD
rpnos     SGMMFDCGGLQFTAAPFNGWYMNSEIGSRNLGDTNRYNMLEKIAQKMELDTRTPVTLWKD
orynos1   SNMLLEIGGLEFSACPFSGWYMGTEIGVRDYCDNSRYNILEEVAKKMNLDMRKTSSLWKD
humnos1   SNMLLEIGGLEFSACPFSGWYMGTEIGVRDYCDNSRYNILEEVAKKMNLDMRKTSSLWKD 632
ratnos1   SNMLLEIGGLEFSACPFSGWYMGTEIGVRDYCDNSRYNILEEVAKKMDLDMRKTSSLWKD
musnos1   SNMLLEIGGLEFSACPFSGWYMGTEIGVRDYCDNSRYNILEEVAKKMDLDMRKTSSLWKD
xennos1   SNLLLEIGGLEFSACPFSGWYMGTEIGVRDYCDNSRYNILEEVALKMNLDTRKTSSLWKD
fugnos1   SNMMLEIGGLEFTACPFSGWYMGTEIGVRDFCDTSRYNMLEEVANKMGLDTRKTSSLWKD
susnos3   SNMLLEIGGLEFPAAPFSGWYMSTEIGTRNLCDPHRYNILEDVAVCMDLDTRTTSSLWKD
bovnos3   SNMLLEIGGLEFSAAPFSGWYMSTEIGTRNLCDPHRYNILEDVAVCMDLDTRTTSSLWKD
humnos3   SNMLLEIGGLEFPAAPFSGWYMSTEIGTRNLCDPHRYNILEDVAVCMDLDTRTTSSLWKD
musnos3   SNMLLEIGGLEFPAAPFSGWYMSSEIGMRDLCDPHRYNILEDVAVCMDLDTRTTSSLWKD
musnos2   ANMLLEVGGLEFPACPFNGWYMGTEIGVRDFCDTQRYNILEEVGRRMGLETHTLASLWKD
ratnos2   ANMLLEVGGLEFPACPFNGWYMGTEIGVRDLCDTQRYNILEEVGRRMGLETHTLASLWKD
humnos2   ANMLLEVGGLEFPGCPFNGWYMGTEIGVRDFCDVQRYNILEEVGRRMGLETHKLASLWKD
ccnos2    SNMLLEIGGLEFPACPFNGWYMGTEIGVRDFCDTKRYNVLERVGRQMGLETQKLPSLWKD
lymnos    SGMLLDCGGLEFPSCPFNGWFMGTMIGSRNLCDPHRYNMLEPIGLKMGLNTETASSLWKD
          :.::::: **::*.. .*.**:*.: ** *:  *  * *:** :.  * *: ..  :****

dnos      KAVVEMNIAVLHSYQSRNVTIVDHHTASESFMKHFENESKLRNGCPADWIWIVPPLSGSI
anonos    KALVEINIAVLHSYQSRNITIVDHHTASESFMKHFENETKLRNGCPADWIWIVPPMSASV
rpnos     LAMVEANVAVLHSFQLHNVTIVDHHSAAESFMKHLENEQRLRGGCPADWVWIVPPISGSA
orynos1   QALVEINIAVLYSFQSDKVTIVDHHSATESFIKHMENEYRCRGGCPADWVWIVPPMSGSI
humnos1   QALVEINIAVLYSFQSDKVTIVDHHSATESFIKHMENEYRCRGGCPADWVWIVPPMSGSI 692
ratnos1   QALVEINIAVLYSFQSDKVTIVDHHSATESFIKHMENEYRCRGGCPADWVWIVPPMSGSI
musnos1   QALVEINIAVLYSFQSDKVTIVDHHSATESFIKHMENEYRCRGGCPADWVWIVPPMSGSI
xennos1   QALVEINIAVLYSFQSDKVTIVDHHSATESFIKHMENEYRCRGGCPADWVWIVPPMSGSI
fugnos1   QALVEVNIAVLHSFQSCKVTIVDHHSATESFMKHMENEYRVRGGCPGDWVWIVPPMSGSI
susnos3   KAAVEINLAVLHSYQLAKVTIVDHHAATASFMKHLENEQKARGGCPADWAWIVPPISGSL
bovnos3   KAAVEINLAVLHSFQLAKVTIVDHHAATVSFMKHLDNEQKARGGCPADWAWIVPPISGSL
humnos3   KAAVEINVAVLHSYQLAKVTIVDHHAATASFMKHLENEQKARGGCPADWAWIVPPISGSL
musnos3   KAAVEINVAVLHSYQLAKVTIVDHHAATASFMKHLENEQKARGGCPADWAWIVPPISGSL
musnos2   RAVTEINVAVLHSFQKQNVTIMDHHTASESFMKHMQNEYRARGGCPADWIWLVPPVSGSI
ratnos2   RAVTEINAAVLHSFQKQNVTIMDHHTASESFMKHMQNEYRARGGCPADWIWLVPPVSGSI
humnos2   QAVVEINIAVLHSFQKQNVTIMDHHSAAESFMKYMQNEYRSRGGCPADWIWLVPPMSGSI
ccnos2    QALVAINVAVMHSFQKNKVTITDHHTAPESFMQHMEMEVRLRGGCPADWVWLVPPMSGSL
lymnos    RVLIEVNVAVLYSFESANVTIVNHHDASTDFISHMDKEIKLRGGCPSDWVRMVPPMSGST
           .    * **::*::  ::** :** *. .*:.::: * : *.***.**  :***:*.*
```

Figure A.3. continued.

```
dnos      TPVFHQEMALYYLKPSFEYQ-DPAWRTHVWKKGRGESK-GKKPRRKFNFKQIARAVKFTS
anonos    TPVFHQEMAVYYLRPSFEYQ-ESAMKTHIWKKGRDSAK-NKKPRRKFNFKQIARAVKFTS
rpnos     TPVFFQEMANYFLYPGYIYQ-EDAWKCHEWKEIDVKHG-LKKEKRKFHFKQIARAVKFTS
orynos1   TPVFHQEMLNYRLTPCFEYQ-PDPWNTHVWKGTNGTP----TKRRAIGFKKLAEAVKFSA
humnos1   TPVFHQEMLNYRLTPSFEYQ-PDPWNTHVWKGTNGTP----TKRRAIGFKKLAEAVKFSA 747
ratnos1   TPVFHQEMLNYRLTPSFEYQ-PDPWNTHVWKGTNGTP----TKRRAIGFKKLAEAVKFSA
musnos1   TPVFHQEMLNYRLTPSFEYQ-PDPWNTHVWKGTNGTP----TKRRAIGFKKLAEAVKFSA
xennos1   TPVFHQEMLNYRLTPSFEYQ-PDPWNTHVWKGVNGTP----TKKRAIGFKKLAKAVKFSA
fugnos1   TPVFHQEMLNYRLTPSYEYQ-LDPWHTHVWKGVNGTP----TKKRAIGFKKLAKAVKFSA
susnos3   TPVFHQEMVNYVLSPAFRYQ-PDPWKGSAAKGT-G------IAR-KKTFKEVANAVKISA
bovnos3   TPVFHQEMVNYILSPAFRYQ-PDPWKGSATKGA-G------ITR-KKTFKEVANAVKISA
humnos3   TPVFHQEMVNYFLSPAFRYQ-PDPWKGSAAKGT-G------ITR-KKTFKEVANAVKISA
musnos3   TPVFHQEMVNYFLSPAFRYQ-PDPWKGSAAKGA-G------ITR-KKTFKEVANAVKISA
musnos2   TPVFHQEMLNYVLSPFYYYQ-IEPWKTHIWQNEKLR-----PRRREIRFRVLVKVVFFAS
ratnos2   TPVFHQEMLNYVLSPFYYYQ-IEPWKTHIWQDEKLR-----PRRREIRFTVLVKAVFFAS
humnos2   TPVFHQEMLNYVLSPFYYYQ-VEAWKTHVWQDEKRR-----PKRREIPLKVLVKAVLFAC
ccnos2    TPVYHQEMLNYILSPFFYYQ-PDPWLTHKWKVKKRN-----ARRHTISFKGLIRAVLFSQ
lymnos    LEVFHQEMLLYNLHPAFVRQDVKPWKKHVWKSDQSVPINSCNPKRKLGFKALARAVEFSA
            *:.***  * * * :  *   .      :            :    :  : ..* ::

dnos      KLFGRALSKRIKATVLYATETGKSEQYAKQLCELLGHAFNAQIYCMSDYDISSIEHEALL
anonos    KLFGRALSRRIKATVLYATETGRSEQYARQLVELLGHAFNAQIYCMSDYDISSIEHEALL
rpnos     KLFGSALSKRIKATILFATETGKSEMYARKLGDIFSHAFHSQVLSMEDYDMSKIEHEALL
orynos1   KLMGQAMAKRVKATILYATETGKSQAYAKTLCEIFKHAFDAKVMSMEEYDIVHLEHETLV
humnos1   KLMGQAMAKRVKATILYATETGKSQAYAKTLCEIFKHAFDAKVMSMEEYDIVHLEHETLV 807
ratnos1   KLMGQAMAKRVKATILYATETGKSQAYAKTLCEIFKHAFDAKAMSMEEYDIVHLEHEALV
musnos1   KLMGQAMAKRVKATILYATETGKSQAYAKTLCEIFKHAFDAKAMSMEEYDIVHLEHEALV
xennos1   KLMGQAMAKRVKATILYATETGKSQVYAKTLCEIFKHAFDAKVMSMDEYDVVHLEHETLV
fugnos1   KLMGHAMAKRVKATILFATETGKSQDYAKTLCEIFKHAFDPKVMSMDDYDVVDLEHETLV
susnos3   SLMATVMPKRVKASILYASETVRAQSYAQQLGRLFRKAFDPRVLCMDEYDVVSLEHETLV
bovnos3   SLMGTLMAKRVKATILYASETGRAQSYAQQLGRLFRKAFDPRVLCMDEYDVVSLEHEALV
humnos3   SLMGTVMAKRVKATILYGSETGRAQSYAQQLGRLFRKAFDPRVLCMDEYDVVSLEHETLV
musnos3   SLMGTVMAKRVKATILYGSETGRAQSYAQQLGRLFRKAFDPRVLCMDEYDVVSLEHEALV
musnos2   MLMRKVMASRVRATVLFATETGKSEALARDLATLFSYAFNTKVVCMDQYKASTLEEEQLL
ratnos2   VLMRKVMASRVRATVLFATETGKSEALARDLAALFSYAFNTKVVCMEQYKANTLEEEQLL
humnos2   MLMRKTMASRVRVTILFATETGKSEALAWDLGALFSCAFNPKVVCMDKYRLSCLEEERLL
ccnos2    TLIKSALTKRVHCTVLYATETGKSHTFAKKLNTMMNCAFKSQVVSMEDYNFSELEKESFL
lymnos    SLMSKALSSRVKCSIFYATETGRSERFARRLSEIFKPVFHSRVVCMDDYAVETLEHESLV
            *:   :. *:: :::::.:** ::.  *  *  ::  .*..:  .*..*    :*.* ::

dnos      IVVASTFGNGDPPENGELFSQELYAMRVQESSEHGLQDSS--------------------
anonos    LVVASTFGNGDPPENGQLFAQDLYAMKLHESGHHQAHSELT--------------------
rpnos     LVVASTFGNGDPPENGQGFAQSLYTIKMDENGLPNGHTNN--------------------
orynos1   LVVTSTFGNGDPPENGEKFRCALMEMRHPNSLQEER------------------------
humnos1   LVVTSTFGNGDPPENGEKFGCALMEMRHPNSVQEER------------------------ 843
ratnos1   LVVTSTFGNGDPPENGEKFGCALMEMRHPNSVQEERKYPEPLRFFPRKGPSLSHVDSEAH
musnos1   LVVTSTFGNGDPPENGEKFGCALMEMRHPNSVQEE-----------R-------------
xennos1   LVVTSTFGNGDPPENGEKFGCALMEMRHPNSNLEER------------------------
fugnos1   LVVTSTFGNGDPPENGEKFGAALMEMREPTSNTEDR------------------------
susnos3   LVVTSTFGNGDPPENGESFAAALMEMSGPYNGSPRP----------E-------------
bovnos3   LVVTSTFGNGDPPENGESFAAALMEMSGPYNSSPRP----------E-------------
humnos3   LVVTSTFGNGDPPENGESFAAALMEMSGPYNSSPRP----------E-------------
musnos3   LVVTSTFGNGDPPENGESFAAALMEMSGPYNSSPRP----------E-------------
musnos2   LVVTSTFGNGDCPSNGQTLKKSLFMLR------ELN------------------------
ratnos2   LVVTSTFGNGDCPSNGQTLKKSLFMMK------ELG------------------------
humnos2   LVVTSTFGNGDCPGNGEKLKKSLFMLK------ELN------------------------
ccnos2    IVVTSTFGNGDCPGNGESFKKQLLSLN------NLR------------------------
lymnos    MVITSTFGNGEPPENGKQFAQSLLDMKRKYDCDLGFLES--------------------
          :*::******: * **: :   *  :
```

Figure A.3. continued.

```
dnos      -------IGSSKSFMKASSRQEFMKLPLQQVKRIDRWDSLRGSTSDTFTEETFGPLSNVR
anonos    ------IAASSKSFIKANSRSDLGKFGPMGGRKIDRLDSLRGSTTDTLSEETFGPLSNVR
rpnos     ------TLASSASFIKANSQTDRQAS-------LERCDSFRGSTGD---ADVFGPLSNVR
orynos1   ----------KSYKVRFNSVSSYSDS----------RKSSGDGPDVRDHFESAGPLANVR
humnos1   ----------KSYKVRFNSVSSYSDS----------QKSSGDGPDLRDNFESAGPLANVR 883
ratnos1   SLVAARDSQHRSYKVRFNSVSSYSDS----------RKSSGDGPDLRDNFESTGPLANVR
musnos1   ----------KSYKVRFNSVSSYSDS----------RKSSGDGPDLRDNFESTGPLANVR
xennos1   ----------KSYKVRFNSVSSSADA----------QKSSANETGPRDNFGSASPLANVR
fugnos1   ----------KSYKVRFNSVSSHSDT----------RKSSSDEPDAKIHFESTGPLANVR
susnos3   --------QHRSYKIRFNSVSCSDPL----------VSSWRRKRKESSNTDSAGALGTLR
bovnos3   --------QHKSYKIRFNSVSCSDPL----------VSSWRRKRKESSNTDSAGALGTLR
humnos3   --------QHKSYKIRFNSISCSDPL----------VSSWRRKRKESSNTDSAGALGTLR
musnos3   --------QHKSYKIRFNSVSCSDPL----------VSSWRRKRKESSNTDSAGALGTLR
musnos2   ------------------------------------------------------HTFR
ratnos2   ------------------------------------------------------HTFR
humnos2   ------------------------------------------------------NKFR
ccnos2    ------------------------------------------------------NQVR
lymnos    -------CSSISTCIKSSILTEGPLA----------ADVIGDRQSLAMG---TGPLCNVR
                                                                  .*

dnos      FAVFALGSSAYPNFCAFGQYVDNILGELGGERLLRVAYGDEMCGQEQSFRKWAPEVFKLA
anonos    FAVFALGSSAYPNFCAFGKYIDNILGELGGERLMKMATGDEICGQEQAFRKWAPEVFKIA
rpnos     FAVFALGSSAYPNFCAFGSYVDNLLGELGGERLVKLTTGDEMCGQAQACNKWAPEVFSVA
orynos1   FSVFGLGSRAYPHFCAFGHAVDTLLEELGGERILKMREGDELCGQEEAFRTWAKKVFKAA
humnos1   FSVFGLGSRAYPHFCAFGHAVDTLLEELGGERILKMREGDELCGQEEAFRTWAKKVFKAA 943
ratnos1   FSVFGLGSRAYPHFCAFGHAVDTLLEELGGERILKMREGDELCGQEEAFRTWAKKVFKAA
musnos1   FSVFGLGSRAYPHFCAFGHAVDTLLEELGGERILKMREGDELCGQEEAFRTWAKKVFKAA
xennos1   FSAFGLGSRAYPHFCAFARAVDTLLEELGGERIMKMGEGDELCGQEESFRTWAKKVFQAA
fugnos1   FSVFGLGSRAYPHFCAFAHAVDTLFEELGGERILRMGEGDELCGQEEAFRTWAKKVFKAA
susnos3   FCVFGLGSRAYPHFCAFARAVDTRLEELGGERLLQLGQGDELCGQEEAFRGWAQAAFQAS
bovnos3   FCVFGLGSRAYPHFCAFARAVDTRLEELGGERLLQLGQGDELCGQEEAFRGWAKAAFQAS
humnos3   FCVFGLGSRAYPHFCAFARAVDTRLEELGGERLLQLGQGDELCGQEEAFRGWAQAAFQAA
musnos3   FCVFGLGSRAYPHFCAFARAVDTRLEELGGERLLQLGQGDELCGQEEAFRGWAQAAFQAA
musnos2   YAVFGLGSSMYPQFCAFAHDIDQKLSHLGASQLAPTGEGDELSGQEDAFRSWAVQTFRAA
ratnos2   YAVFGLGSSMYPQFCAFAHDIDQKLSHLGASQLAPTGEGDELSGQEDAFRSWAVQTFRAA
humnos2   YAVFGLGSSMYPRFCAFAHDIDQKLSHLGASQLTPMGEGDELSGQEDAFRSWAVQTFKAA
ccnos2    YCVFGLGSRMYPHFCAFAHAVDDRFAALGAIRVSATGEGDELNGQEEAFSAWACTVFKDA
lymnos    FAVFGLGSKAYPYYAAYGKYIYLMLQELGAERLVNYCAGDALYGQEQSFRAWSEEVFKAS
          :..*.***  ** :.*:.  :   :  **. ::     ** : ** ::   *:  .*  :

dnos      CETFCLDPEE--SLSDASLALQNDSLTVNTVRLVPSANK--GSLDSSLSKYHNKKVHCCK
anonos    CETFCLDPEE--TLSDAAFALQS-ELSENTVRYAPVAEY--ESLDRALSKFHNKKSMECS
rpnos     CDTFCLDSDE--TFLEATQMLHSEAVTASTVRFVESAT---QDLCKALSHLHNKKVWKCP
orynos1   CDVFCVGDDVNIEKANNSLISNDRSWKRNKFRLTYVAEA--PGLTQGLSSVHKKRVSAAR
humnos1   CDVFCVGDDVNIEKANNSLISNDRSWKRNKFRLTFVAEA--PELTQGLSNVHKKRVSAAR 1001
ratnos1   CDVFCVGDDVNIEKANNSLISNDRSWKRNKFRLTYVAEA--PDLTQGLSNVHKKRVSAAR
musnos1   CDVFCVGDDVNIEKANNSLISNDRSWKRNKFRLTYVAEA--PELTQGLSNVHKKRVSAAR
xennos1   CDVFCVGDDVSIEKANNSLISNDRSWKRSKYRISFVAEA--PELTQALYSIHKKKVYGAR
fugnos1   CDVFCVGDDVNIEKANNSLISNDRSWKKNKFRLTYTAEA--PSLTKALYGVHKKKVHAAK
susnos3   CETFCVGEDA--KAAARDIFSPKRSWKRQRYRLSAQVEG--LQLLPGLVHVHRRKMFQAT
bovnos3   CETFCVGEEA--KAAAQDIFSPKRSWKRQRYRLSTQAEG--LQLLPGLIHVHRRKMFQAT
humnos3   CETFCVGEDA--KAAARDIFSPKRSWKRQRYRLSAQAEG--LQLLPGLIHVHRRKMFQAT
musnos3   CETFCVGEDA--KAAARDIFSPKRSWKRQRYRLSTQAES--LQLLPGLTHVHRRKMFQAT
musnos2   CETFDVRSKH--HIQIPKRFTSNATWEPQQYRLIQSPEP--LDLNRALSSIHAKNVFTMR
ratnos2   CETFDVRSKH--CIQIPKRYTSNATWEPEQYKLTQSPEP--LDLNKALSSIHAKNVFTMR
humnos2   CETFDVRGKQ--HIQIPKLYTSNVTWDPHHYRLVQDSQP--LDLSKALSSMHAKNVFTMR
ccnos2    CKEFNIQC------ELPGKEGMADSWDPQRHRVQNDSCT--VDRITALSALHSKAVVPMK
lymnos    CEAFCLDNR-------NDAPGPQTKGDCSKVRIVPVENCQEPDLCQVLRNIHGKEVMPLI
          *. * :                          :               *   * :
```

Figure A.3. continued.

```
dnos      AKAKPHNLTRLSEG-AKTTMLLEIC--APGLEYEPGDHVGIFPANRTELVDGLLNRLVGV
anonos    VKRNPINLHCEMNGTERSTILVEIM--AEGIDYEPGDHVGIFPANRKEIVDGIIERLTGV
rpnos     LLGK-RNLHGKGST--RATLLLEIER-NENISYQPGDHVGVLACNRKELVEGIISHLESA
orynos1   LLSR-QNLQSPKSSRSTIFVRLHTNG-SQELQYQPGDHLGVFPGNHEDLVNALIERLEDA
humnos1   LLSR-QNLQSPKSSRSTIFVRLHTNG-SQELQYQPGDHLGVFPGNHEDLVNALIERLEDA 1059
ratnos1   LLSR-QNLQSPKSSRSTIFVRLHTNG-NQELQYQPGDHLGVFPGNHEDLVNALIERLEDA
musnos1   LLSR-QNLQSPKSSRSTIFVRLHTNG-NQELQYQPGDHLGVFPGNHEDLVNALIERLEDA
xennos1   LLSR-QNLQSPKSSRSTIFLKLHSNG-HKELCYKPGDHLGVFPGNHEDLVNALIERLDDA
fugnos1   MLDS-QNLQSPKSNRSTILVRLDTNN-QDSLKYKPGDHLGIFPGNHEDLVSALIDKLEDA
susnos3   VLSV-ENLQSSKSTRATILVRLDTEG-QEGLQYQPGDHIGICPPNRTGLVEALLSRVEDP
bovnos3   VLSV-ENLQSSKSTRATILVRLDTAG-QEGLQYQPGDHIGICPPNRPGLVEALLSRVEDP
humnos3   IRSV-ENLQSSKSTRATILVRLDTGG-QEGLQYQPGDHIGVCPPNRPGLVEALLSRVEDP
musnos3   ILSV-ENLQSSKSTRATILVRLDTGG-QEGLQYQPGDHIGVCPPNRPGLVEALLSRVEDP
musnos2   LKSQ-QNLQSEKSSRTTLLVQLTFEG-SRGPSYLPGEHLGIFPGNQTALVQGILERVVDC
ratnos2   LKSL-QNLQSEKSSRTTLLVQLTFEG-SRGPSYLPGEHLGIFPGNQTALVQGILERVVDC
humnos2   LKSR-QNLQSPTSSRATILVELSCED-GQGLNYLPGEHLGVCPGNQPALVQGILERVVDG
ccnos2    LKRR-QNLQSPTSSRATILVELGMDGNTEPLNIVPGDHVGIFPGNSPELVAGILKHLANA
lymnos    LAER-IQLQAKDSDQQTILIKLDAHN-ATDLKYAPGDHVAIFPANSPEIVDAILVRLDTS
                :*    .      : :             **:*:.: . *   :* .:: ::

dnos      --DNPDEVLQLQLLKEKQTSNGIFKCWEPHDKIPPDTLRNLLARFFDLTTPPSRQLLTLL
anonos    --NDPDEMLQLQVLKEKQTQNGVYKSWEPHERLPVCTLRTLLTRFLDITTPPTRQLLTYL
rpnos     --IDPDKSVQLQILKENTTPDGIVRNWIPHDRLPTCSLRTMLTRFLDITTPPSPNLLQFF
orynos1   --PPANQMVKVELLEERNTALGVISNWKDEPRLPPCTVFQAFKYYLDITTPPTPLQLQQF
humnos1   --PPVNQMVKVELLEERNTALGVISNWTDELRLPPCTIFQAFKYYLDITTPPTPLQLQQF 1117
ratnos1   --PPANHVVKVEMLEERNTALGVISNWKDESRLPPCTIFQAFKYYLDITTPPTPLQLQQF
musnos1   --PPANHVVKVEMLEERNTALGVISNWKDESRLPPCTIFQAFKYYLDITTPPTPLQLQQF
xennos1   --PPVNQMVRVEMLEERNTALGVISNWTEEERIPPCTIFQAFKYFLDITTPPTPLLMQQI
fugnos1   --PPVNQIVKVEFLEERNTALGVISNWTNETRVPPCTINQAFQYFLDITTPPSPVLLQQF
susnos3   --TPPTESVGVEQLE-KGSPGGPPPSWVRDPRLPPYTLRQALTFFLDITSPPSPRLLRVL
bovnos3   --PPPTESVAVEQLE-KGSPGGPPPSWVRDPRLPPCTLRQALTFFLDITSPPSPRLLRLL
humnos3   --PAPTEPVAVEQLE-KGSPGGPPPGWVRDPRLPPCTLRQALTFFLDITSPPSPQLLRLL
musnos3   --PPSTEPVAVEQLE-KGSPGGPPPGWVRDPRLPPCTLRQALTYFLDITSPPSPRLLRLL
musnos2   --PTPHQTVCLEVLDES----GSY--WVKDKRLPPCSLSQALTYFLDITTPPTQLQLHKL
ratnos2   --SSPDQTVCLEVLDES----GSY--WVKDKRLPPCSLRQALTYFLDITTPPTQLQLHKL
humnos2   --PTPHQTVRLEALDES----GSY--WVSDKRLPPCSLSQALTYFLDITTPPTQLLLQKL
ccnos2    --PPINQSLRLEFLSACPD--GER--WQRVERIPPCSLAQALTYYLDVTTPPSQSLLRKL
lymnos    KGPSPDQVVKTEIST----QLGTNDTWRSH--LPICTSRTAFSFLLDVTTPPSQEILQVL
                . :  :          *    *     :*  :    :   :*:*:**:   :  :

dnos      AGFCEDTADKERLELLVNDSSAYEDWRHWRLPHLLDVLEEFPSCRPPAPLLLAQLTPLQP
anonos    ASCCGDKADEERLLMLANESSVYEDWRYWKLPHLLEVLEEFPSCRPPAAVFVAQLNALQP
rpnos     ASCATNSEDQEKLTELATDSAAYEDWRYWKYPNLLEVLEEFPSVRVLPALLIAQLTPLQP
orynos1   ASLASNEKEKQRLLVLSKGLQEYEEWKWGKNPTIVEVLEEFPSIQMPATLLLTQLSLLQP
humnos1   ASLATSEKEKQRLLVLSKGLQEYEEWKWGKNPTIVEVLEEFPSIQMPATLLLTQLSLLQP 1177
ratnos1   ASLATNEKEKQRLLVLSKGLQEYEEWKWGKNPTMVEVLEEFPSIQMPATLLLTQLSLLQP
musnos1   ASLATNEKEKQRLLVLSKGLQEYEEWKWGKNPTMVEVLEEFPSIQMPATLLLTQLSLLQP
xennos1   ASLATDEKEKKRLEILSKGLQEYEQWKWYKNPTIVEVLEEFPSIQMPSSLLLTQLPLLQP
fugnos1   AALATNEKEKRKLEVLSKGLQEYEEWKWYNNPTLVEVLEEFPSIQMPSTLLLSQLPLLQP
susnos3   STLAEEPSEQQELETLSQDPRRYEEWKWFRCPTLLEVLEQFPSVALPTPLLLTQLALLQP
bovnos3   STLAEEPSEQQELETLSQDPRRYEEWKWFRCPTLLEVLEQFPSVALPAPLLLTQLPLLQP
humnos3   STLAEEPREQQELEALSQDPRRYEEWKWFRCPTLLEVLEQFPSVALPAPLLLTQLPLLQP
musnos3   STLAEESSEQQELEALSQDPRRYEEWKWFSCPTLLEVLEQFPSVALPAPLILTQLPLLQP
musnos2   ARFATDETDRQRLEALCQPSE-YNDWKFSNNPTFLEVLEEFPSLHVPAAFLLSQLPILKP
ratnos2   ARFATEETHRQRLEALCQPSE-YNDWKFSNNPTFLEVLEEFPSLRVPAAFLLSQLPILKP
humnos2   AQVATEEPERQRLEALCQPSE-YSKWKFTNSPTFLEVLEEFPSLRVSAGFLLSQLPILKP
ccnos2    SKMAKQEDHRQRLLALATDFQVYATWKEFYKPTSLEVLEEFSSLELSADFLLSQLPLLKP
lymnos    ATQASSDMDKHKLEQLASNSEAYEKWRLDLSPNILEILDEFPSLKIPPSLLLTQLPLLQP
          :  . .  ....*  *      *  *:    *  :::*::*.*    . .:::**  *:*
```

Figure A.3. continued.

```
dnos      RFYSISSSPRRVSDEIHLTVAIVKYRCEDGQGDERYGVCSNYLSGLRADDELFMFVRSAL
anonos    RFYSISSSPRKYSNEIHLTVAIVTYRAEDGEGAEHYGVCSNYLANLQSDDKIYLFVRSAP
rpnos     RFYSISSAPSLYANQIHLTVAVVQYCTQDGKGPIHYGVASNYLYDVTIGDSIYLFTRSAP
orynos1   RYYSISSSPDMYPDEVHLTVAIVSYHTRDGEGPIHHGVCSSWLNRIPADEVVPCFVRGAP
humnos1   RYYSISSSPDMYPDEVHLTVAIVSYRTRDGEGPIHHGVCSSWLNRIQADELVPCFVRGAP 1237
ratnos1   RYYSISSSPDMYPDEVHLTVAIVSYHTRDGEGPVHHGVCSSWLNRIQADDVVPCFVRGAP
musnos1   RYYSISSSPDMYPDEVHLTVAIVSYHTRDGEGPVHHGVCSSWLNRIQADDVVPCFVRGAP
xennos1   RYYSISSSPEMYPEEVHLTVAVVSYRTRDGDGPIHHGVCSSWLNRIHPDEVVPCFVRGAP
fugnos1   RYYSISSSPDLHPGEIHLTVAVVSYRTRDGAGSIHHGVCSSWLSRIEKGEMVPCFVRSAP
susnos3   RYYSVSSAPSTYPGEIHPTVAVLAYRTQDGLGPLHYGVCSTWLGQLKPGDPVPCFIRAAP
bovnos3   RYYSVSSAPNAHPGEVHLTVAVLAYRTQDGLGPLHYGVCSTWLSQLKTGDPVPCFIRGAP
humnos3   RYYSVSSAPSTHPGEIHLTVAVLAYRTQDGLGPLHYGVCSTWLSQLKPGDPVPCFIRGAP
musnos3   RYYSVSSAPSAHPGEIHLTIAVLAYRTQDGLGPLHYGVCSTWMSQLKAGDPVPCFIRGAP
musnos2   RYYSISSSQDHTPSEVHLTVAVVTYRTRDGQGPLHHGVCSTWIRNLKPQDPVPCFVRSVS
ratnos2   RYYSISSSQDHTPSEVHLTVAVVTYRTRDGQGPLHHGVCSTWINNLKPEDPVPCFVRSVS
humnos2   RFYSISSSRDHTPTEIHLTVAVVTYHTRDGQGPLHHGVCSTWLNSLKPQDPVPCFVRNAS
ccnos2    RLYSISSSPDLHPQELHLTVAVVSYCTQEGKGPLHFGLCSTWLNTIKEGDMVPFFAHSSD
lymnos    RYYSISSSQQKNPNEVHATIAVVRFKTQDGDGPVHEGVCSSWLNRSPIGTVVPCFLRSAP
          * **:**:    . ::* *:*:: :  .:* *  : *:.*.::        :  * :

dnos      GFHLPSDRSRPIILIGPGTGIAPFRSFWQEFQVLRDLDPT--------------------
anonos    SFHMSKDRTKPVILIGPGTGIAPFRSFWQEWDHIKTEMVD--------------------
rpnos     NFHLPKSDTAPIIMVGPGTGIAPFRGFWQHRLAQRSLNGP--------------------
orynos1   SFRLPRNPQVPCILVGPGTAFAPFRSFWQQRQFDIQHKG--------------------
humnos1   SFHLPRNPQVPCILVGPGTGIAPFRSFWQQRQFDIQHKG-------------------- 1276
ratnos1   SFHLPRNPQVPCILVGPGTGIAPFRSFWQQRQFDIQHKG--------------------
musnos1   SFHLPRNPQVPCILVGPGTGIAPFRSFWQQRQFDIQHKG--------------------
xennos1   SFQMPEDPQVPCILIGPGTGIAPFRSFWQQRLYDMQHRG--------------------
fugnos1   SFQLPKNNQTPCILVGPGTGIAPFRSFWQQRLYDLEHNG--------------------
susnos3   SFRLPPDPSLPCILVGPGTGIAPFRGFWQERLHDIESKG--------------------
bovnos3   SFRLPPDPYVPCILVGPGTGIAPFRGFWQERLHDIESKG--------------------
humnos3   SFRLPPDPSLPCILVGPGTGIAPFRGFWQERLHDIESKG--------------------
musnos3   SFRLPPDPNLPCILVGPGTGIAPFRGFWQDRLHDIEIKG--------------------
musnos2   GFQLPEDPSQPCILIGPGTGIAPFRSFWQQRLHDSQHKG--------------------
ratnos2   GFQLPEDPSQPCILIGPGTGIAPFRSFWQQRLHDSQRRG--------------------
humnos2   GFHLPEDPSHPCILIGPGTGIAPFRSFWQQRLHDSQHKG--------------------
ccnos2    GFHLPSDPSAPCILVGVGSGIAPFRSFWQQQFHDMKKTG--------------------
lymnos    HFHLPEDPSLPIIMIGPGSGIAPFRSFWQQRLGEIENTMPSCENTMLSCETTIPSCENSM
           *::. .   * *::* *:.:****.***.

dnos      ------------------------------------------------------------
anonos    ------------------------------------------------------------
rpnos     ------------------------------------------------------------
orynos1   ------------------------------------------------------------
humnos1   ------------------------------------------------------------ 1276
ratnos1   ------------------------------------------------------------
musnos1   ------------------------------------------------------------
xennos1   ------------------------------------------------------------
fugnos1   ------------------------------------------------------------
susnos3   ------------------------------------------------------------
bovnos3   ------------------------------------------------------------
humnos3   ------------------------------------------------------------
musnos3   ------------------------------------------------------------
musnos2   ------------------------------------------------------------
ratnos2   ------------------------------------------------------------
humnos2   ------------------------------------------------------------
ccnos2    ------------------------------------------------------------
lymnos    PSCENTMPSCENTMPSCENTIPSCENTIPSCENTMPSCENTIPSWERTMQPCQIILPSQT
```

Figure A.3. continued.

```
dnos      -AKLPKMWLFFGCRNRDVD-LYAEEKAELQKDQILDRVFLALSREQAIPKTYVQDLIEQE
anonos    -CKIPKVWLFFGCRTKNVD-LYRDEKEEMVQHGVLDRVFLALSREENIPKTYVQDLALKE
rpnos     -GKFGKMSLFFGCRLRNLD-LYQEEKESMLKEGILSKVFLALSREPSIPKTYVQDLLRVE
orynos1   -MSPCPMVLVFGCRQSKIDHIYREEALQAKNKGVFRELYTAYSREPDKPKKYVQDILQEQ
humnos1   -MNPCPMVLVFGCRQSKIDHIYREETLQAKNKGVFRELYTAYSREPDKPKKYVQDILQEQ 1325
ratnos1   -MNPCPMVLVFGCRQSKIDHIYREETLQAKNKGVFRELYTVYSREPDRPKKYVQDVLQEQ
musnos1   -MNPCPMVLVFGCRQSKIDHIYREETLQAKNKGVFRELYTAYSREPDRPKKYVQDVLQEQ
xennos1   -LKPCPMILVFGCRESKIDHIYKEETMHAKNKGVFKELFTAYSREPNKPKKYVQDVLKEQ
fugnos1   -IESCPMILVFGCRQSEIDHIYNEETIQAKNKNVFKELYTAYSREPGKPKKYVQDALREQ
susnos3   -LQPAPMTLVFGCRCSQLDHLYRDEVQDAQQRGVFGRVLTAFSREPDSPKTYVQDILRTE
bovnos3   -LQPAPMTLVFGCRCSQLDHLYRDEVQDAQERGVFGRVLTAFSREPDSPKTYVQDILRTE
humnos3   -LQPTPMTLVFGCRCSQLDHLYRDEVQNAQQRGVFGRVLTAFSREPDNPKTYVQDILRTE
musnos3   -LQPAPMTLVFGCRCSQLDHLYRDEVLDAQQRGVFGQVLTAFSRDPGSPKTYVQDLLRTE
musnos2   -LKGGRMSLVFGCRHPEEDHLYQEEMQEMVRKRVLFQVHTGYSRLPGKPKVYVQDILQKQ
ratnos2   -LKGGRMTLVFGCRHPEEDHLYQEEMQEMVRKGVLFQVHTGYSRLPGKPKVYVQDILQKE
humnos2   -VRGGRMTLVFGCRRPDEDHIYQEEMLEMAQKGVLHAVHTAYSRLPGKPKVYVQDILRQQ
ccnos2    -LKGNPVTLVFGCRGSDTDHLYKEETLDMRDNGTLSSITTAYSRQTGQPKVYVQDILREQ
lymnos    KKHFGEMVLYTGCRTAKHM-IYAAELEEMKRLGVLSNYHVALSREAALPKMYVQDIIIKN
                 : *  ***  .   :*  *         :       **    ** ****    :

dnos      FDS-LYQLIVQERGHIYVCGDVTMAEHVYQTIRKCIAGKEQKSEAEVETFLLTLRDESRY
anonos    AES-ISELIMQEKGHIYVCGDVTMAEHVYQTLRKILATREKRTETEMEKYMLTLRDENRY
rpnos     CKS-VYIQIVQEGGHFYVCGDCTMAEHVFRTLRQIIQDQGNMTDHQVDNFMLAMRDENRY
orynos1   LAEQVYRALKEQGGHIYVCGDVTMAADVLKAVQRIMAQQGKLSAEDAGVFISRLRDDNRY
humnos1   LAESVYRALKEQGGHIYVCGDVTMAADVLKAIQRIMTQQGKLSAEDAGVFISRMRDDNRY 1385
ratnos1   LAESVYRALKEQGGHIYVCGDVTMAADVLKAIQRIMTQQGKLSEEDAGVFISRLRDDNRY
musnos1   LAESVYRALKEQGGHIYVCGDVTMAADVLKAIQRIMTQQGKLSEEDAGVFISRLRDDNRY
xennos1   LSEVTYKALKDQGGHIYICGDVNMAGDVLKSLQHVVKESGNLTIEEAGAFISKLRDDNRY
fugnos1   LSERVYQCLREEGGHIYVCGDVTMAGDVLKNVQQIIKQEGNMSLEEAGLFISKLRDENRY
susnos3   LAAEVHRVLCLERGHMFVCGDVTMATSVLQTVQRILATEGNMELDEAGDVIGVLRDQQRY
bovnos3   LAAEVHRVLCLERGHMFVCGDVTMATSVLQTVQRILATEGDMELDEAGDVIGVLRDQQRY
humnos3   LAAEVHRVLCLERGHMFVCGDVTMATNVLQTVQRILATEGDMELDEAGDVIGVLRDQQRY
musnos3   LAAEVHRVLCLEQGHMFVCGDVTMATSVLQTVQRILATEGGMELDEAGDVIGVLRDQQRY
musnos2   LANEVLSVLHGEQGHLYICGDVRMARDVATTLKKLVATKLNLSEEQVEDYFFQLKSQKRY
ratnos2   LADEVFSVLHGEQGHLYVCGDVRMARDVATTLKKLVAAKLNLSEEQVEDYFFQLKSQKRY
humnos2   LASEVLRVLHKEPGHLYVCGDVRMARDVAHTLKQLVAAKLKLNEEQVEDYFFQLKSQKRY
ccnos2    LSDKVFEVLHHNPGHLYVCGGMNMANDVAATIKEILVSRLGITLAQAEEYLSRLKNEKRY
lymnos    AAA-VYEIVMKKGGHFYVSGDVSMAHDVTRALELVLCQQGGR---EASQQVMSLRDENLF
                 :  : **:::.*.  **  *   :.  :          :    .  ::.:. :

dnos      HEDIFGITLRTAEIHTKSRATARIRMASQP-----------------------
anonos    HEDIFGITLRTAEIHNKSRATARIRMASQP-----------------------
rpnos     HEDIFGITLRTAEVHNRSRESARIRMASQSQP---------------------
orynos1   HEDIFGVTLRTYEVTNRLRSESIAFIEESKKDTDEVFSS---------------
humnos1   HEDIFGVTLRTYEVTNRLRSESIAFIEESKKDTDEVFSSLDPLAQPAASLARDR 1439
ratnos1   HEDIFGVTLRTYEVTNRLRSESIAFIEESKKDADEVFSS---------------
musnos1   HEDIFGVTLRTYEVTNRLRSESIAFIEESKKDTDEVFSS---------------
xennos1   HEDIFGVTLRTYEVTNRLRSESIAFIEESKKDSDEVFCS---------------
fugnos1   HEDIFGVTLRTYEVTNRIRSESIAYIEENKKDSDEVFCS---------------
susnos3   HEDIFGLTLRTQEVTSRIRTQSFSLQERHLRGAVPWTFDPPGPDTPGP------
bovnos3   HEDIFGLTLRTQEVTSRIRTQSFSLQERHLRGAVPWAFDPPGPDTPGP------
humnos3   HEDIFGLTLRTQEVTSRIRTQSFSLQERQLRGAVPWAFDPPGSDTNSP------
musnos3   HEDIFGLTLRTQEVTSRIRTQSFSLQERQLRGAVPWSFDPPGPEIPGS------
musnos2   HEDIFGAVFSYGAKKGSALEEPKATRL---------------------------
ratnos2   HEDIFGAVFSYGVKKGNALEEPKGTRL---------------------------
humnos2   HEDIFGAVFPYEAKKDRVAVQPSSLEMSAL-----------------------
ccnos2    HEDIFGS-----------------------------------------------
lymnos    HEDIFGSFVRKAG-GQRSEDE---------------------------------
          ******
```

Figure A.3. continued.

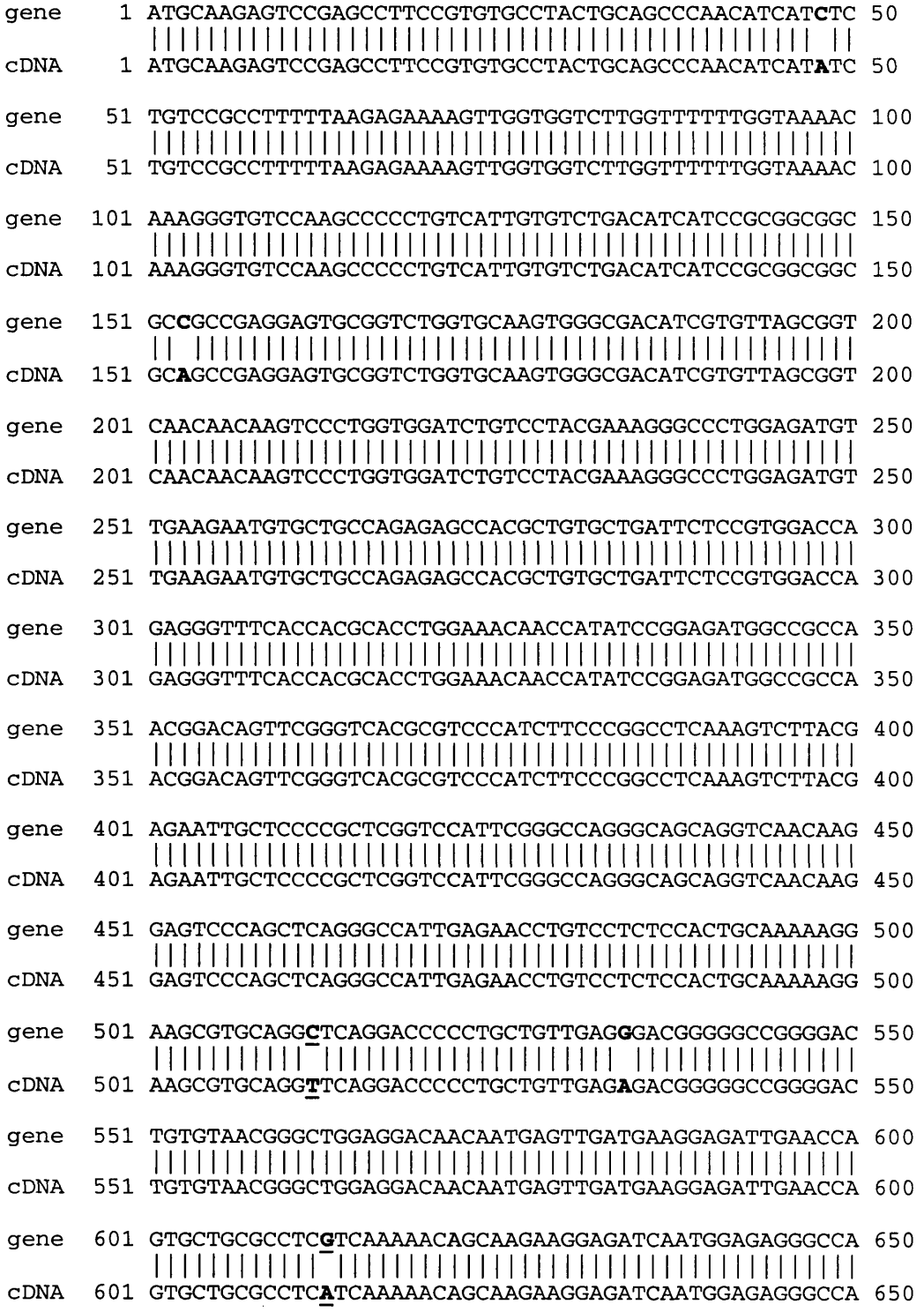


Figure A.4. Alignment of the *Fugu* NOS1 gene predicted coding sequence and the *Fugu* NOS1 cDNA (*f*NOS) sequence. Nucleotide difference are highlighted in bold and those resulting in amino acid changes are also underlined.

```
gene   651 GAGGCACGTAGGGAGAAGAGATGCTGAGATTCAAGTGACCTGGGGCGCTG 700
           ||||||  || ||||||||||||||||||||||||||||||||||||||
cDNA   651 GAGGCATGTGGGGAGAAGAGATGCTGAGATTCAAGTGACCTGGGGCGCTG 700

gene   701 GTGTAGGAATCGACACATCTCTGCAGTTGGACTCCTGTAAGAACAAAATG 750
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA   701 GTGTAGGAATCGACACATCTCTGCAGTTGGACTCCTGTAAGAACAAAATG 750

gene   751 CCGGAAAAAGAGCCCGGTGTGCCGCAGAACGCCGACAATGACAAGCCTCC 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA   751 CCGGAAAAAGAGCCCGGTGTGCCGCAGAACGCCGACAATGACAAGCCTCC 800

gene   801 TGCCGAGGCCAGGACCTCACCTACTAAATCCCTGCAGAACGGAAGCCCCT 850
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA   801 TGCCGAGGCCAGGACCTCACCTACTAAATCCCTGCAGAACGGAAGCCCCT 850

gene   851 CCAAATGCCCTCGCTTCCTGAAGATCAAGAACTGGGAGACCGGCGCCATC 900
           |||||||||||||||||||||||||||||||||||||||||||||| ||
cDNA   851 CCAAATGCCCTCGCTTCCTGAAGATCAAGAACTGGGAGACCGGCGCCGTC 900

gene   901 CAAAATGACACATTACACAACAGCTCCACCAAGACGCCAATGTGCCCCGA 950
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA   901 CAAAATGACACATTACACAACAGCTCCACCAAGACGCCAATGTGCCCCGA 950

gene   951 GAACGTGTGCTACGGCTCCCTTATGATGCCCAACCTGCACGCCCGCAAAC 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA   951 GAACGTGTGCTACGGCTCCCTTATGATGCCCAACCTGCACGCCCGCAAAC 1000

gene  1001 CGGAGGAGGTCAGGAGCAAAGAGGAGCTTCTGAAACTGGCCACCGACTTC 1050
           |||||||||||||||||||||||||||||||||||||||||||| ||||||
cDNA  1001 CGGAGGAGGTCAGGAGCAAAGAGGAGCTTCTGAAACTGGCCACTGACTTC 1050

gene  1051 ATTGACCAGTATTACACCTCCATCAAAAGGTACGGCTCCAAGGCCCATAC 1100
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA  1051 ATTGACCAGTATTACACCTCCATCAAAAGGTACGGCTCCAAGGCCCATAC 1100

gene  1101 GGACAGACTGGAGGAGGTGACCAAGGAGATTGAAGCGACTGGAACCTACC 1150
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA  1101 GGACAGACTGGAGGAGGTGACCAAGGAGATTGAAGCGACTGGAACCTACC 1150

gene  1151 AGCTGAAAGACACAGAACTGATTTATGGAGCCAAACATGCCTGGAGGAAT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA  1151 AGCTGAAAGACACAGAACTGATTTATGGAGCCAAACATGCCTGGAGGAAT 1200

gene  1201 GCTGCCCGTTGTGTGGGAAGAATCCAGTGGTCTAAACTACAGGTTTTTGA 1250
           |||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA  1201 GCTGCCCGTTGTGTGGGAAGAATCCAGTGGTCTAAACTACGGGTTTTTGA 1250

gene  1251 TGCCAGAGACTGCACAACAGCTCATGGAATGTACAACTACATCTGTAACC 1300
           ||| ||||||||||||||||||||||||||||||||||||||||||||||
cDNA  1251 TGCTAGAGACTGCACAACAGCTCATGGAATGTACAACTACATCTGTAACC 1300

gene  1301 ACATCAAGTACGCTACCAACAAAGGGAATCTGAGGTCAGCTATCACCATA 1350
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA  1301 ACATCAAGTACGCTACCAACAAAGGGAATCTGAGGTCAGCTATCACCATA 1350

gene  1351 TTTCCTCCGAGGACAGATGGCAAACATGACTTTCGAGTGTGGAACAGTCA 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA  1351 TTTCCTCCGAGGACAGATGGCAAACATGACTTTCGAGTGTGGAACAGTCA 1400
```

Figure A.4. continued.

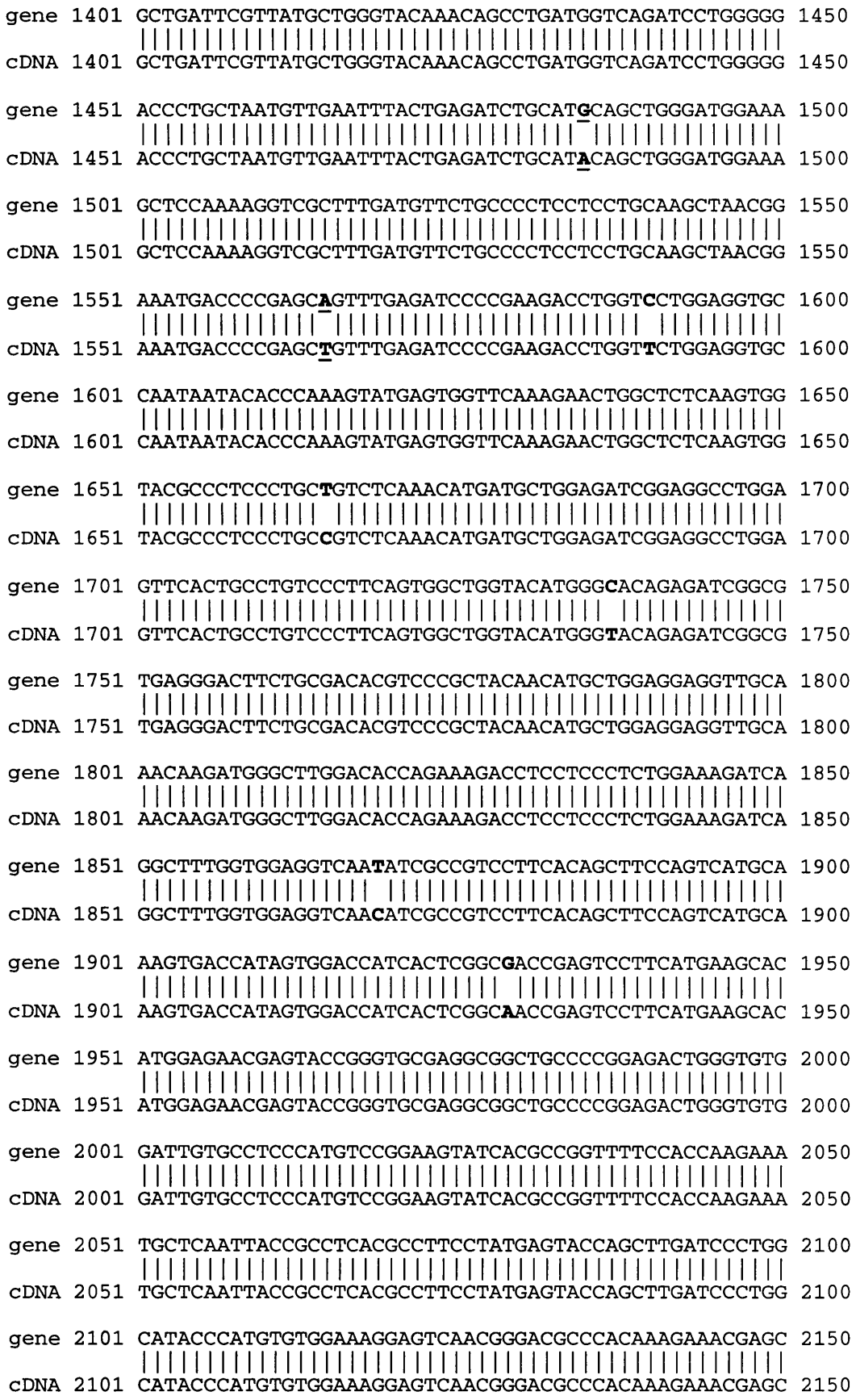


Figure A.4. continued.

```
gene 2151 CATTGGATTCAAAAAGCTTGCCAAAGCGGTGAAGTTTTCAGCCAAACTCA 2200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2151 CATTGGATTCAAAAAGCTTGCCAAAGCGGTGAAGTTTTCAGCCAAACTCA 2200

gene 2201 TGGGTCACGCAATGGCCAAGAGAGTCAAAGCCACCATATTGTTCGCCACC 2250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2201 TGGGTCACGCAATGGCCAAGAGAGTCAAAGCCACCATATTGTTCGCCACC 2250

gene 2251 GAGACGGGAAAATCGCAGGATTACGCCAAAACTCTCTGCGAAATCTTCAA 2300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2251 GAGACGGGAAAATCGCAGGATTACGCCAAAACTCTCTGCGAAATCTTCAA 2300

gene 2301 GCACGCATTCGACCCAAAGGTCATGTCTATGGACGATTATGATGTGGTGG 2350
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2301 GCACGCATTCGACCCAAAGGTCATGTCTATGGACGATTATGATGTGGTGG 2350

gene 2351 ATCTGGAGCATGAGACGCTGGTGTTAGTGGTGACCAGCACGTTTGGCAAC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2351 ATCTGGAGCATGAGACGCTGGTGTTAGTGGTGACCAGCACGTTTGGCAAC 2400

gene 2401 GGCGACCCCCCTGAGAACGGAGAGAAATTTGGAGCCGCCTTAATGGAGAT 2450
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2401 GGCGACCCCCCTGAGAACGGAGAGAAATTTGGAGCCGCCTTAATGGAGAT 2450

gene 2451 GCGGGAGCCGACGTCCAACACAGAAGACAGAAAGAGCTACAAGGTCCGTT 2500
          |||||||||||| |||||| ||||||||||||||||||||||||||||||
cDNA 2451 GCGGGAGCCGACATCCAACGCAGAAGACAGAAAGAGCTACAAGGTCCGTT 2500

gene 2501 TCAACAGCGTGTCCTCCCACTCCGACACTCGCAAGTCCTCAAGTGACGAA 2550
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2501 TCAACAGCGTGTCCTCCCACTCCGACACTCGCAAGTCCTCAAGTGACGAA 2550

gene 2551 CCGGACGCCAAGATTCACTTTGAAAGCACCGGACCTCTGGCCAATGTCAG 2600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2551 CCGGACGCCAAGATTCACTTTGAAAGCACCGGACCTCTGGCCAATGTCAG 2600

gene 2601 GTTCTCTGTGTTTGGCCTTGGATCCAGAGCCTACCCACACTTCTGCGCCT 2650
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2601 GTTCTCTGTGTTTGGCCTTGGATCCAGAGCCTACCCACACTTCTGCGCCT 2650

gene 2651 TTGCCCACGCCGTGGACACGCTGTTTGAAGAGCTCGGGGGGGGAGCGCATC 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2651 TTGCCCACGCCGTGGACACGCTGTTTGAAGAGCTCGGGGGGGGAGCGCATC 2700

gene 2701 CTACGCATGGGAGAAGGGGATGAGCTGTGTGGACAAGAGGAGGCTTTCAG 2750
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2701 CTACGCATGGGAGAAGGGGATGAGCTGTGTGGACAAGAGGAGGCTTTCAG 2750

gene 2751 AACCTGGGCGAAAAAGGTTTTTAAGGCCGCCTGCGACGTGTTCTGTGTCG 2800
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2751 AACCTGGGCGAAAAAGGTTTTTAAGGCCGCCTGCGACGTGTTCTGTGTCG 2800

gene 2801 GCGATGACGTGAACATCGAGAAGGCCAACAACTCCCTGATCAGCAACGAC 2850
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2801 GCGATGACGTGAACATCGAGAAGGCCAACAACTCCCTGATCAGCAACGAC 2850

gene 2851 CGCAGCTGGAAGAAAAACAAGTTCCGTCTGACATACACGGCCGAGGCGCC 2900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2851 CGCAGCTGGAAGAAAAACAAGTTCCGTCTGACATACACGGCCGAGGCGCC 2900
```

Figure A.4. continued.

```
gene 2901 GAGCCTCACAAAAGCTTTGTACGGCGTGCACAAGAAGAAGGTCCATGCAG 2950
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2901 GAGCCTCACAAAAGCTTTGTACGGCGTGCACAAGAAGAAGGTCCATGCAG 2950

gene 2951 CAAAGATGCTCGATTCTCAGAATTTACAGAGTCCGAAATCCAATCGTTCC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 2951 CAAAGATGCTCGATTCTCAGAATTTACAGAGTCCGAAATCCAATCGTTCC 3000

gene 3001 ACCATTCTCGTACGGCTGGACACAAATAACCAAGACAGCCTGAAATACAA 3050
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3001 ACCATTCTCGTACGGCTGGACACAAATAACCAAGACAGCCTGAAATACAA 3050

gene 3051 GCCAGGAGACCATCTGGGCATCTTCCCTGGCAACCACGAGGACCTGGTGT 3100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3051 GCCAGGAGACCATCTGGGCATCTTCCCTGGCAACCACGAGGACCTGGTGT 3100

gene 3101 CGGCTCTCATAGATAAGCTGGAGGATGCGCCGCCTGTCAATCAGATTGTA 3150
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3101 CGGCTCTCATAGATAAGCTGGAGGATGCGCCGCCTGTCAATCAGATTGTA 3150

gene 3151 AAAGTGGAGTTCTTAGAGGAGAGGAACACTGCCCTGGGTGTGATAAGTAA 3200
          |||||||||||||||||||| ||||||||||||||||||||||||||||
cDNA 3151 AAAGTGGAGTTCTTAGAGGAGAGAAACACTGCCCTGGGTGTGATAAGTAA 3200

gene 3201 CTGGACCAATGAGACTCGGGTCCCTCCCTGCACCATCAACCAGGCCTTCC 3250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3201 CTGGACCAATGAGACTCGGGTCCCTCCCTGCACCATCAACCAGGCCTTCC 3250

gene 3251 AGTACTTCCTGGACATCACCACCCCGCCCAGCCCCGTACTGCTGCAGCAG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3251 AGTACTTCCTGGACATCACCACCCCGCCCAGCCCCGTACTGCTGCAGCAG 3300

gene 3301 TTCGCTGCTCTGGCCACTAATGAGAAAGAGAAACGGAAACTCGAGGTCCT 3350
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3301 TTCGCTGCTCTGGCCACTAATGAGAAAGAGAAACGGAAACTCGAGGTCCT 3350

gene 3351 CAGCAAGGGCTTGCAGGAGTATGAGGAGTGGAAGTGGTACAACAACCCCA 3400
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3351 CAGCAAGGGCTTGCAGGAGTATGAGGAGTGGAAGTGGTACAACAACCCCA 3400

gene 3401 CCCTTGTGGAGGTACTAGAGGAATTCCCCTCCATCCAGATGCCTTCTACG 3450
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3401 CCCTTGTGGAGGTACTAGAGGAATTCCCCTCCATCCAGATGCCTTCTACG 3450

gene 3451 CTGCTGCTCTCCCAGCTTCCCCTGCTGCAGCCTCGTTACTACTCCATCAG 3500
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3451 CTGCTGCTCTCCCAGCTTCCCCTGCTGCAGCCTCGTTACTACTCCATCAG 3500

gene 3501 CTCCTCTCCAGACCTGCACCCGGGAGAGATCCACCTCACGGTTGCTGTGG 3550
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3501 CTCCTCTCCAGACCTGCACCCGGGAGAGATCCACCTCACGGTTGCTGTGG 3550

gene 3551 TCTCCTACCGTACCAGAGATGGAGCAGGGTCGATCCACCATGGAGTGTGT 3600
          ||||||||||||||| ||||||||||||||||||||||||||||||||||
cDNA 3551 TCTCCTACCGTACCCGAGATGGAGCAGGGTCGATCCACCATGGAGTGTGT 3600

gene 3601 TCGTCGTGGCTCAGCAGGATAGAGAAGGGGGAGATGGTGCCGTGTTTTGT 3650
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3601 TCGTCGTGGCTCAGCAGGATAGAGAAGGGGGAGATGGTGCCGTGTTTTGT 3650
```

Figure A.4. continued

```
gene 3651 CCGAAGTGCTCCATCCTTCCAGCTTCCCAAAAACAACCAAACACCTTGCA 3700
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3651 CCGAAGTGCTCCATCCTTCCAGCTTCCCAAAAACAACCAAACACCTTGCA 3700

gene 3701 TCCTGGTGGGTCCGGGAACCGGAATCGCCCCATTCCGGAGCTTTTGGCAA 3750
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3701 TCCTGGTGGGTCCGGGAACCGGAATCGCCCCATTCCGGAGCTTTTGGCAA 3750

gene 3751 CAGCGACTGTATGACCTTGAACACAACGGCATCGAGTCATGCCCAATGAT 3800
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3751 CAGCGACTGTATGACCTTGAACACAACGGCATCGAGTCATGCCCAATGAT 3800

gene 3801 CCTGGTGTTTGGCTGTCGACAGTCTGAGATTGACCACATCTACAATGAGG 3850
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3801 CCTGGTGTTTGGCTGTCGACAGTCTGAGATTGACCACATCTACAAAGAGG 3850

gene 3851 AGACCATCCAAGCCAAGAACAAGAACGTGTTCAAGGAGCTGTACACGGCC 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3851 AGACCATCCAAGCCAAGAACAAGAACGTGTTCAAGGGGCTGTACACGGCC 3900

gene 3901 TATTCCAGAGAGCCCGGCAAACCAAAGAAATATGTGCAGGATGCACTGCG 3950
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3901 TATTCCAGAGAGCCCGGCAAACCAAAGAAATATGTGCAGGATGCACTGCG 3950

gene 3951 TGAGCAGCTGTCGGAGCGGGTGTACCAGTGCCTGAGGGAGGAGGGAGGAC 4000
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 3951 TGAGCAGCTGTCGGAGCGGGTGTACCAGTGCCTGAGGGAGGAGGGAGGAC 4000

gene 4001 ACATCTACGTGTGCGGGGATGTTACGATGGCCGGGGATGTTCTCAAGAAC 4050
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 4001 ACATCTACGTGTGCGGGGATGTTACGATGGCGGGGGATGTTCTCAAGAAC 4050

gene 4051 GTCCAGCAGATCATCAAGCAAGAGGGCAACATGAGCCTGGAGGAAGCCGG 4100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 4051 GTCCAGCAGATCATCAAACAAGAGGGCAACATGAGCCTGGAGGAAGCCGG 4100

gene 4101 CTTGTTCATCAGCAAGCTTCGGGATGAGAACCGCTACCATGAGGACATCT 4150
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 4101 CTTGTTCATCAGCAAGCTTCGGGATGAGAACCGCTACCATGAGGACATCT 4150

gene 4151 TTGGAGTCACCCTGCGCACCTACGAGGTCACCAACAGGATTCGGTCAGAG 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 4151 TTGGAGTCACCCTGCGCACCTACGAGGTCACCAACAGGATTCGGTCAGAG 4200

gene 4201 TCCATTGCCTACATCGAAGAGAATAAAAAGGACTCCGATGAGGTGTTCTG 4250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
cDNA 4201 TCCATTGCCTACATCGAAGAGAATAAAAAGGATTCCGATGAGGTGTTCTG 4250

gene 4251 CTCATAG 4257
          |||||||
cDNA 4251 CTCATAG 4257
```

Figure A.4. continued.

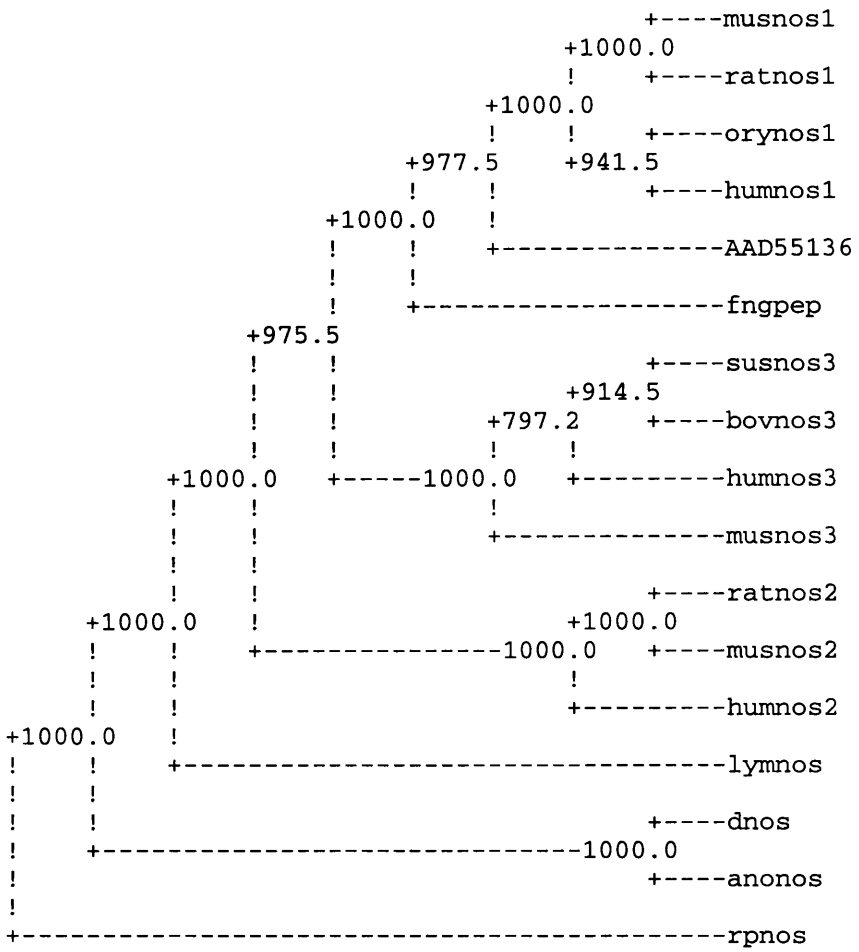


Figure A5. Consensus tree. The numbers at the forks indicate the number of times the group consisting of the species which are to the right of that fork occurred among the trees, out of 1000.00 trees.