Durham E-Theses

Gene Expression in Neuroblastoma

Said, Sohair S.S.

How to cite:

Said, Sohair S.S. (1996) *Gene Expression in Neuroblastoma*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/5394/

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

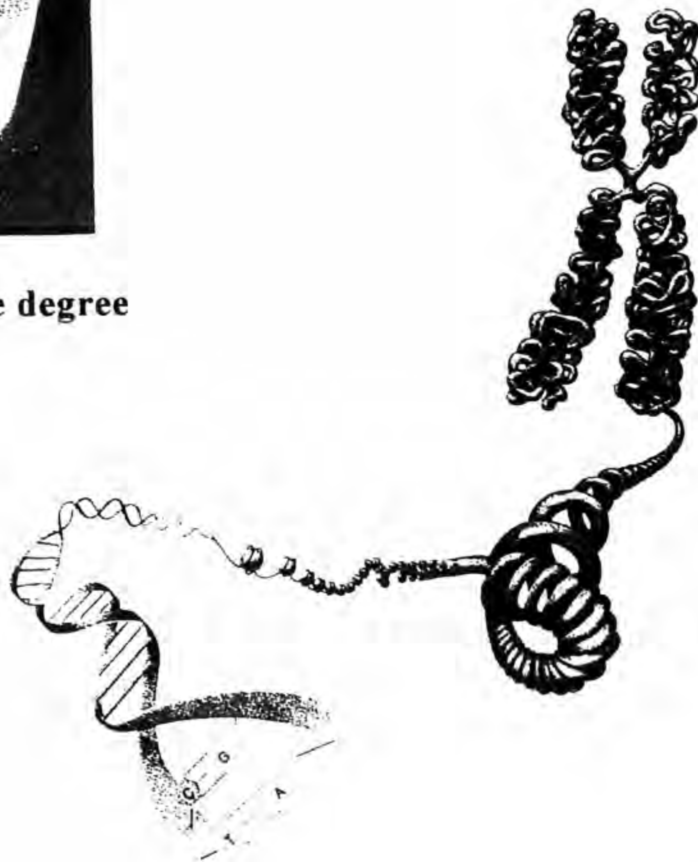
Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

The copyright of this thesis rests with the author. No quotation from it should be published without the written consent of the author and information derived from it should be acknowledged.

University of Durham

Gene Expression in Neuroblastoma

By
Sohair S. S. SAID B.Sc.



A thesis submitted for the degree of Master of science

Department of Biological Sciences

April 1996

- 6 OCT 1997

Thesis
1996/
SAI

قال الله تعالى :

وَمَآ أُوتِيتُم مِّنَ ٱلْعِلْمِ إِلَّا قَلِيلًا

سورة الإسراء : الآية ٨٥

Allah the Exalted said

"and of knowledge you (mankind) have been given but little."

Qur'an , 17:85

I. Abstract

The aim of this project is to identify potential target genes (NTG's) which may be regulated by N-*myc* in the cell line SK-N-SH which has no amplification of the oncogene N-*myc*.

N-*myc* target genes were isolated using a modification of the protocol developed to identify p53 binding sites; this was done by generating a NTG library by isolating genomic DNA from SK-N-SH cells, digesting to completion with *Sau3A*, and mixing with nuclear extract from the same cell line. N-myc antibody was added and DNA-protein complexes immunoprecipitated to isolate DNA with N-Myc bound. DNA recovered from these precipitates was then cloned into pBluescript SK+. 400 clones were isolated, 4 of which were partially sequenced and compared to sequence data bases.

The NTG clones showed substantial sequence identity to a number of potentially interesting genes, among them several implicated in tumour suppression, and the expression of cell surface markers. Their significance for neuroblastoma is discussed.

II. Acknowledgement

I would like to give special thanks to my supervisor Dr. CH. Shaw for his constant support, encouragement and helpful advice.

Particular thanks to my research colleague Dr. D. Pearson for her assistance and helpful discussion through out the work.
Dr. B. Teakin , Miss V. Parker for their kind help.

I wish to express my deepest gratitude to my husband and parents for their constant love, encouragement and moral support which has made this research possible.

Thanks are also due to Durham Training and Enterprise council for their financial support.

II. Contents

I. Abstract 1
II. Acknowledgement 2
III. Contents 3
IV. List of figures 5
V. List of tables 6
1. Introduction 7
1.1. Description 8
1.2. Stage explanation 8
1.3. Treatment options 8
1.4. Molecular characteristics of neuroblastoma cells 9
1.5. Summary 14
1.6. Aim of research 15
2. Materials and Methods 16
2.1. Materials 16
2.1.1. NTG library 16
2.1.2. Colony blots 21
2.1.3. Filter hybridization 22
2.1.4. Protein analysis 22
2.2. Methods 25
2.2.1. Tissue culture 25
2.2.1.1. Recovering SK-N-SH cells from liquid Nitrogen .. 25
2.2.1.2. Trypsinisation of cells for passaging 25
2.2.1.3. Preservation of cells in liquid Nitrogen 25
2.2.2. DNA Manipulation 26
2.2.2.1. Complete sure cells 26
2.2.2.2. Competent DH5 α cells 26
2.2.2.3. Isolation of genomic DNA 26
2.2.2.4. Nuclear protein extraction 28
2.2.2.5. pBluescript SK+ 29
2.2.2.6. Preparation of N-myc binding library 31
2.2.2.7. Isolation of plasmid DNA 35
2.2.2.8. *SacI / EcoRI* double digest 35
2.2.2.9. Insert size calculation 36
2.2.2.10. Colony blots 36
2.2.2.11. Colony hybridization 37

2.2.2.12. Glycerol stocks 40
2.2.2.13. Protein analysis 40

3. Results 42
3.1. NTG clones and homology search 42
3.2. Protein analysis 76
3.3. Colony hybridization 78
4. Discussion 79
4.1. Conclusion 84
5. References 85
5.1. Journals 85
5.2. Other bibliographic sources 86
5.3. Books 87
6. Appendices 88

V.I. List of figures

1.1. The development of cancer 7
1.2. Catecholamine Metabolism 9
1.3. Functional domains of N-*myc* 11
1.4. Speculative model for interaction and regulation of BR-HLH-Zip proteins in neuroblastoma 12
2.1. DNA extraction (5μl) on 0.8% Agarose gel 27
2.2. *Sau3A* cut genomic DNA (2μl) on 0.8% Agarose gel 28
2.3. BamH1 pBluescript (2μl) on 0.8% Agarose gel 30
2.4. CIP BamH1 pBluescript (2μl) on 0.8% Agarose gel 31
2.5. Stage explanation of preparation of N-myc binding library 31
2.6. Plasmid DNA for CIP transformants (2μl) on 0.8% Agarose gel 34
2.7. NTG-21 37
2.8. NTG-21 *SacI / EcoRI* digest (5μl) on 0.8 Agarose gel 38
2.9. *Sau3A* cut NTG-21 (5μl) on 0.8% Agarose gel 39
3.1. A number of plasmid DNA digested with *EcoRI/SacI* (2μl) on 0.8% Agarose gel 42
3.2. Pure plasmid for +13, +5, +8, +33 (1μl) on 1.2% Agarose gel 43
3.3. Homology search of NTG+5 44
3.4. Restriction map & DNA sequence of NTG+5 47
3.5. NTG+5 sequenced regions 48
3.6. NTG+5 *SacI*, *EcoRI*, *SacI/EcoRI* digest 2μl on 1.2% Agarose gel 48
3.7. *SacI/EcoRI* digestion of NTG+5 49
3.8. 10 different digests of NTG+5 (2μl) on 1.2% Agarose gel 49
3.9. *HindII* digestion of NTG+5 50
3.10. *PstI* digestion of NTG+5 50
3.11. Sequencing strategy for NTG+5 sub-clones 51
3.12. *SacI/PstI* NTG+5 digest (100μl) on 1.2% gel 51
3.13. *PstI* NTG+5 digest & *SacI/PstI* cut pBluescript (100μl) 51
3.14. *SacI/PstI* cut pBluescript 52
3.15. Homology search for 800bp *SacI/PstI* sub-clone of NTG+5 52
3.16. Restriction map & DNA sequence for 800bp *SacI/PstI* sub-clone NTG+5 53
3.17. *PstI* cut pBluescript 54

3.18. Possible recombinants 55
3.19. 1.4Kb sub-clone (2μl) on 0.8% Agarose gel 55
3.20. Homology search for 1.4Kb *PstI* sub-clones of NTG+5 56
3.21. Restriction map & DNA sequence for 1.4Kb *PstI* sub-clone NTG+5 58
3.22. Homology search 59
3.23. Restriction map & DNA sequence for 1.4Kb *PstI* sub-clone NTG+5 62
3.24. Sequenced regions of NTG+8 62
3.25. *EcoRI/SacI* digestion of NTG+8 63
3.26. 10 different digest of NTG+8 (2μl) on 1.2% Agarose gel 63
3.27. Fragments of NTG+8 64
3.28. Products of *HindII* digestion of NTG+8 64
3.29. Sub-cloning strategy for NTG+8 65
3.30. NTG+8 *SacI/HindII* (150μl) on 2% agarose gel 65
3.31. *SacI/HindII* cut pBluescript 66
3.32. Homology search for 700bp *SacI/HindII* NTG+8 sub-clone 66
3.33. Restriction map & DNA sequence for 700bp *SacI/HindII* NTG+8 sub-clone 68
3.34. *HindII* cut pBluescript 69
3.35. NTG+13 sequenced region 69
3.36. Homology search for NTG+13 70
3.37. Restriction map & DNA sequence for NTG+13 74
3.38. NTG+33 sequenced regions 75
3.39. Restriction map & DNA sequence for NTG+33 75
3.40. SDS PAGE of protein extraction from SK-N-SH & SK-N-Be2c 77
3.41. Colony blots probed with NTG-21 78

V. List of tables

2.1. Dilutions for the standard curve of protein analysis 40
3.1. Example of some isolated NTG clones and their inserts sizes 43
3.2. Data table for NTG's 5, 8, 13, 33 44
3.3. Standard curve results 76
3.4. Value in μg protein/ml for the cell line SK-N-SH & SK-N-Be2c 77

1. Introduction

Cancer is a group of more than 100 different diseases. It occurs when cells become abnormal and keep dividing without control or order forming a mass of excess tissue called a tumour.

Tumours can be benign (not cancer) or malignant (cancer). The cells in malignant tumours can invade and damage nearby tissues and organs. Cancer cells can also break away from a malignant tumour and travel through the bloodstream or the lymphatic system to form new tumours in other parts of the body (Metastasis).

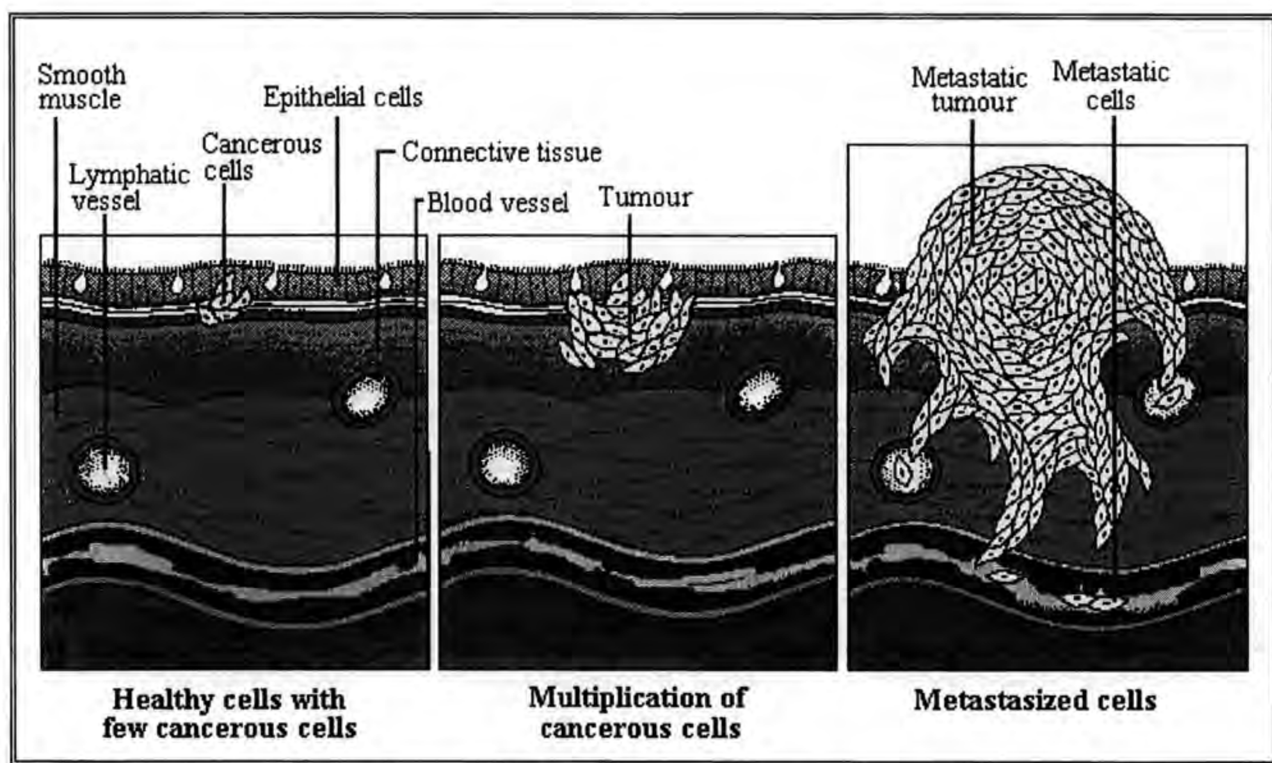


Figure 1.1, The development of cancer.

Cancer is not a new disease, as sometimes humans and fossil bones of dinosaurs show evidence of the disease. Its apparent increase may be due to the increased longevity of humans.

Knowledge about the development of tumours is progressing, and new types of treatments are expected soon.

It is only in the past 50 years that generalised hypotheses about cancer have led to specific findings that shape the treatment people receive.

Cancer in children is not common, but the most common solid tumour is Neuroblastoma.

1.1. Description

Neuroblastoma occurs at a rate of about 1/100,000 per year, most often in babies, very young children (under the age of 15 years), and in the unborn in rare cases. It is slightly more common in white children than in black and slightly more common in boys than girls. About 75% of cases are diagnosed before the age of 5 years.
Although most frequently presenting as a large abdominal mass, this tumour is metastasised in 70% of patients at diagnosis, and there has been little advance in survival over the last 20 years for advanced disease.
The most common primary site is the adrenal gland (40%), pelvis (5%) and neck (5%) are the other sites seen. The tumours arise from neural crest cells which form part of the sympathetic nervous system.

1.2. Stage explanation:

StageI. Localized resectable: The cancer is found only in the place where it started and can be removed by surgery.

StageII. Localized unresectable: The cancer is found only in the place where it started but it cannot be totally removed by surgery.

StageIII. Regional : The cancer has spread from where it started to the tissue around it or to the lymph nodes in the abdomen.

StageIV. Disseminated : The cancer has spread to lymph nodes outside the abdomen or to the bone, liver, or other organs.

StageV. The cancer is in the place where it began and has spread only to the liver, skin, and / or bone marrow.

StageVI. Recurrent neuroblastoma means that the cancer has come back after it has been treated.

1.3. Treatment options

Treatment for Neuroblastoma depends on: the stage of the child's disease, the child's age and the general health.
In general, four types of treatment are used:

- Surgery.
- Radiation therapy.
- Chemotherapy.
- Bone marrow transplantation (using high doses of chemotherapy with or without radiation therapy to destroy the cancer cells and putting bone marrow back into the bone.

1.4. Molecular Characteristics of Neuroblastoma Cells

1.4.1. Decreased expression of the major histocompatibility complex class 1 antigen MHC (they are cell surface glycoproteins that are required for the recognition of target cells by cytotoxic T lymphocytes). This decrease is thought to correlate with high level of expression of the N-*myc* gene.

1.4.2. Urinary Catecholamines

At least 85-90% of tumours have increased levels of catecholamine in the serum and urine [dopa, dopamine, norepinephrine, normetanephrine, homovanillic acid (HVA), vanillylamaidlic acid (VMA)].

The levels are not a guide to prognosis but measurement of total catecholamines,(VMA) and (HVA) can be used as a marker of effective treatment if initially elevated.

Some evidence suggests that the ratio of VMA to HVA may correlate with prognosis in disseminated disease , the higher the ratio the better the prognosis.

The presence of vanillyl acetic acid (VAA) in the urine may be associated with a poor prognosis as may be the elevation of urinary cystathionine levels. Plasma dopamine may be elevated but is rarely measured.

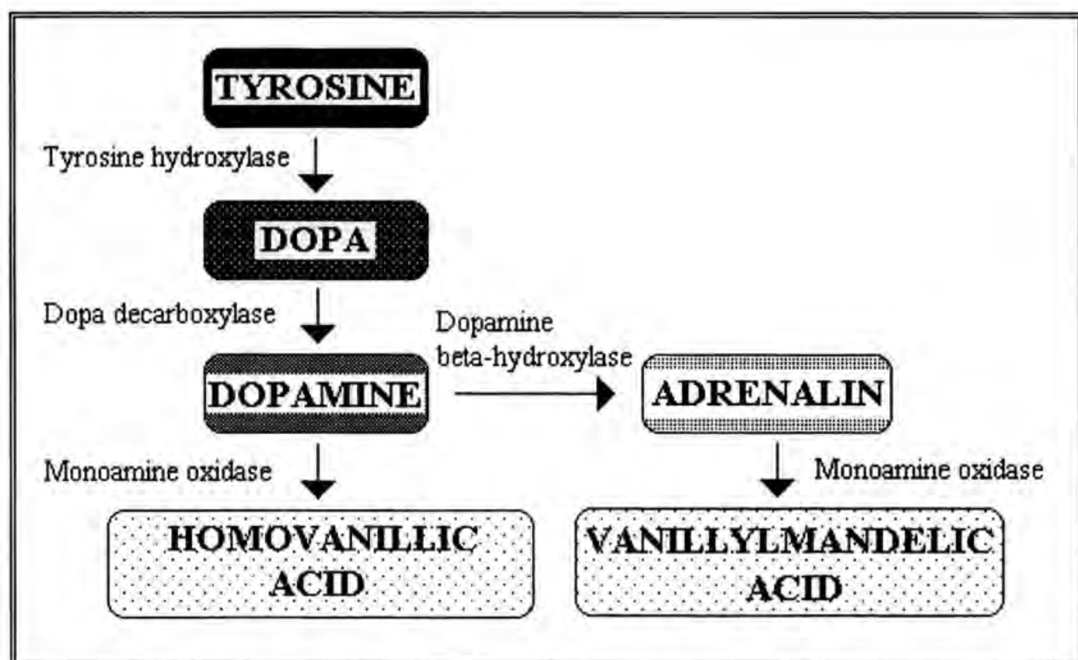


Figure 1.2, Catecholamine Metabolism

1.4.3. Deletion and loss of putative suppresser genes on 1p and 14q

The most characteristic cytogenetic abnormality in neuroblastoma is the deletion at the distal end of the short arm of chromosome 1 from 1p 36.1 to 1p 36.3 , and it is

thought that the p arm of chromosome 1 probably harbours a series of suppresser genes that play a key role in neuroblastoma differentiation.

There is also recent evidence that a deletion or loss of the long arm of chromosome 14 also accrues with increased frequency in neuroblastomas [1,2] and probably represents loss of another suppresser gene .

Another allelic loss has been found involving 11q and 13q in some studies, but not in others[1,2] so the significance of these findings is unclear.

1.4.4. Expression of the nerve growth factor receptors

As mentioned before, neuroblastoma is derived from the sympathoadrenal lineage of the neural crest. Neurotrophic factors and their receptors have been implicated in the pathogenesis of neuroblastoma. Neurotrophins are a family of related molecules that promote neuronal survival and differentiation in the central and peripheral nervous system, these includes nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophin-3 (NT-3) and neurotrophin-4/5 (NT-4/5) [3].

NGF, the first of the neurotrophins to be discovered, has been shown to be essential for survival and differentiation of neural crest-derived peripheral sympathetic and sensory neurones, as well as cholinergic neurones in the central nervous system [4].

Two (NGF) receptor sub-units, the high affinity NGF receptor with tyrosine kinase activity ($P140^{TRKA}$), and a transmemberane glycoprotein, the low affinity NGF receptor ($P75^{NGFR}$), are thought to be required for high affinity NGF binding, and probably for signal transduction [5,6], and that detectability of the NGF receptor sub-unit $P140^{TRKA}$), expressed at different levels by the majority of neuroblastoma's, was associated with lower stage of disease, lower age of patients, and was inversely related to N-*myc* amplification [7].

1.4.5. N-*myc* amplification and expression

Amplification of cellular oncogenes is one of the major genetic alterations that appears to contribute to tumorigenesis [8,9], while most tumours show sporadic amplification of cellular oncogenes, human neuroblastoma tumours show frequent amplification of the N-*myc* gene which correlates with aggressive tumour growth.

A study involving 64 tumours of patients not subjected to therapy revealed that amplification of N-*myc* gene was limited to stage III and IV tumours [10]. Amplification was discovered in approximately 50% of all stage III and IV tumours. This implies that amplification of N-*myc* is a late event in the development of neuroblastomas.

The amplification sequence (N-*myc*) is found either on extrachromosomal double minutes (DMs) or on the homogenous staining regions (HSRs) on different

chromosomes in neuroblastoma cell lines, and the normal single-copy locus was mapped to the distal short arm of chromosome 2 band p23-24 [11,12].
Subsequent studies revealed amplification of N-*myc* also in small cell lung cancers, retinoblastoma, malignant gliomas and astrocytomas, also low expresion of N-Myc is detectable in neuroblastoma cells lacking amplification such as the human neuroblastoma line SK-N-SH or the murine neuroblastoma line Neuro2a, the SH-SY5Y cell line, and G1-ME-N cell line.
The process of gene amplification in general usually involves co-amplification of extensive regions of DNA that share a common origin but which become re-arranged with subsequent amplification events [13].
Recently another gene DDX1, has been shown to be co-amplified with N-*myc* in two retinoblastoma cell lines [14], however rarely in neuroblastoma cell lines, and if so limited to stage 4, 4s.
DDX1 is a member of the DEAD box gene family, has eight highly conserved amino acid motif and maps to the chromosomal region 2p24, the same chromosomal band that N-*myc* maps to (2p-23-24).
It is still unclear whether the DDX1 protein directly affects N-*myc* processing or it's stability or the rate of translation of the N-Myc protein.
Another gene which has been shown to be co-amplified with N-*myc* in neuroblastoma cell lines is the ornithine decarboxylase (ODC1) gene which maps to chromosome 2p24-25, it co-amplifies with N-*myc* in one of six primary neuroblastomas [15].
As a result of N-*myc* amplification, elevated levels of the N-Myc protein are expressed.
N-Myc is like other proteins of the MYC gene family, a short lived protein, located in the nucleus of the cell, phosphorylated by Casein Kinase II (CKII) and binds DNA. It contains a carboxy-terminus, that has a nuclear localisation signal (N), a basic region (BR) responsible for DNA binding, helix-loop-helix (HLH) and a leucine zipper domain (Zip).

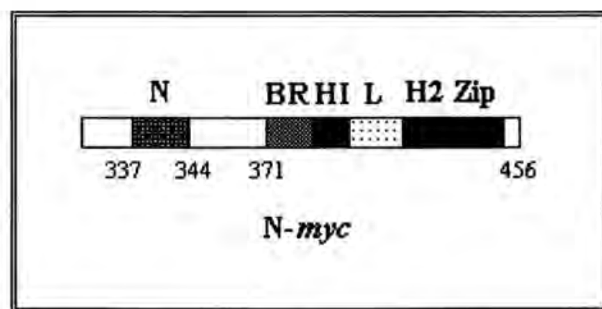


Figure 1.3, Functional domains of N-*myc*

The HLH-Zip domain is responsible for the physical interaction with another HLH-Zip protein, Max, this principal structure of the C-terminus is shared among all members of the myc family of proteins (C-,L-, and N-Myc).
The MYC family of proteins are thought to be involved in transcription because of having both a carboxy-terminal basic-helix-loop-helix-zipper (bHLH-Z) domain, common to a large class of transcription factors [16], and an amino-terminal fragment which, for C-Myc has transactivating function when assayed in chimaeric constructs [17].
As mentioned before, C-, N- and L-Myc proteins heterodimerize in vitro and in vivo with the bHLH-Z protein Max. In vitro, Max homodimerizes but preferentially associates with Myc, which homodimerizes poorly [18].
The resulting Myc-Max heterodimer, binds the nucleotide sequence CACGTG with higher affinity than either homodimer alone [20].

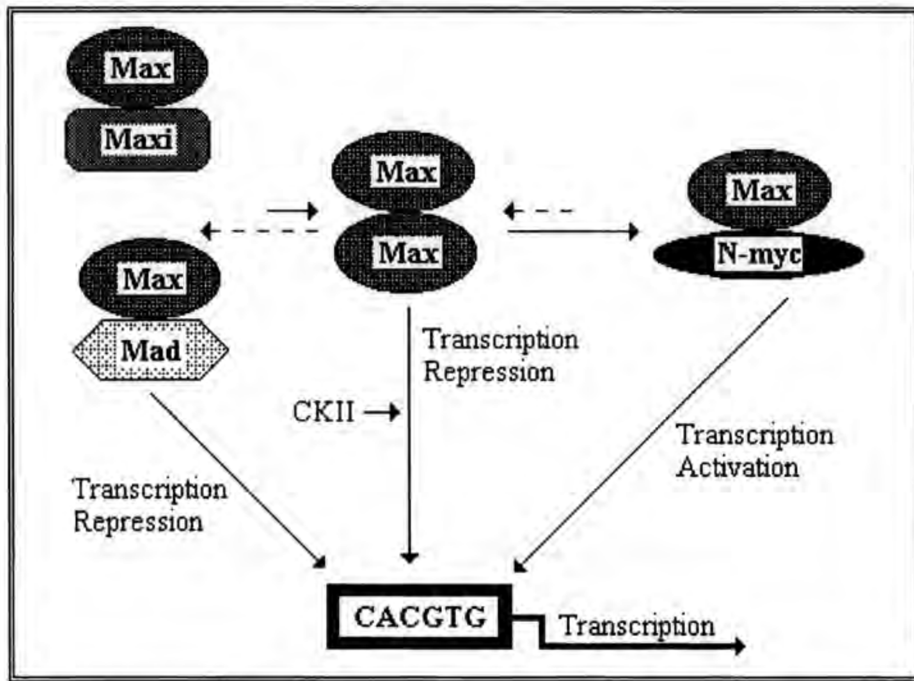


Figure 1.4, A speculative model for interaction and regulation of BR-HLH-Zip proteins in neuroblastoma.

It is thought that Max-Max and Myc-Max complexes have distinct transcriptional functions, thus the relative amounts of Max homodimers and Myc-Max hetrodimers may affect the transcriptional activity of a reporter gene containing promoter-proximal Myc-Max binding sites.
Transcription from such a reporter gene is stimulated by Myc over expression and repressed by Max over expression, activation is dependent on N- and C- terminal regions of Myc, while repression is dependent on the Max DNA-binding domain.

Max and Myc differ in several important biochemical properties, Myc is a highly unstable protein, whose synthesis is rapidly induced upon cell cycle entry, and maintained at a constant level in proliferating cells.

By contrast Max is highly stable and is expressed at equivalent levels in both resting and proliferating cells [21].

As Max seems to be in excess of Myc in vivo and Myc's short half-life is unaltered by dimerization with Max, the rate of Myc synthesis determines the ratio of Max-Max to Myc-Max complexes.

In continuously growing neuroblastoma cells, where N-*myc* is overexpressed as the result of amplification, the presence of N-*myc* forces Max into abnormal amounts of N-Myc-Max heterodimeric complexes, which could perturb the fine tuned expression of unidentified genes or possibly genes involved in control of the cell cycle and differentiation, see figure 4.

The p21- and $p22^{max}$ are the major variants that are expressed in a wide variety of species and cell types, and probably the predominantly expressed forms of Max in neuroblastoma cells.

Expression of Max proteins is not restricted to cells with high levels of N-Myc, but was also observed in neuroblastoma cells without detectable N-*myc* expression.

Max also interacts with C-myc. The C-myc proto-oncogene is a key regulator of cell growth and differentiation, it is localized in the nucleus and can activate transcription when brought to DNA by heterologous DNA binding domains [22], it contains an activation domain in its amino terminus, the integrity of which is correlated with its ability to cause oncogenic transformation [23].

Max protein is expressed in all tissues in which C-myc is expressed, and in some, including the adult brain in which C-myc is not.

Max-Max homodimers and C-myc-Max heterodimers bind tightly to the E-box CACGTG, and while DNA bound C-myc-Max hetrodimers activate transcription, Max-Max homodimers do not. This fact strengthened the hypothesis that transcription regulation is important to Myc function [24,25].

The interaction between Max and C-myc is dependent on the integrity of the C-myc HLH-Zip domain, but not on the basic region or other sequences outside the domain. As for C-myc there is substantial evidence that the bHLH region and the adjacent leucine zipper motif are functionally important, deletions within these regions result in loss or alteration of transforming activity as well as reduction of the capacity to autoregulate endogenous Myc expression and to inhibit cell differentiation [26,27].

It is suggested that because most C-myc in vivo is associated with Max and C-myc-Max heterodimers bind CACGTG sites more tightly than C-myc-C-myc homodimers, that one function of Max is to facilitate the binding of C-myc to these

sites, and the fact that N-Myc and L-Myc as well as C-Myc specifially associate with Max suggests that Max may also serve to integrate the functions of these three proteins that are differentially expressed during development, differentiation, and neoplasia.

Max (but not Myc) also interacts with other HLH-Zip proteins, Mad and Mxil and may also probably contribute to transcriptional control in neuroblastoma Mad-Max and Mxil-Max hetrodimers bind efficiently to the Myc-Max E-box. Thus, (Mad / Max heterodimers formed during the differentiation of cells are thought to be more effective repressors than max homodimers because of their higher DNA binding activity).

It is also thought that Mad-Max and Mxil-Max heterodimers inhibit Myc-Max transactivation by sequestering Max and therefore preventing formation of Myc-Max heterodimers, and also by competing with Myc-Max heterodimers for common target sites at promoters (E-Box), see figure 4.
The outcome of this whole system is that; In the normal situation both homodimeric and heterodimeric populations of BR-HLH-Zip proteins exist at a level that is strictly regulated, and if the level of one member of this system is altered as a result of a genetic change, regulation is lost with the consequence of neoplastic cell growth.

1.5. Summary

From what has been written so far, the following out-lines are drawn:

- N-*myc* is a cellular oncogene, which seems to contribute to the growth control of cells.
- The process of gene amplification usually involves co-amplification of extensive regions of DNA.
- Co-amplification of the DDX1 gene is rare in neuroblastoma cells, and often limited to stage 4-4s.
- N-*myc* co-amplifies with the ODC1 gene in one of six primary neuroblastomas.
- The protein produced by N-*myc* oncogene is N-Myc, which is thought to be a putative transcription regulator (may be responsible for either "turning on" or "turning off" other genes).

- N-*myc* normal activity may require a fine tuned interplay with at least two or more proteins, and might be regulated at least in part, through interaction with these proteins.
- Enhanced levels of N-Myc could result in disturbance of the balanced interaction (with the other proteins).
- N-Myc and Max work together in binding DNA, and probably enhanced levels of N-Myc lead's to a greater number of transcription activating complexes.

1.6. Aim of research

The aim of this project is to identify potential targets for N-*myc* regulation using the cell line SK-N-SH.

SK-N-SH are human neuroblastoma cells; a metastasis from a bone marrow sample and comprise of two distinct populations of cells:

- N - type cells show neuroblastic morphology.
- S - type cells are large and epitheliod.

Karyotype shows chromosomal translocation between chromosome 7 and chromosome 22, and finally this cell line has no amplification of the oncogene N-*myc* (low levels of N-Myc protein).

N-*myc* target genes were isolated using a modification of the protocol developed to identify p53 binding sites [28], and as follows (see figure 5):

- Generating NTG library, by isolating genomic DNA from SK-N-SH cells digested to completion with *Sau3A* and mixed with nuclear extract from SK-N-SH cells, and finally immunoprecipitated with N-*myc* antibodies.
- DNA recovered from these precipitates was cloned into pBluescript SK+, and repeated immunoprecipitation on the primary N-myc selected library either:
  - In the presence of poly dIdC → +NTG clones.
  - Without competition → -NTG clones.
- 400 clones have been isolated, their size of insert calculated, and pure plasmid prepared for inserts larger than 300bp.
- Inserts were then sequenced using reverse and universal primers.
- DNA sequence homology searches have been carried out using the Daresbury data base.

2. Materials and Methods

2.1. Materials:

2.1.1- NTG library

⇒ Solution I (Cell resuspension solution)

- Glucose 50 mM
- Tris.Cl (pH 8.0) 25 mM
- EDTA (pH 8.0) 10 mM
- Autoclave at 15 lb. in^{-2} for 20 min.

⇒ Solution III (Potassium acetate)

- Per 100 ml:
- KOAc (5 M) 60 ml.
- Glacial acetic acid 11.5 ml.
- Autoclave at 15 lb. in^{-2} for 20 min.

⇒ Sodium Acetate (3M)

- Per litre:
- $NaOAc.3H_2O$ 408.1 g.
- Adjust to pH 5.2 with glacial acetic acid.
- Autoclave at 15 lb. in^{-2} for 20 min.

⇒ 1 M Tris.Cl pH 7.5

- Per litre:
- Tris 121.1 g
- Adjust to pH 7.5 with HCl.
- Autoclave at 15 lb. in^{-2} for 20 min.

⇒ Trypsin

1 ml of Trypsin (Gibco BRL) + 10ml PBS/EDTA. Make up immediately prior to use.

⇒ 1 M TRIS. pH (8)

- Per 100 ml H_2O:
- 12.11 g

⇒ Solution II (Cell lysis solution)

- SDS 1%
- NaOH 0.2M
- Prepare a 2% solution of SDS and a 0.4 M solution of NaOH, separately. Autoclave at 15 lb. in^{-2} for 20 min. Mix 1:1 prior to use to give a 1% and 0.2 M solution.

⇒ 0.5 M EDTA pH 8.0

- Per litre:
- EDTA $Na_2.2H_2O$ 186.1 g
- Adjust to pH 8.0 in order that the EDTA will dissolve.
- Autoclave at 15 lb. in^{-2} for 20 min.

⇒ TFBI.

- KOAc (1M) 6 ml (fc 30 mM)
- RbCl 2.418 g (fc 100 mM)
- $CaCl_2$ (1M) 2 ml (fc 10 mM)
- $MgCl_2$ (1M) 10 ml (fc 50 mM)
- Glycerol 30 ml (fc 15%)
- H_2O to 200 ml. Adjust to pH 5.8 with 0.2 acetic acid (few drops). Filter sterilize, store 4°C.

⇒ TFBII.

- MOPS 0.2093 g (fc 10 mM)
- CaCl2 (1M) 7.5 ml (fc 75 mM)
- RbCl 0.1209 g (fc 10 mM)
- Glycerol 15 ml
- H_2O to 100 ml. Adjust to pH 6.5 with KOH (drop or two). Filter sterilize, store 4°C.

⇒ Ammonium Acetate 10M

- 1.92 gm. / 25 ml H_2O.

⇒ Leupeptin 0.5 mg/ml

- Dissolve 1 mg in 2 ml of sterile H_2O, store at -20°C.

⇒ Pepstatin A 5 mg/ml

- Dissolve 5 mg in 5 ml of ethanol, store at -20°C.

⇒ NETN

♦ NaCl	100 mM
♦ EDTA	1 mM
♦ Tris.Cl pH 8.0	20 mM
♦ Nonidet P-40	0.5 %
♦ PMSF	1 mM

⇒ PMSF 100 mM

- ♦ Dissolve 17.4 mg in 1 ml of isopropanol, store at -20°C.

⇒ Leupeptin 0.5 mg/ml

- ♦ 7.5 μg in 100 μl NETN.

⇒ RNAase A

- 10 mg / 1ml buffer:

♦ Tris.Cl (pH 7.5)	10mM
♦ NaCl	15mM

⇒ Proteinase K

- ♦ 20 mg / 1ml

⇒ Ampicillin

- ♦ 0.25 gm / 10ml of sterile water.

⇒ TBS

♦ Tris. Cl pH 7.5	20 mM
♦ NaCl	500 mM

⇒ λ Hind III DNA marker

- Set up the following digestion in an eppendorf:

♦ H_2O	780 μl
♦ Buffer 10	100 μl
♦ λ DNA	100 μl
♦ Hind III	20 μl

- ♦ Leave in the 37°C heat block overnight.
- ♦ Add 100 μl of 6X loading buffer, store at -20°C

⇒ X Gal [DMFX Galactoside]

- ♦ 0.4 gm / 10 ml DMF.

⇒ Agarose gel

♦ 0.8 % gel	0.48g Agarose / 60 ml 1X TAE buffer
♦ 1.2 % gel	0.72g Agarose / 60 ml 1X TAE buffer
♦ 1.5 % gel	0.9g Agarose / 60 ml 1X TAE buffer
♦ 2.0 % gel	1.2g Agarose / 60 ml 1X TAE buffer

⇒ Nutrient Broth (LABM)

- 2.5g of labM Nutrient / 100ml H_2O
- Autoclave at 15 lb. in^{-2} for 20 min.
- Add 200 µl of Ampicillin before use

⇒ Nutrient Agar (LABM)

- 2.8g / 100ml H_2O
- Autoclave at 15 lb. in^{-2} for 20 min.
- Add Ampicillin-Xgal before use.

⇒ PBS / EDTA

- 5 PBS tablets (Dulbeccos phosphate w/o calcium or magnesium) + 0.1g EDTA in 500ml of H_2O.
- Aliquot into 100ml bottles and autoclave. Store at 4°C.

Buffers:

⇒ T.E. Buffer (pH 8.0).

- Per litre:

♦ Ttis.Cl (pH 7.5)	10 mM
♦ EDTA (pH 8.0)	1.0 mM

- Autoclave at 15 lb. in^{-2} for 20 min.

⇒ 1X TAE buffer

- Per 5 litres:

♦ 50X TAE	100 ml.

⇒ 6X Loading buffer

♦ Bromophenol blue	0.25% (w/v)
♦ Xylene cyanol ff	0.25% (w/v)
♦ Glycerol	30% (w/v)

⇒ 50X Tris-acetate electrophoresis buffer (TAE)

- Per litre:

♦ Tris	242.0 g.
♦ Glacial acetic acid	57.1 ml.
♦ EDTA (0.5 M pH 8.0)	100 ml.

⇒ Wash buffer

- NENT containing:

♦ PMSF	0.5 mM
♦ Leupeptin	1μg / ml
♦ Pepstatin	1μg / ml

⇒ Dissociation buffer

♦ Tris.Cl pH 9.0	500 mM
♦ EDTA	20 mM
♦ NaCl	10 mM
♦ SDS	0.2%

⇒ Extraction buffer

♦ Tris.HCl (pH 8.0)	10mM
♦ EDTA (pH 8.0)	0.1mM
♦ Pencreatic RNAase	20 μg/ml
♦ SDS	0.5 %

⇒ Buffer A

♦ 100mM Hepes pH 7.9	10 ml (fc 10mM)
♦ 250 mM KCl	4 ml (fc 10mM)
♦ 100 mM EDTA	20μl (fc 0.1mM)
♦ 10 mM EGTA	1ml (fc 0.1mM)
♦ 100 mM DTT	1ml (fc 1mM)

♦ Add to H_2O final volume100ml

♦ Add PMSF to 0.5mM before use.

⇒ 10X CIP dephosphorylation buffer

♦ $ZnCl_2$	10mM
♦ $MgCl_2$	10mM
♦ Tris.Cl (pH 8.3)	100 mM

⇒ Binding buffer

- Tris.HCl pH 7.5 20mM
- Glycerol 4%
- DTT 1 mM
- NaCl 50mM
- EDTA 0.1mM
- Spermidine 2.5mM
- $MgCl_2$ 7.5mM
- Leupeptin 1μg / ml
- Pepstatin 1μg / ml
- PMSF 0.5mM

⇒ CA/Mn buffer (pH 5.5)

- $CaCl_2.2H_2O$ 100 mM (0.735g / 50ml)
- $MnCl_2.4H_2O$ 70 mM (0.69265g / 50ml)
- NAOAc(anhyd) 40 mM (0.1641g / 50ml)

Prepare fresh and filter sterilize. Start with a solution of the NaOAc at pH 7.0, add the Ca and Mn, this will bring the pH down to nearer 6.0. Finally adjust the pH down to 5.5 with acetic acid, *never adjust the pH up with NaOH.*

Make up to 42.5 ml filter sterilize and then add 7.5 ml of sterile glycerol for the buffer with 15% glycerol.

⇒ Buffer C

- 100 mM Hepes pH 7.9 20ml (fc 20mM)
- 4M NaCl 10ml (fc 0.4M)
- 100 mM EDTA 1 ml (fc 1mM)
- 10 mM EGTA 10ml (fc 1mM)
- 100 mM DTT 1ml (fc 1mM)
- Add to H_2O final volume100ml
- Add PMSF to 1mM before use.

2.1.2- Colony blots

- *Preparation of solutions:*

⇒ SDS 10%

⇒ Denaturing solution

- NaOH 0.5 N
- NaCl 1.5 M

⇒ Neutralizing solution

- NaCl 1.5 M
- Tris.Cl (pH 7.5) 0.5 M

⇒ 20X SSC

- Per litre:
- NaCl 175.3 g
- Sodium citrate 88.2 g

Adjust to pH 7.0 with 10 N NaOH.

2.1.3- Filter hybridization

- *Preparation of solutions:*

⇒ Pre- hybridization

Solution	Volume	Concentration
20 X SSPE	12.5 ml	5 X SSPE
100 X Denhardt's	2.5 ml	5 X Denhardt's
10% [w/v] SDS	2.5 ml	0.5% [w/v] SDS

- + 3.5 ml H_2O + 100 μl herring sperm [10mg/ml].

⇒ Hybridization

Solution	Volume	Concentration
20 X SSPE	12.5 ml	5 X SSPE
100 X Denhardt's	2.5 ml	5 X Denhardt's
10% [w/v] SDS	2.5 ml	0.5% [w/v] SDS

- + 3.5 ml H_2O + 100 μl 0.5 M EDTA.

2.1.4- Protein Analysis

- *Preparation of solutions:*

⇒ Protein stain

- Per litre:
- H_2O 450ml
- Methanol 450ml
- Acetic acid 100ml
- Coomassie brilliant blue R250 1g

⇒ SDS page sample buffer

- H_2O 4ml
- 0.5 M Tris.Cl (pH 6.8) 1.0ml
- Glycerol 0.8ml
- 10% SDS 1.6ml
- β Mercaptoethanol 0.4ml
- 0.1% Bromophenol blue 0.2ml [of 1mg in 1ml H_2O]
- Store in -20°C

⇒ 45% Acrylamide solution

- 22.5g in 50 ml H_2O

⇒ 1.6% Bisacrylamide

- 0.32 in 50 ml H_2O

⇒ 1.5M Tris.Cl (pH 8.8)

⇒ 4% SDS

- 0.1 SDS

2gm SDS + 50ml H_2O

⇒ SDS page electrophoresis buffer

- Per 3 litre:
- Tris 9.1g
- Glycine 43.2g
- SDS 3.0g

⇒ 12.5% Resolving gel

- 45% Acrylamide 2.78ml
- 1.6% Bisacrylamide 0.64ml
- 1.5M Tris pH 8.8 2.5ml
- 4% SDS 0.26ml
- H_2O 3.82ml
- TEMED 13.2μl
- 0.56% APS 0.5ml

⇒ Protein stain

- Per litre:

◆ H_2O	450ml
◇ Methanol	450ml
◆ Acetic acid	100ml

⇒ 0.65% APS (ammonium persulphate)

- Prepare fresh when needed:

◆ 11.2mg in 2ml H_2O

⇒ 4% Stacking gel

◆ 45% Acrylamide	0.7ml
◆ 1.6% Bisacrylamide	0.55ml
◆ 0.5M Tris pH 6.8	2.0ml
◆ 4% SDS	0.4ml
◆ H_2O	4.6ml
◆ TEMED	12µl
◆ 0.56% APS	0.55ml

2.2. Methods:

2.2.1. Tissue culture

2.2.1.1- Recovering SK-N-SH cells from liquid Nitrogen

Cells were removed from the freezer, thawed by gently swirling the tube in a beaker of water at 37°C and then transferring the contents to a sterile 25ml universal. 5ml of media (RPMI 1640 Medium with Glutamax-1, from GIBCO BRL) was then added slowly over a period of 2 minutes.

Cells were then pelleted at 1500rpm for 2 minutes at room temperature, and re-suspended in 1ml of medium containing 10% FCS. The contents were then transferred to a $25cm^2$ flask, a further 2ml of medium was added and incubated at 37°C in a 5% CO_2 incubator overnight (until cells attach).

[Media was changed every 3-4 days]

2.2.1.2- Trypsinisation of cells for passaging

Medium was decanted from the cells, which were then washed twice with PBS/EDTA, ensuring all surfaces of the flask as well as the monolayer were covered.

2.5ml of Trypsin was added to rinse the cells and then poured off leaving a thin film of trypsin covering the cell surface.

The sealed flask was then placed in 5% CO_2 incubator for one minute or until cells were rounded up and detached.

3 ml of medium + 10% FCS was added to the cell suspension to inhibit trypsin and then decanted into a sterile 25ml universal.

The flask was then rinsed with a further 5ml of medium and decanted into the 25ml universal. The cells were centrifuged at 1500rpm at room temperature for 2 minutes, then, after pouring off the medium 1ml of fresh media was added and aspirated to a single cell suspension. The contents were then transferred to a $25cm^2$ flask and an appropriate volume of medium containing 10% FCS (2ml) was added. The flask was then kept in a 5% CO_2 incubator at 37°C overnight (changing medium every 3-4 day).

2.2.1.3- Preservation of cells in liquid Nitrogen

Cells had to be trypsinized first (previous method) but re-suspended in 1ml of 90% FCS / 10% DMSO. 0.5ml of cell suspension was added to NUNC Cryovials (screw tops finger tight), and placed at -20°C for 30 minutes and then transferred to an insulated box at –70°C overnight, then to a liquid Nitrogen storage container.

2.2.2. DNA Manipulation

To begin the research, a number of NTG clones are required (which may have N-Myc target genes). These clones are generated from what is called the N-Myc binding library. In order to begin the library, a number of preparations are required, for example: Competent sure cells, Genomic DNA, Nuclear extraction, Competent DH5 ∝ cells, pBluescript SK+.

2.2.2.1- Complete sure cells

5ml of culture was grown overnight (E.Coli) and then using 1ml (5ml) to inoculate 100ml of pre-warmed LB, left to grow at 37°C until OD_{600}=0.5, the cells were then centrifuged in 50ml aliquots in falcon tubes.

The pellets were re-suspended in ½ the volume of ice-cold Ca/Mn buffer and after 1 hour incubation on ice, centrifuged at 200rpm for 5 minutes and re-suspended very gently in 1/20 the volume of ice-cold Ca/Mn buffer + 15% glycerol and finally dispensed in 200 μl aliquots into cold eppendorfs, frozen in liquid Nitrogen and transferred to ice (stored at -80°C).

2.2.2.2- Competent DH5 ∝ cells

5ml of culture was grown overnight, and 1ml was used to inoculate 100ml of pre-warmed LB, it was then left to grow at 37°C until OD_{600} reached 0.3-0.35. The cells were then chilled on ice for 5 minutes and pelleted using the MSE mistral for 7 minutes at 3600rpm, 4°C and re-suspended in 2/5 the volume using TFBI, left again on ice for 5 minutes, pelleted down, re-suspended in 1/25 the volume using TFBII and finally dispensed in 200μl aliquots into pre-chilled eppendorfs and frozen in liquid Nitrogen (stored at -80°C).

2.2.2.3- Isolation of Genomic DNA

I- Preparation of cells

SK-N-SH cells growing in monolayer:

a- Cells were detached from the dishes by trypsinisation and recovered by centrifugation at 1500 rpm at 4°C for 10 minutes.
b- Cells were re-suspended in 5-10 volumes of ice-cold TBS and recovered again at 1500rpm at 4°C for 10 minutes.
c- Re-suspended in 1ml TE, 10ml of extraction buffer was added and incubated for 1 hour at 37 °C.

II- DNA Extraction

a. After 1 hour of incubation proteinase K was added to a final concentration of 100 μg/ml (50 μl of 20 ml/ml stock) and incubated at 50 °C for 3 hours with periodic mixing.

b. After 3 hours of incubation the solution was left to cool down at room temperature and extracted with an equal volume of Phenol equilibrated with 0.5 M Tris-HCl pH 8.0 and was left to mix gently for 2 minutes on a rocking platform.
c. The solution was then centrifuged at 5000g for 50 minutes at room temperature, then the aqueous layer was carefully transferred to a fresh tube.
d. Step b & c was repeated twice more.
e. For isolating DNA in the size range 100-150Kb a 0.2 volume of 10M ammonium acetate and 2 volumes of ethanol was added at room temperature (swirled until thoroughly mixed), and the DNA precipitate was removed with a sterile needle or by centrifugation if (DNA fragments).
f. DNA pellet was washed twice with 70% of ethanol and re-dissolved in 1ml of TE for each $5X10^6$ cells by gently rocking the solution for 12-24 hours.
g. 5µl of DNA Extraction was run on a 0.8% gel to check DNA presence and purification. (see figure 2.1)

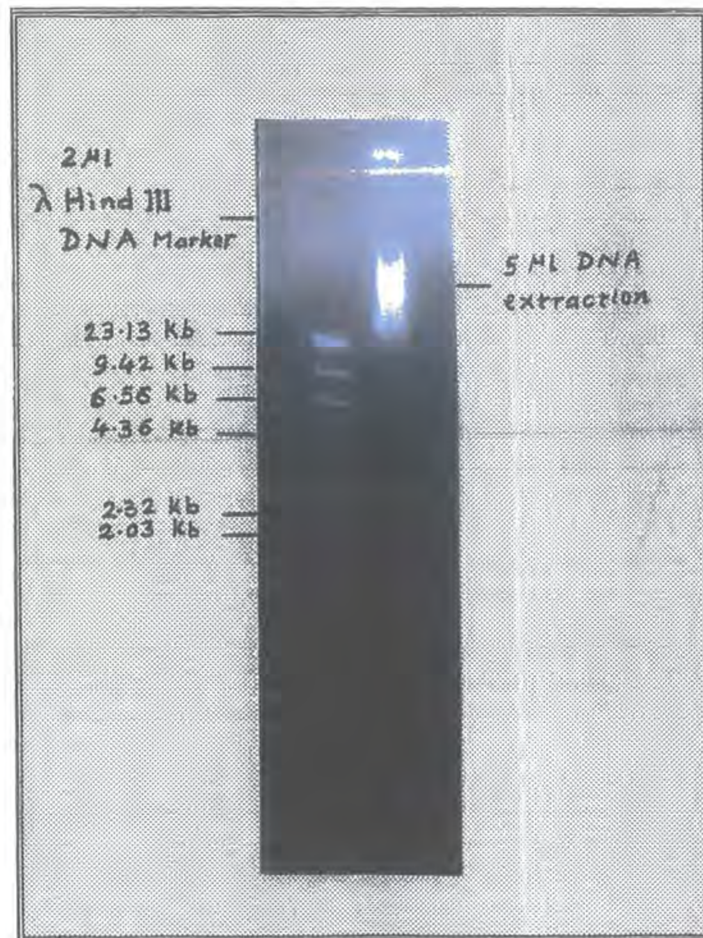


Figure 2.1, 5µl of DNA extraction on 0.8% Agarose gel.

III- *Sau3A* Genomic DNA

The following were mixed:

a. 100μl of Genomic DNA.
b. 20μl of buffer 2.
c. 10μl of *Sau3A*.
d. 70μl of H_2O.

and incubated for 5 hours at 37°C, then precipitated with 20μl of Sodium Acetate and (Ethanol, 70% Ethanol at -20°C). The pellet was then re-suspended in 20μl of H_2O. (see figure 2.2)

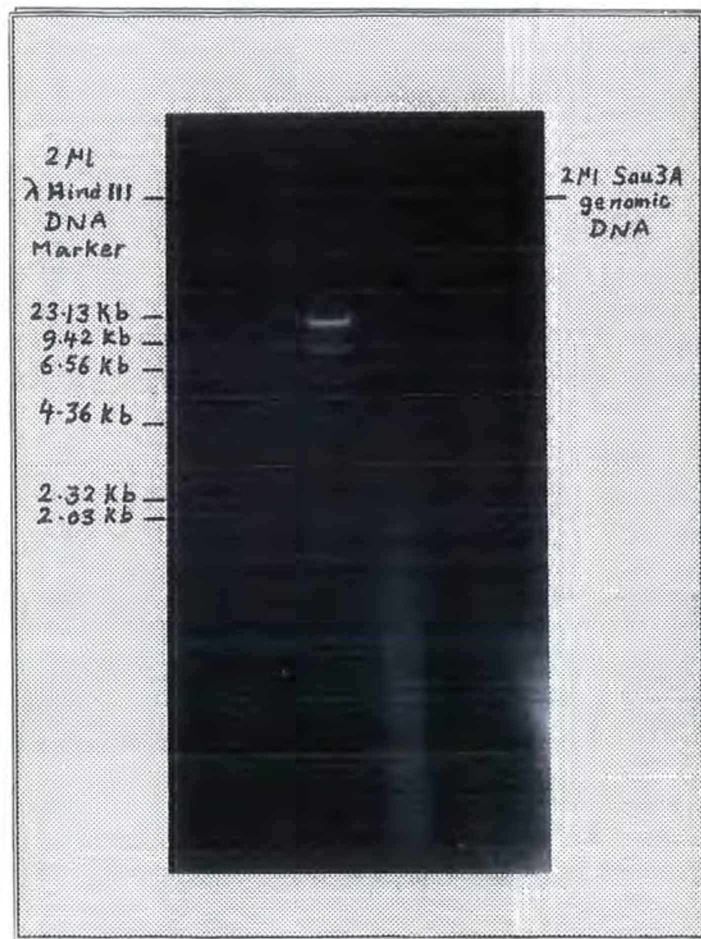


Figure 2.2, 2μl of *Sau3A* cut genomic DNA on 0.8% Agarose gel.

2.2.2.4- <u>Nuclear Protein Extraction</u>

a. Cells (SK-N-SH) were scraped from tissue culture flasks into TBS (5ml), and the flasks were washed with a further 2ml of TBS.
b. Cells were then pelleted at 6500rpm in a refrigerated centrifuge for 2 minutes and re-suspended in 2ml of TBS, then transferred to an eppendorf tube, pelleted in MSE mistral at 6500rpm for 15 seconds.

c. TBS was then removed and the pellet re-suspended in 400µl of cold buffer A, swelled on ice for 15 minutes then 25ml of 10% Nonidet NP-40 was added.
d. After 10 seconds of vortexing the nuclei were pelleted at 6500rpm in MSE microcentrifuge for 30 seconds and re-suspended in 50µl of ice cold buffer C, the nuclei were then left on ice for 15 minutes with occasional shaking.
e. Lysed nuclei and other debris were removed by centrifugation at 13000rpm, 40°C for 50 minutes (store at -70 °C).

2.2.2.5- pBluescript SK+

I- Preparation of pBluescript

a. 3ml of culture was grown overnight (pBluescript SK+ cells) centrifuged for 3 minutes and re-suspended in 200µl of ice cold solution.
b. 200µl of solution II was added and mixed by gentle inversion, then mixed with 200µl of ice cold solution III. The solution was then centrifuged at full speed in a microfuge for 5 minutes to pellet the precipitate. The supernatant was then transferred to a clean eppendorf (avoiding pipetting up any of the precipitate) and centrifuged once more for 1 minute.
c. The supernatant was transferred to a clean eppendorf tube and 2 volumes of absolute ethanol (at -20 °C) was added then left on ice at least 15 minutes.
d. The solution was centrifuged at full speed for 15 minutes. The ethanol poured off and 1 ml of 70% ethanol (at -20 °C) was added, centrifuged for 3 minutes and ethanol poured off.
e. The pellet was then re-suspended in 30µl of H_2O.
f. 2µl was run on a 0.8% gel to check the presence of pBluescript.

II- *BamH1* pBluescript

The following were mixed:

a. 50µl pBluescript.
b. 10µl Buffer 4.
c. 10µl *BamH1*.
d. 30µl H_2O.

Left for 3 hours at 37°C, then precipitated with 2 volumes of Ethanol, centrifuged and the pellet washed with 70% Ethanol at -20°C. The pellet was then re-suspended with 40µl of H_2O.

10µl is kept as a non CIP (Calf Intestinal Alkaline Phosphates) *BamH1* pBluescript (2µl was run on a 0.8% gel). (see figure 2.3)

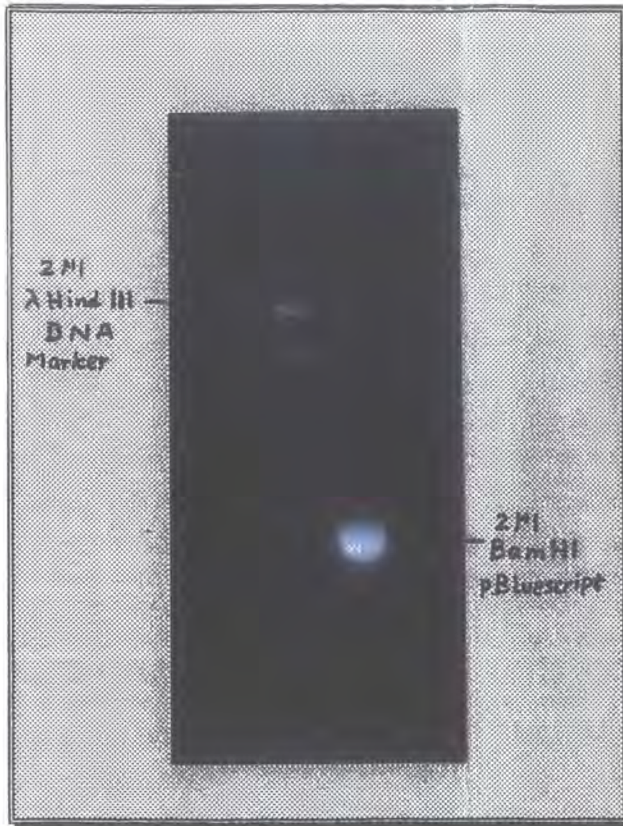


Figure 2.3, 2μl of *BamH1* pBluescript on 0.8% agarose gel.

The remaining 3μl was dephosphorylated with CIP. This was done by the following method. The following were mixed:

a. 30μl of *BamH1* pBluescript.

b. 10μl of (10 X CIP dephosphorylation buffer).

c. 60μl of H_2O.

d. 1μl of CIP.

After 30 minutes incubation at 37°C the following was added:

a. 50μl of 2% SDS

b. 10μl of 100mM EDTA

c. 2μl of 10mg/ml Proteinase K.

d. 38μl of H_2O

Incubate again at -56°C for 30 minutes. The solution was then extracted with 200μl of Phenol : Chloroform, centrifuged for 3 minutes, transferring aqueous layer to a fresh tube, and a further 200μl of Phenol was added.

The solution was then precipitated with 20μl of Sodium acetate and 2 volume of Ethanol, 70% Ethanol, and finally the pellet was re-dissolved in 15μl of H_2O.

(2μl of CIP *BamH1* pBluescript was run on a 0.8% gel). (see figure 2.4)

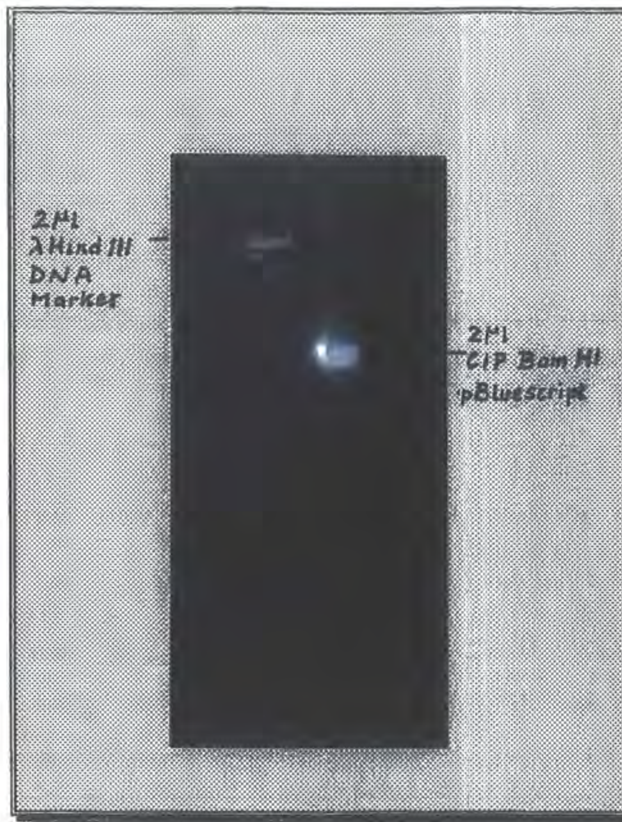


Figure 2.4, 2μl of CIP *BamH1* pBluescript on 0.8% Agarose gel.

2.2.2.6- Preparation of N-*myc* Binding library

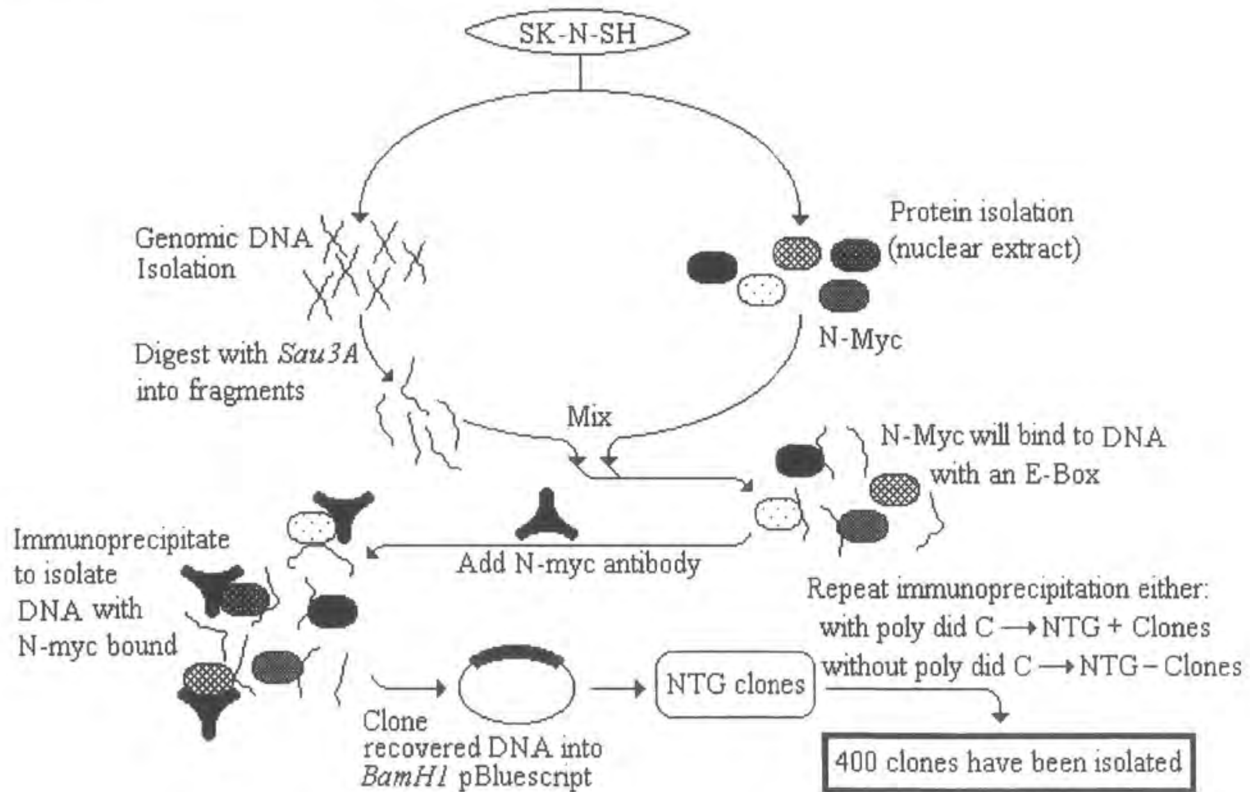


Figure 2.5, Stage explanation of preparation of N-myc binding library.

I- First round of screening

a. The library was started by mixing the following:
    1. 10μg of *Sau3A* DNA
    2. 50μl of Nuclear extract
    3. In a total volume of 100μl of binding buffer and incubating at room temperature for 1 hour. Then, 2 μg of N-myc antibody was added.
b. After 1 hour of incubation at room temperature, protein A- sepharose (7.5mg in 100μl NETN buffer) was added, then mixed and incubated for a further 30 minutes. The solution was then washed twice in washing buffer (30μl) and the DNA dissociated in 200μl of dissociation buffer at 55°C for 30 minutes.
c. 200μl of Phenol-Chloroform was used to extract the DNA (aqueous layer) which was then precipitated with $^1/_3$ Volume of ethanol for 16 hour at -80°C.
d. The DNA was recovered by 15 minutes centrifugation and re-dissolved in 10μl H_2O, then cloned into *BamH1* pBluescript (half treated with CIP), and then sure cells transformed with vector.

II- Ligation

a. CIP BamH1 pBluescript

   The following were mixed:
    1. 5μg of DNA.
    2. 10μl of CIP pBluescript.
    3. 4μl of ligase buffer (5X).
    4. 1μl of ligase.
b. Non CIP *BamH1* pBluescript

   The following were mixed:
    1. 5μg of DNA.
    2. 8 μl pBluescript.
    3. 4 μl of ligase buffer.
    4. 1μl of ligase
    5. 2μl of H_2O.

Both were stored in 4°C overnight to ligate.

III- Sure cell transformation

a. The 20µl of CIP ligation was split into 2 eppendorts, each of them had the following:
1. 10µl of ligation.
2. 40µl of TE.
3. 100µl of Sure cells.

They were left on ice for 30 minutes, then heat shocked at 37°C for 5 minutes, diluted in to 1ml of pre-warmed LB (each eppendorf), and finally left shaking for 1-2 hours (100 minutes optimum) at 37°C, then both tubes were centrifuged, the media decanted and re-suspended in 50µl of LB.
Both 150µl were plated out on the same plate and incubated at 37°C overnight. (Plates had Ampicillin + X Gal)
[The same method was used for the non CIP *BamH1* pBluescript ligation]

b. Transformants were scraped off into 1.5ml LB and prepared plasmid DNA from 750µl, for each CIP treated and non CIP, the other 750µl was stored as glycerol at -80°C.

IV- Preparation of plasmid DNA

a. The transformants were pelleted at full speed in a microfuge for 3 minutes and discarding the supernatant, 200µl of solution 1 was added to resuspend the cells plus 4µl of RNAseA of (10mg RNA/1ml Tric.Cl+NaCl buffer).

A further 200µl of solution was added and mixed by gentle inversion, and finally 200µl of solution III was added and mixed by shaking.

b. The precipitate was pelleted at full speed in a microfuge for 5 minutes, and extracted with an equal volume of Phenol: Chloroform.

The aqueous layer was transferred to a clean eppendorf tube and Ethanol precipitated, finally, re-suspending the DNA pellet in 22µl H_2O. This method was used for both CIP transformants and non CIP transformants.
(a 2µl of each was run on a 0.8% gel). (see figure 2.6)

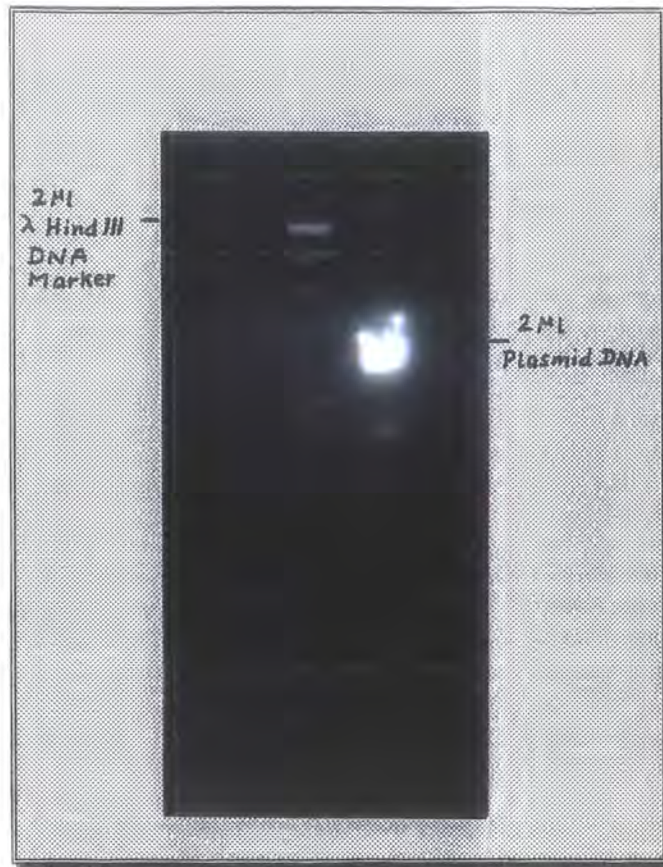


Figure 2.6, 2µl of plasmid DNA for CIP transformants on 0.8% Agarose gel.

V- Second round of screening

The following were mixed:

1. 15µl of CIP (whole plasmid).
2. 50µl of Nuclear Extract.
3. 0.5µl of 1µg/µl poly (dIdC)-[to generate NTG + colonies or without competitor to generate NTG-colonies]
4. In total volume of 100µl of binding buffer.

Then the procedure from 6-1- to 6-3-a of the first round was repeated (large ampicillin -X Gal plates were used, 500ml agar on each plate).

The plates had a large number of NTG clones (whites). A total of 400 clones were picked, restreaked on to new ampicillin plates and given a number (Exp NTG + 1). For each NTG clone a 3 ml. overnight culture was grown in order to isolate the plasmid DNA. A 2µl aliquot of each plasmid was run on a 0.8% gel and then after digesting the insert DNA with *SacI/EcoRI*, run on a 1.2% gel and calculating the size of insert using a standard graph.

Different sizes of insert were chosen (larger than 300 bp), and pure plasmid DNA was prepared using a QIA prep plasmid kit (By QIAGEN).

The inserts were then sequenced by Julea Bartley using an AB1 373A automatic DNA sequencer using the reverse (RV) and universal (UP) primers, then the sequence was transmitted to SEQNET Daresbury and compared to DNA and protein databases using the FASTA and GAP algorithms. Candidate homology were selected based on high similarity over an extended length compared with the actual length of the inserts and an assessment of their potential significance to neuroblastoma.

2.2.2.7- Isolation of plasmid DNA

This method was used for each of the 400 clones, stored at -20°C until needed.

Cells were pelleted from a 3ml overnight culture at full speed in a microfuge for 3minutes, and then re-suspended in 200μl of solution 1 pipetting.

A further 200μl of solution II was added and mixed by gentle inversion, and finally 200μl of solution III is added and mixed by shaking. The solution was then centrifuged at full speed in a microfuge for 5 minutes to pellet the precipitate.
The supernatant was transferred to a clean Eppendorf tube, and Ethanol precipitated with (Ethanol/ 70% Ethanol). The DNA pellet then re-suspended in 30 μl of T.E. containing RNAase (900μl of T.E. + 100μl of RNAase).
A small sample was then run on a 0.8% agarose gel (2μl DNA, 8μl of sterile water, 2μl of loading buffer).

2.2.2.8- *SacI / EcoRI* double digest

The following were mixed:

a. 14μl of H_2O.
b. 2μl of DNA (NTG).
c. 2μl of buffer 2.
d. 1μl of *SacI.*
e. 1μl of *EcoRI.*

Incubated at 37°C for 1-2 hours, then 4μl of loading buffer were added and run on a 1.2% agarose gel using the λ *Hind III* DNA marker.

2.2.2.9- Insert size calculation

After digesting out the insert DNA with *SacI/EcoRI*, a standard curve was drawn using the standard sizes of the λ *Hind III* DNA marker and the distance in millimetres for each size.

114 NTG clones have been plasmid prepped, digested with *SacI/EcoRI*, and the size of insert calculated for each one. Out of the 114, 5 NTG's were chosen (with different size of insert) partially or fully sequenced, each had been sequenced using RV-UP and sent for Homology search.
The procedure for each NTG clone is explained in details in later sections.

2.2.2.10- Colony Blots

Colony Blots were done in order to do check if any clones had been previously isolated
NTG clones were plated onto new ampicillin plates (100 on each plate) and incubated at 37°C overnight.
A nylon membrane was cut into pieces slightly smaller than the agar plates, and each membrane was labelled for identification purposes, then placed label side down on to the re-streaked plate, making sure no air bubbles were trapped. 3 asymmetrical marks were made by stabbing through the membrane and the agar (with a sterile needle) for later realignment, from this point the membrane was handled with blunt forceps.

The membranes were then transferred (bacteria side up) to a piece of whatman 3MM paper soaked in 10% SDS and left for 3 minutes (This will break the cell wall and release the DNA). After 3 minutes the excess SDS was wiped off and the membrane transferred to 3MM paper soaked in Denaturing solution, and left for 5 minutes. Denaturing solution converts the double strand DNA to single strands so it sticks to the membrane.
After the excess denaturing solution was wiped off and the membrane transferred to a piece of 3MM paper soaked in neutralizing solution for 5 minutes and finally to a 3MM paper soaked in 2XSSC for 5 minutes.
The membranes were then transferred to a piece of dry 3MM paper and allowing to air dry for 30 minutes, wrapped in cling film, and placed bacteria side down on the UV transilluminator for 5 minutes. To store at room temperature they were wrapped in tin foil.

2.2.2.11- Colony Hybridization

Colony Hybridization was done to check if any of the 400 NTG clones have picked the same insert as the one in clone NTG-21.

NTG-21 is thought to be a target gene for N-*myc*, and has been isolated from a NTG library using the cell line SK-N-Be2c (This cell line has 150-fold N-myc oncogene amplification). NTG-21 contains an 803bp insert and shows 65% identity over 415bp to the 3' end of intron 1 of the human aromatic L-amino acid decarboxylase gene (DOPA decarboxylase) gene, and as mentioned before, one of Neuroblastoma's distinguished molecular characteristics is the increased catecholamine metabolism and high expression of DOPA decarboxylase. This appears to be a link between N-Myc expression and catecholamine metabolism.

I-

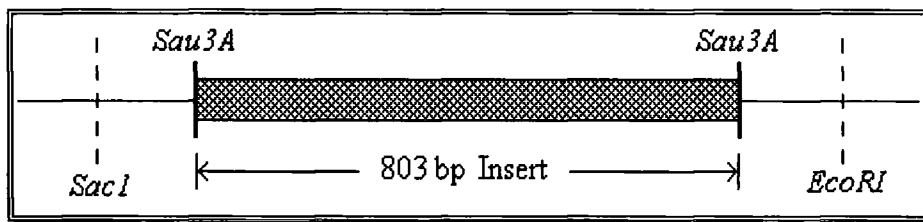


Figure 2.7, NTG-21

NTG-21 was digested first with *SacI/EcoRI* as follows:

a. 17µl DNA.
b. 4µl *SacI.*
c. 4µl *EcoRI.*
d. 8µl buffer (2).
e. 47 H_2O.

After 3 to 4 hours in 37°C, 16µl of loading buffer was added, and run on a 0.8% agarose gel. The fragment was then isolated from the gel as follows.

II- Isolation of DNA from Agarose gels

After electrophoresis, the band of interest was cut using a clean razor blade, trimming off all excess gel. The slice of gel was then put into a labelled eppendorf and 800µl of sodium iodide added, then placed in a water bath at 70°C for 5 minutes, or until the agarose had completely melted.

After leaving it for 5 minutes at room temperature 5µl of silica fines was added (to re-suspend) and vortexed for 1 second, the tube was then inverted continuously for

After the *SacI/EcoRI* digest isolation it was then digested with *Sau3A* as follows:

a. 48µl DNA.
b. 4µl *Sau3A*.
c. 8µl buffer.
d. 40µl H_2O.
e. 3 to 4 hours in 37°C heat block, run on 0.8% agarose gel and then isolating the 800pb band using the previous method (isolation of DNA from agarose gels).
f. 5µl of the 800 NTG-21 was run on a 0.8% agarose gel. (see figure 2.9)

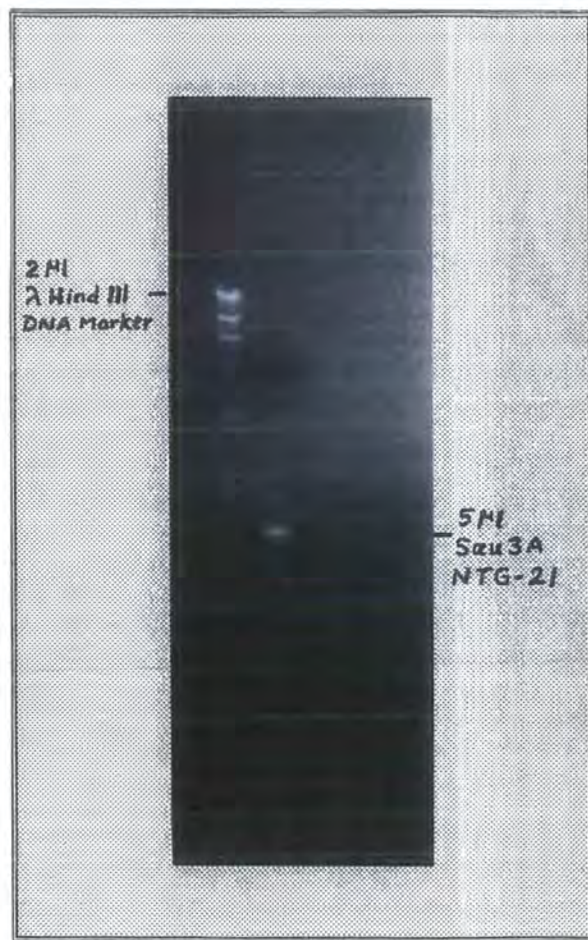


Figure 2.9, 5µl of *Sau3A* cut NTG-21 on 0.8% Agarose gel.

Colony hybridization was then carried out using the multiprime DNA labelling reaction kit.

2.2.2.12- Glycerol stocks

Glycerol stocks were prepared in order to store all the 400 NTG clones until needed.

Method:

5ml of broth was incubated in a 15ml falcon tube, with bacteria (NTG clone) and grown overnight with shaking.

The cells were pelleted by centrifugation for 2min, pouring off the medium and resuspending the pellet in 0.5ml of fresh medium, and then transferred to sterile eppendorf tube and adding 0.5ml of slightly warmed sterile glycerol, shaking well to mix thoroughly, and stored at -70°C (Aseptic technique was used throughout and close to the bunsen flame as possible).

2.2.2.13- Protein analysis

This was done using protein extract from SK-N-SH cells and SK-N-Be2c to compare the protein levels in both cells.

1. Bradford protein assay

a- Prepared 100ml of a 1:5 dilution of Bio-Rad protein assay Dye reagent concentrate (20ml dye + 80ml H_2O).

b- Dissolved 11mg of Bovine serum albumin fraction V in 10 ml of H_2O.

c- Prepared the following dilutions in eppendorf tubes:

Tube No.	Volume of BSA (1.1 mg/ml) (μl)	Volume of Assay buffer (μl)	Protein (μg)
1	0	500	0
2	45.5	454.5	5
3	90.9	409.1	10
4	159.1	340.9	17.5
5	227.3	272.7	25
6	318.2	181.8	35
7	409.1	90.9	45
8	500	0	55

Table 2.1.

d- Prepared 50μl of each standard (above) in to duplicate test tubes, 2.5ml of diluted dye was added and incubated for 10-15 minutes at room temperature, read OD_{595nm}, graph was produced·

For sample between 5 and 50μl of protein extract was added to a 2.5ml of diluted dye, incubated for 10-15 minutes at room temperature, and then read OD_{595}.

The value for protein μg was read from the standard curve and calculated the value for μg protein/μl (explained in chapter three).

II. Protein analysis

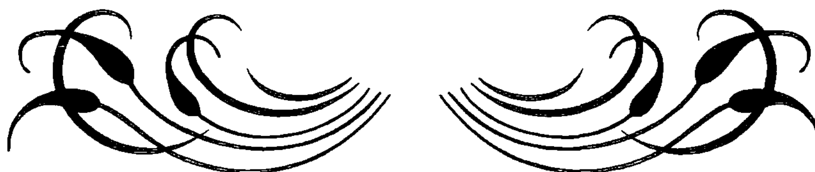
Method:

The SDS-Polyacrylamide gel was performed using a resolving gel, on top the stacking gel overlaid with 0.1% SDS which is the poured off once gel sets.
The samples were prepared as follows:

SK-N-SH	SK-N-Be2c
10μl N/E	6.25μl N/E
+	+
10μl Dye (SDS buffer)	10μl Dye (SDS buffer)

Both samples were boiled for 5 minutes to allow the SDS to denature the N/E (Neuroblastoma protein extract).

A 16μl of each was loaded in to the SDS gel wells with 16μl of the standards (molecular size marker from Sigma). After 1-2 hours, the gel was stained by immersing the gel in a minimum of 5 volumes of staining solution, and placed on a slow rotating platform for at least 4 hours at room temperature, and then de-stained by soaking it in the destaining solution for 4-8 hours on a rocking platform, changing the destaining solution three or four times. The result is in chapter III.



3. Results

3.1- NTG clones and homology search

The aim of the research was to identify potential N-Myc target genes (NTGs) using the cell line SK-N-SH which has no N-*myc* amplification. This was done by generating an NTG library /(see Materials and Methods)

The NTG library was prepared by isolating genomic DNA from SK-N-SH cells, digesting to completion with *Sau3A*, and then mixing with nuclear extract from SK-N-SH cells, to allow N-Myc to bind to DNA with an E-box (either high, or low affinity).

N-myc antibody was added and DNA-protein complexes immunoprecipitated to isolate DNA with N-Myc bound. DNA fragments were then inserted into cloning vectors [*BamH1* cut pBluescript], then sure cells transformed with vector, see figure 2.5 .

This procedure was repeated four times using different amounts of material each time to generate the genomic library with 400 white clones. Plasmid DNA was then isolated from 114 white transformants, digested with *EcoRI/SacI*, and run on agarose gels to estimate the insert sizes, (figures 3.1.a & b shows examples of these gels) using λ*HindIII* DNA marker.

Figure 3.1a,

Figure 3.1b,

2µl of a number of plasmid DNA digested with *EcoRI/SacI* on 0.8% Agarose gel.

The size of insert in 114 clones were calculated for the second time using the 100bp ladder as a marker to give accurate reading (photos not included).

NTG clone	Approx. size	NTG clone	Approx. size
1	1700 bp	26	1220 bp
2	680 bp	27	735+660+195+185
3	1400+720+320+180	33	1350+140 bp
5	1450 bp +390 bp	36	1550 bp
7	2250 bp	40	1550 bp
8	1300+760+240 bp	45	1450 bp +370 bp
9	1490 bp	48	1600 bp
12	1450 bp	52	1500 bp
13	740+660+200+190	54	1600 bp
17	1150 bp +420 bp	56	1400 bp
19	1450 bp		

Table 3.1.

Out of the 114 NTG's, the following were chosen randomly: NTG+13, NTG+5, NTG+8, NTG+33. From each, pure plasmid was prepared and partially or fully sequenced.

(see figure 3.2.a, b, and c).

Figure 3.2a,

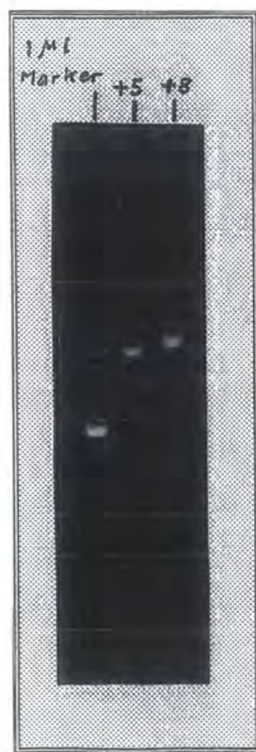


Figure 3.2b,

Figure 3.2c,

1μl of pure plasmid (+13, +5, +8, +33) on 1.2% Agarose gel.

Each insert was partially or fully sequenced using the reverse (RV) and universal (UP) primers then compared to DNA databases at SEQNET Daresbury.

Clone No.	Total insert size sequenced	Restriction map	Sequence	E.Box	Homology search
5	1137 bp	✓	Partial	✓	✓
8	956 bp	✓	Partial	*	✓
13	1014 bp	✓	Partial	*	✓
33	1210 bp	✓	RV.	*	*

Table 3.2.

NTG+5

- Insert size approximately 1137bp.
- Sequenced initially using RV and UP shown in figure 3.5, (DNA sequence and predicted restriction map as in figure 3.4.I & II).
- Homology search using the Daresbury databases shows the following:

```
5u
gb_pr:hummha3c7b

LOCUS       HUMMHA3C7B   8892 bp   DNA            PRI        31-MAR-1995
DEFINITION  Homo sapiens MHC class I alpha chain gene, leader exon, alpha 1 and
            2 exons (HLA A3, C7, B7, BfS, C4A3, C4B1, DR2, DQ6).
ACCESSION   L29411
KEYWORDS    cell surface antigen; cell surface glycoprotein; class I gene;
            integral membrane protein; major histocompatibility complex. . . .

SCORES           Init1:   481 Initn:   738 Opt:   592
            70.9% identity in 316 bp overlap
```

Figure 3.3.I,

```
5u /rev
gb_pr:humtnfz

LOCUS       HUMTNFZ      10950 bp    DNA            PRI        25-MAR-1992
DEFINITION  Human tumor necrosis factor receptor, 3' flank.
ACCESSION   M33L45
KEYWORDS    tumor necrosis factor receptor.
SOURCE      Homo sapiens (library: E80pcD) DNA.
  ORGANISM  Homo sapiens . . .

SCORES           Init1:   250 Initn:   733 Opt:   566
            73.7% identity in 281 bp overlap
```

Figure 3.3.II,

```
5u /rev
gb_pr:humbmyh7

LOCUS       HUMBMYH7     23438 bp    DNA               PRI        31-OCT-1994
DEFINITION  Homo sapiens (clones lambda gMHC 1,2,3 and 4) beta-myosin heavy
            chain (MYH7) gene, complete cds.
ACCESSION   M57965 M30603 M30604 M30605 M57747 M57748 M57749
KEYWORDS    beta-myosin heavy chain; cardiac myosin heavy chain.
SOURCE      Human Leukocyte DNA. . . .

SCORES           Initl:   296 Initn:   326 Opt:   660
            75.9% identity in 294 bp overlap
```

Figure 3.3.III,

```
gb_pr:humatpsyb

LOCUS       HUMATPSYB    10186 bp    DNA               PRI        31-OCT-1994
DEFINITION  Human ATP synthase beta subunit (ATPSB) gene, complete cds.
ACCESSION   M27132
KEYWORDS    ATP synthase beta subunit.
SOURCE      Human fetal liver DNA, clone g-beta-lambda 1.
  ORGANISM  Homo sapiens . . .

SCORES           Initl:   306 Initn:   724 Opt:   512
            69.1% identity in 330 bp overlap
```

Figure 3.3.IV,

```
5u /rev
gb_pr:hshw2

LOCUS       HSHW2        22677 bp    DNA               PRI        12-MAY-1995
DEFINITION  Human DNA sequence from cosmid HW2, Huntington's Disease Region
            chromosome 4p16.3.
ACCESSION   Z49250
KEYWORDS    4p16.3; CpG island; helix-loop-helix; transcription repressor.
SOURCE      human. . . .

SCORES           Initl:   346 Initn:   777 Opt:   645
            70.2% identity in 372 bp overlap
```

Figure 3.3.V,

```
5r
gb_pr:numtcrb

LOCUS       HUMTCRB      684973 bp    DNA              PRI       14-JAN-1995
DEFINITION  Homo sapiens (clones: K41A, K35, K23, K56, X218, G54, H137, H18,
            H18/G15gap, G15, X14, A27, A212, A14, H7.1, H12.18, H130, Alo,
            C215, G1, C68, C21, X11, X6A, CBG1, CBG1/C29gap, and C29) germline
            T-cell receptor beta chain, complete gene.
ACCESSION   L36092 . . .

SCORES           Init1:   136 Initn:   254 Opt:   306
            70.6% identity in 137 bp overlap
```

Figure 3.3.VI,

```
5u /rev
gb_pr:humldlrac

LOCUS       HUMLDLRAC     2382 bp    DNA             PRI        07-JAN-1995
DEFINITION  Human acetylated low density lipoprotein (AcLDL) receptor (LDLR)
            gene, promoter and exon 1.
ACCESSION   M93189
KEYWORDS    LDL receptor; acetylated LDL receptor;
            low density lipoprotein receptor. . . .

SCORES           Init1:   231 Initn:   729 Opt:   736
            69.5% identity in 469 bp overlap
```

Figure 3.3.VII,

```
5u
gb_pr:huminsrd

LOCUS       HUMINSRD      7240 bp    DNA             PRI        06-JAN-1995
DEFINITION  Human insulin receptor (allele 2) gene, exons 14, 15, 16 and 17.
ACCESSION   M29930 M22986 M24639
KEYWORDS    Alu repeat; insulin receptor.
SOURCE      Human lymphoblastoid cell line or peripheral blood lymphocyte DNA.
  ORGANISM  Homo sapiens . . .

SCORES           Init1:   549 Initn:   802 Opt:   703
            75.4% identity in 317 bp overlap
```

Figure 3.3.VIII,

```
5u
gb_pr:humpselect

LOCUS        HUMPSELECT   4855 bp    DNA          PRI       03-JAN-1993
DEFINITION   Homo sapiens P-selectin CD62 (GRMP) gene, 5' end and promoter
             region.
ACCESSION    L01574
KEYWORDS     P-selectin.
SOURCE       Homo sapiens (tissue library: EM3L3) colonic mucosa DNA. . . .

SCORES           Initl:   264 Initn:   723 Opt:   676
             70.0% identity in 400 bp overlap
```

Figure 3.3.IX,

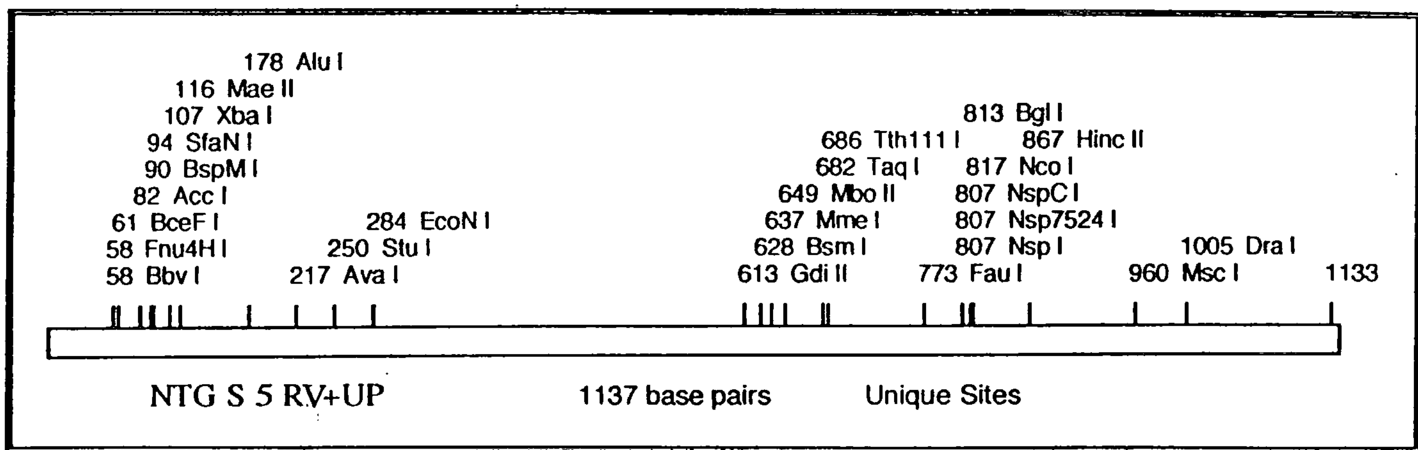


Figure 3.4.I, Restriction map of NTG+5 sequenced region.

```
          |  10      |  20      |  30      |  40        |  50      |  60      |  70      |  80
   1 GATCACGGGC GTTGGTGAAG GATTGGCCCA AGACGGGGTG  GAATCTGGAA CCACTTTGCT GCCGTAATAC AGGGCCCCTG 80
  81 CGTAGACAAA CCTGCATCTG TTAGTTTCTA GATTTACGTT  TCCACGGAAA TTGCTCACCA TGTCAAGAAT CTGTTTTATG 160
 161 CACCCCCAGG TATTGTCAGC TTAGGTTTTT CTCCTCTGTT  TTTATTCCAG AGTTCTCTCG GGCTCCCTCT GTCATGTTGC 240
 241 TGAGGGAGGA GGCCTGTGAT ACAGTTCTCT CTGGAGTCTG  GAACCTGAGC AAGGGTGGCA GTGGGAGGAA CCGCTTTATA 320
 321 GGGCCGGGGA AGGGCCCCGG TGACTCCAGT GCAGGCCGGG  GAGCGAAGTC TGGGAGAAGG CCCAGAATAT AATATATAAT 400
 401 ATTATATAGA ATATAATATA TAATATTATA TAGAATATAA  TATATAATAT TTTATAGAAT ATAATATATA ATATTTTATA 480
 481 GAATATAATA TATAATATTT TATAGAATAT AATATATAAT  ATTTTATAGA ATATAATATA TAATATTTTA TAGAATATAT 560
 561 ATATAATATA TAGAATATAA TATATGATAT T***CAATAG  GATAAGTTCT GCTGGCCGAG TGGAGTGGAA TGCAAGTCCA 640
 641 ACTAGAGGGA AGAATAATAT AAAACTTCCA ATAATTTTTT  TTCGAGACGA GGTCTTATTC TGTCACTCAG GCTGGAGTGC 720
 721 AGTGGTGTGA TTGGCTCACT GCAACCTTCA CCTTCCAGGC  TCAAGCAATC CTCCCGCCTC AGCCTCCTGA GTATTTGGGA 800
 801 CTATGGGCAT GTGCCACCAT GGCTGGCTAA TTTTAGTATT  TTTTTCTAAA GATGGGGTTT TGCTATGTTG ACCAGGCTCG 880
 881 TCTCATGTTC ATAGGCTCAA GCAATTAACC CACTCCAGCC  TCCCAAAGTG CTGGGATTAT AGGCATGAGC CACCATGACT 960
 961 GGCCAGGTTC TAGTAATTGT AATATAGGAT TTGCTTATGT  ATTTTTTAAA AATGAGTGAT AGTGAGTAAC CAATTAAGTA 1040
1041 TTCTGTTTTA TTAGATGTAT TGAAAACAGT GATATGGTCA  GGCACGGTGG CTCACGCCTG TAACCCCAGC ACTTTGGGAG 1120
1121 GCCAAGGCAA GTGGATC                                                                       1137
          |  10      |  20      |  30      |  40        |  50      |  60      |  70      |  80
```

Figure 3.4.II, DNA sequence of NTG+5 RV+UP (*** = unsequenced region).

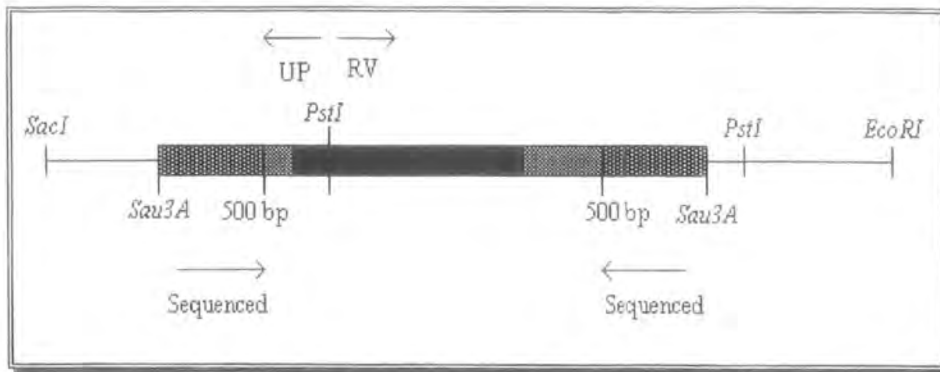


Figure 3.5, NTG+5 sequenced regions.

- NTG+5 was digested using the *SacI/EcoRI* double digest (see method 2.2.2.8), to check if it possesses *SacI* or *EcoRI* sites. (see figure 3.6)

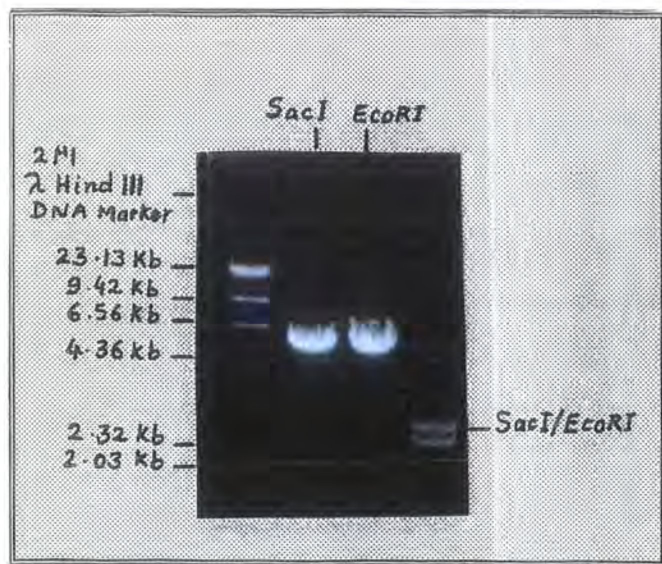


Figure 3.6, 2µl of NTG+5 *SacI, EcoRI, SacI/EcoRI* digest on 1.2% Agarose gel.

Digest results:

SacI ⇒ 5.0 Kb

EcoRI ⇒ 5.0 Kb

SacI/EcoRI ⇒ 2.5Kb (vector) + 2.3Kb (insert)

Summary

The NTG5 insert does not contain *SacI* or *EcoRI* sites, the only sites present in NTG+5 are those in the polylinker.

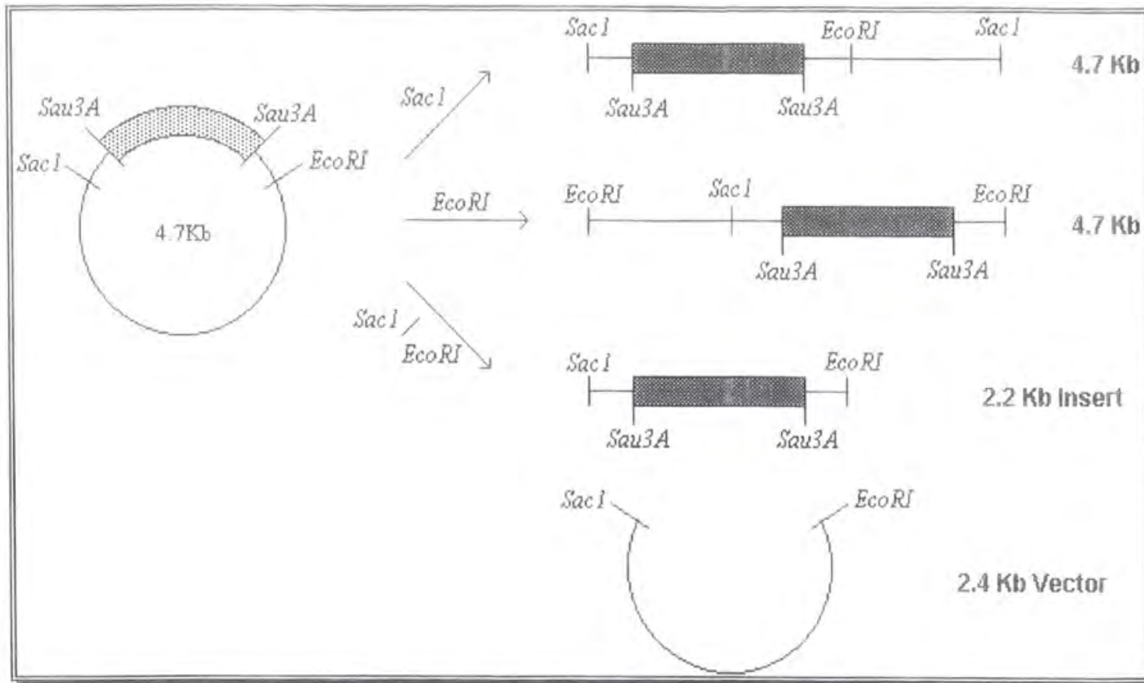


Figure 3.7, *SacI/EcoRI* digestion of NTG+5

- 10 different enzymes were used for digesting NTG+5 to identify the presence of restriction sites (see method 2.2.2.8).

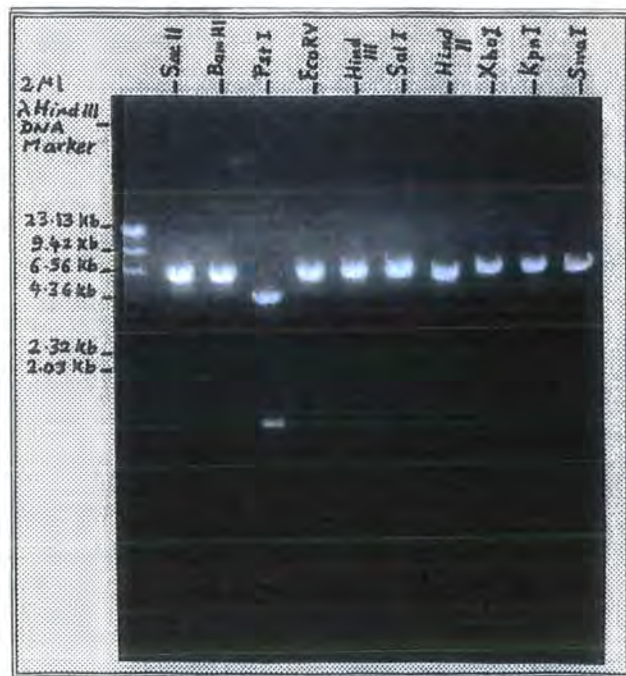


Figure 3.8, 2µl of each of the 10 different digests of NTG+5 on 1.2% Agarose gel.

PstI excises a fragment of 1.4Kb, *HindII* (*HincII*) excises a fragment of 290bp Looking at the restriction map, there is a *HincII* site at 867, therefore the total length 1137-867 ⇒ 270bp corresponds well with the above digest.

Summary - *HindII (HincII)*

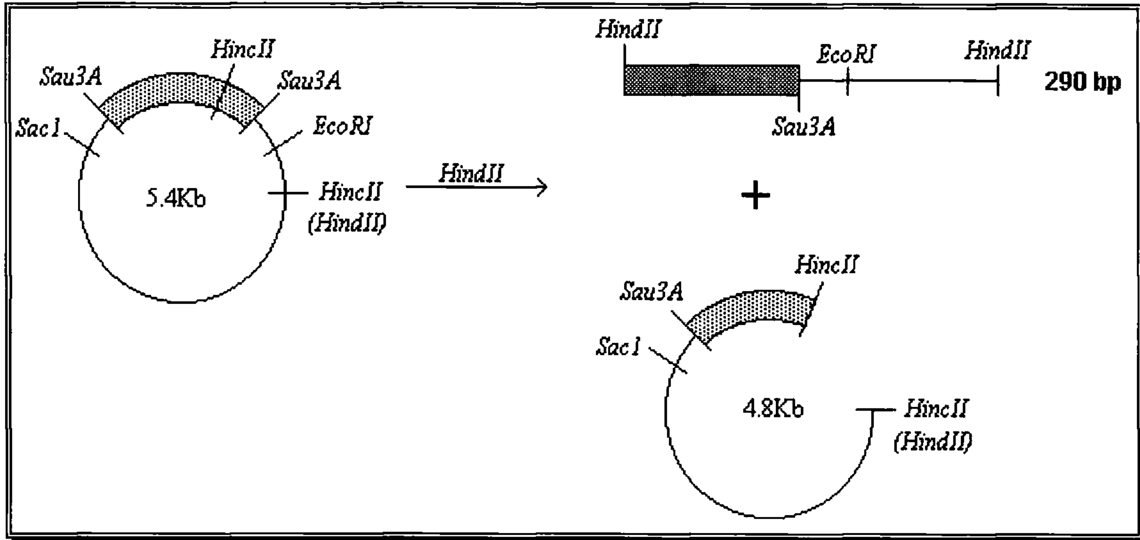


Figure 3.9, *HindII* digestion of NTG+5.

Summary - *PstI*

⇒ The *Sau3A* fragment in NTG+5 is approximately 2.2Kb from the previous *SacI/EcoRI* digests.

⇒ When NTG+5 is cut with *PstI* it excises a 1.4Kb fragment, therefore:

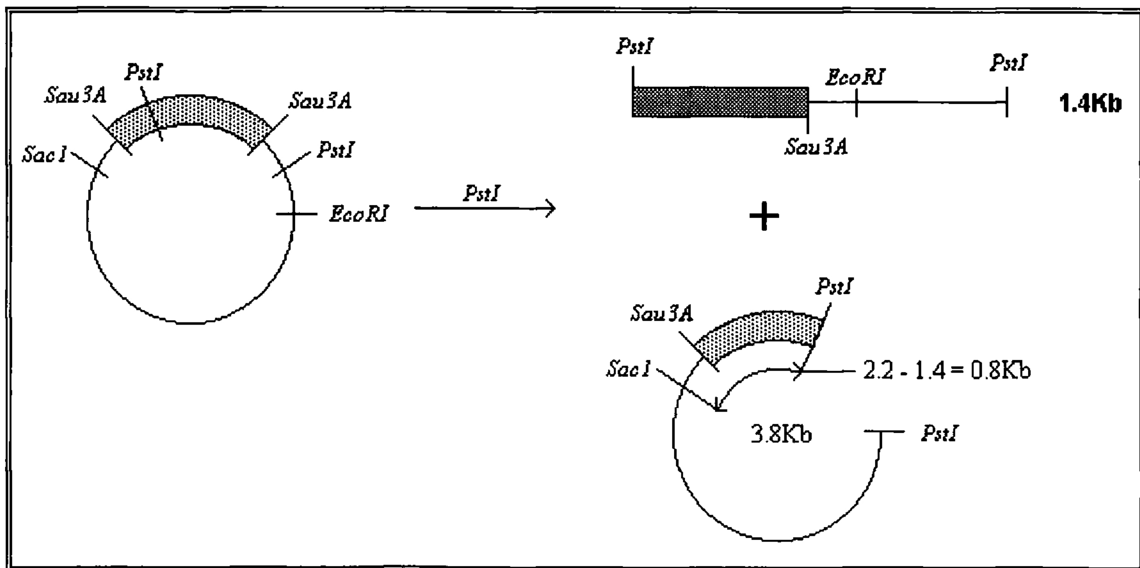


Figure 3.10, *PstI* digestion of NTG+5.

* In order to make sub-clones of NTG+5 for sequencing the following was done:

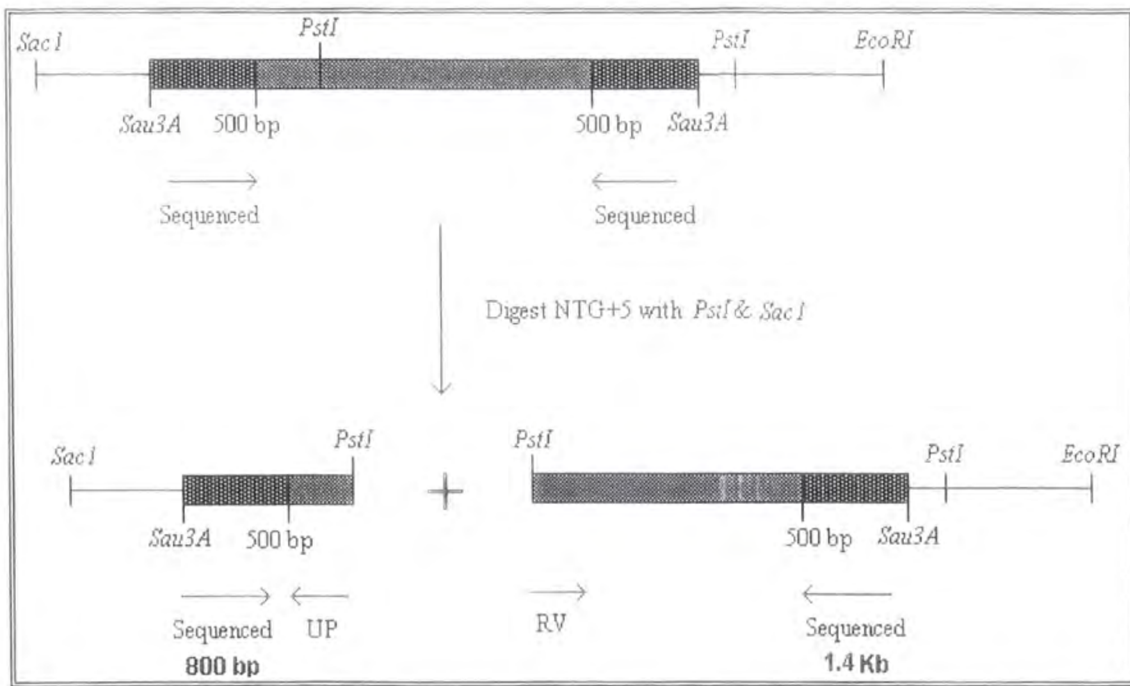


Figure 3.11, Sequencing strategy for NTG+5 sub-clones.

⇒ The 800bp and 1.4Kb bands were isolated from agarose gels using method 2.2.2.11 (see figure 3.12 & 3.13)

Figure 3.12, 100µl of NTG+5 *SacI/PstI* digest on 1.2% gel.

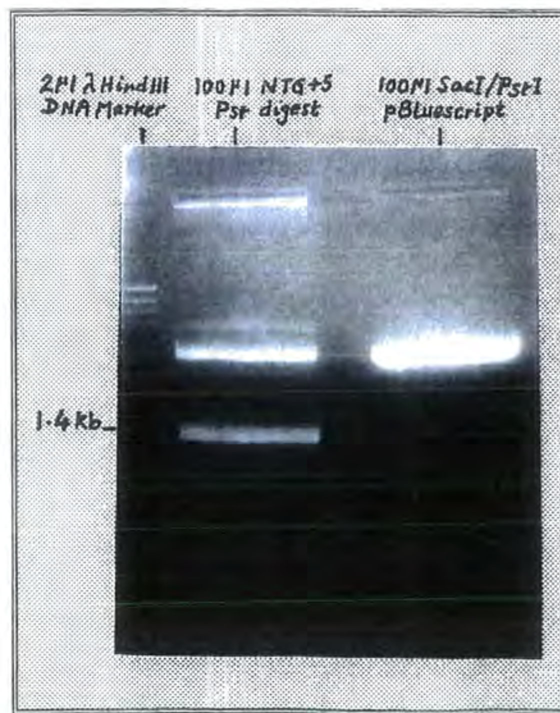


Figure 3.13, 100µl of NTG+5 *PstI* digest 100µl of *SacI/PstI* cut pBluescript on 1.2% gel.

⇒ The 800bp fragment was ligated into *SacI/PstI* cut pBluescript (see figure 3.14):

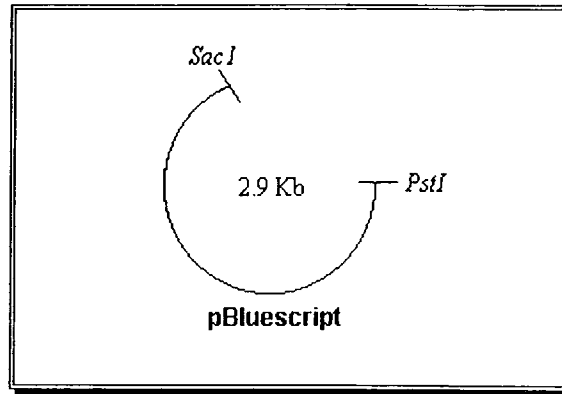


Figure 3.14, *SacI/PstI* cut pBluescript.

⇒ Sequenced using the UP and homology search using the Daresbury database showed the following (see figure 3.16.I & II for predicted restriction map and sequence:

```
5sc /rev
gb_pr:humcolrepu

LOCUS       HUMCOLREPU     801 bp    DNA              PRI       01-NOV-1994
DEFINITION  Human variable number tandem repeat (VNTR) region, allele 17R1 3'
            to collagen type II (COL2A1) gene.
ACCESSION   L10171
KEYWORDS    VNTR; collagen; collagen type II;
            variable number tandem repeat allele. . . .

SCORES           Init1:    334 Initn:    343 Opt:    686
            70.5% identity in 661 bp overlap
```

Figure 3.15.I,

```
5sc
gb_pr:humtcrb

LOCUS       HUMTCRB     684973 bp    DNA              PRI       14-JAN-1995
DEFINITION  Homo sapiens (clones: K41A, K35, K26, K56, X21B, G54, H137, H18,
            H18/G15gap, G15, X1A, A27, A212, A14, H7.1, H12.18, H130, A16,
            C215, G1, C68, C21, X11, X6A, CBG1, CBG1/C29gap, and C29) germline
            T-cell receptor beta chain, complete gene.
ACCESSION   L36092 . . .

SCORES           Init1:    161 Initn:    388 Opt:    585
            63.7% identity in 630 bp overlap
```

Figure 3.15.II,

```
5sc /rev
gb_pr:humifnb2a

LOCUS       HUMIFNB2A     544 bp    DNA              PRI       08-NOV-1994
DEFINITION  Human interleukin 6 (IL6) gene, 3' flank.
ACCESSION   J03049
KEYWORDS    interferon-beta-2; interleukin 6.
SOURCE      Human DNA.


  ORGANISM  Homo sapiens . . .


SCORES           Initl:   212 Initn:   342 Opt:   744
            72.6% identity in 493 bp overlap
```

Figure 3.15.III,

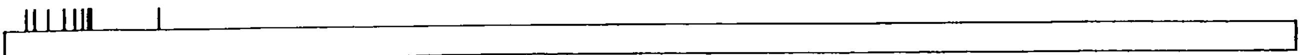
```
5sc
gb_pr:hscolii2

LOCUS       HSCOLII2      2461 bp    DNA              PRI       04-NOV-1994
DEFINITION  Human alpha 1(II) collagen gene 3' untranslated region.
ACCESSION   X02670
KEYWORDS    AT rich region; collagen; collagen alpha 1 type II;
            collagen type II; tandem repeat.
SOURCE      human. . . .


SCORES           Initl:   153 Initn:   221 Opt:   814
            70.3% identity in 632 bp overlap
```

Figure 3.15.IV,

41 Rma I
35 BsiY I
25 HgiA I
25 Bsp1286 I
25 ApaL I
18 Ava I
12 Nla IV
12 Ban I



5 .800.up   739 base pairs   Unique Sites

Figure 3.16.I, Restriction map for 800bp *SacI/PstI* sub-clone of NTG+5

```
      |  10      |  20      |  30      |  40        |  50      |  60      |  70      |  80
  1 GAAAGTATCC AGGCACCCCC GAGGGTGCAC ACACCCAGGG  CTAGGAGACC TCAGGTTTAT ACGTCTCTTT CTTTTCTCTT 80
 81 TCTTGGGGAC TGGTTCTACC GAAACATATA TACATATATG  ATATATACAC ACACGTGTAT ATATCTCATA TATATATATA 160
161 TATATATATA CACGTGTATA TATATATATA TAACAGATAA  CGTATAATTA TAATTATATT ATCTGTAATA TATAATTATA 240
241 TATATTTTGT ATTCTATATA TTATATACTA TGTATTCTAT  ATAATATATA TTATGTATTC TATATAATAT TATATATTAT 320
321 ATTCTATATT CTATATTATA TTCTATATTC TATAGAATAT  TCTATATTAT ATTCTATATT CTATAGAATA TTCTATATTA 400
401 TATTCTATAT TCTATAGAAT ATTCTATATT ATATTCTATA  TTATATTCTA TATATTCTAT ATAATATTAT ATATTATATT 480
481 CTATATAATA TTATATATTA TATTCTATAT ATTATATATT  ATATTCTATA TATTATATAT TATATTCTAT ATATTATATA 560
561 TTATATTCTA TATATTATAT ATTAATATTC TATATATTAT  ATATTATATT CCTATATATT ATATATTATA TTCCTATATA 640
641 TTATATATTA TATTCTATAT ATTATATATT ATATTCTATA  TATTATATAT TATATTCTAA TAAAATATTA TATATTAATA 720
721 TCCCATAAAA TATTATATA                                                                    739
      |  10      |  20      |  30      |  40        |  50      |  60      |  70      |  80
```

Figure 3.16.II, DNA sequence for 800bp *SacI/PstI* sub-clone of NTG+5.

- The 1.4 Kb *PstI* fragment was ligated into *PstI* cut pBluescript:

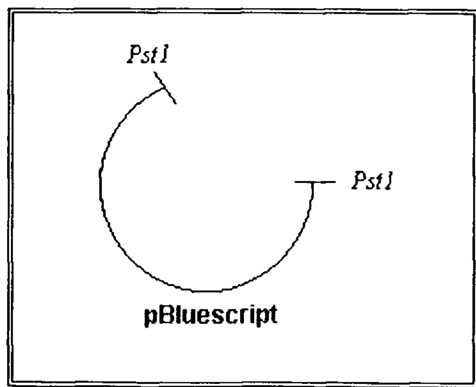


Figure 3.17, *PstI* cut pBluescript.

- To determine the orientation of the fragment, 1μl of 1.4Kb *PstI* was digested with *HindII*, as the *PstI* fragment can ligate into the vector in either orientation to give 2 possible recombinants, see figure 3.18:

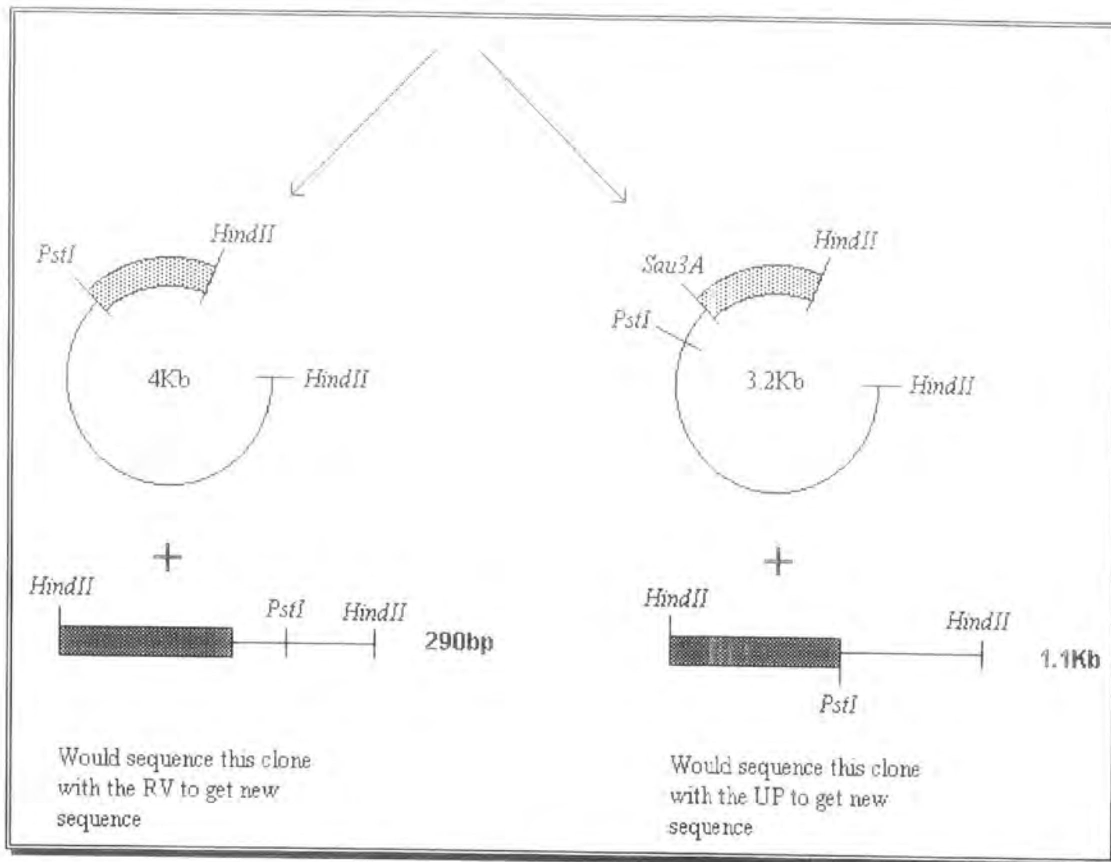


Figure 3.18. Possible recombinants.

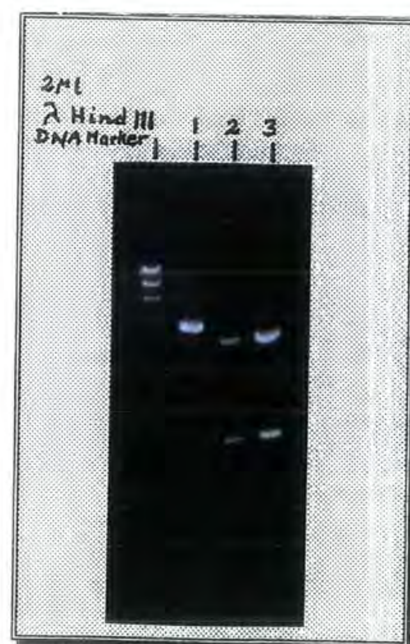


Figure 3.19. 2µl of 1.4Kb *PstI* sub-clone on 0.8% Agarose gel.

3 sub-clones of 1.4Kb *PstI* were tested

- 1 had 4Kb + 290bp
- 2 had 1.11Kb + 3.2Kb

The first sub-clone (4Kb + 290bp) was sequenced using RV. The homology search shows the following (see figure 3.20 I & II for predicted restriction map and sequence):

```
5sc3 /rev
gb_pr:hsrb2p130

LOCUS        HSRB2P130     3249 bp    RNA              PRI        14-MAR-1994
DEFINITION   H.sapiens mRNA for Rb2/p130 protein.
ACCESSION    X74594
NID          g397147
KEYWORDS     E1A binding protein; retinoblastoma-associated protein.
SOURCE       human. . . .

SCORES          Init1:    68 Initn:    111 Opt:    112
             55.7% identity in 219 bp overlap
```

Figure 3.20.I,

```
5sc3
gb_pr:humcd36a

LOCUS        HUMCD36A      2561 bp    DNA              PRI        03-MAR-1995
DEFINITION   Human antigen CD36 gene, alternate exons 2a and 2b and exon 3.
ACCESSION    L06349
NID          g691762
KEYWORDS     Alu repeat; antigen CD36.
SOURCE       Homo sapiens (tissue library: EMBL 4) DNA. . . .

SCORES          Init1:    69 Initn:    112 Opt:    69
             94.7% identity in 19 bp overlap
```

Figure 3.20.II,

```
5sc3 /rev
gb_or:s67171

LOCUS       S67171        3448 bp    mRNA              PRI       04-FEB-1994
DEFINITION  Rbr-2=retinoblastoma susceptibility gene [human, HeLa S3 suspension
            cells, mRNA, 3448 nt].
ACCESSION   S67171
NID         g453131
KEYWORDS    . . . .

SCORES           Init1:    68 Initn:   111 Opt:   112
            55.7% identity in 219 bp overlap
```

Figure 3.20.III,

```
gb_pr:hshcgii

LOCUS       HSHCGII       3642 bp    RNA               PRI       13-AUG-1994
DEFINITION  H.sapiens HCG II mRNA.
ACCESSION   X81001
NID         g531407
KEYWORDS    HCG II gene.
SOURCE      human. . . .

SCORES           Init1:    57 Initn:   131 Opt:    78
            53.7% identity in 231 bp overlap
```

Figure 3.20.IV,

```
5sc3
gb_pr:hsl25g12

LOCUS       HSL25G12      37153 bp   DNA               PRI       03-OCT-1995
DEFINITION  Human DNA sequence from cosmid L25G12, Huntington's Disease Region,
            chromosome 4p16.3.
ACCESSION   Z54336
NID         g1015498
KEYWORDS    4p16.3. . . .

SCORES           Init1:   115 Initn:   115 Opt:   124
            62.5% identity in 120 bp overlap
```

Figure 3.20.V,

```
            153 Xcm I                    318 NspC I
           150 Fok I                     318 Nsp7524 I
         137 ScrF I                      318 Nsp I
         137 EcoR II               275 Alu I
         137 Dsa V               261 Bsp1286 I
         137 BstN I              261 Ban II      339 Xmn I
     112 Acc I        212 Rsa I            316 Nsi I
   101 Hae III     189 PflM I     274 Pvu II                        465 SfaN I
39 BsmA I   137 BstK I   212 Csp6 I    302 Hga I          414 Ssp I
33 Xba I   100 Hae I     189 BsiY I   274 NspB II   359 Mae III    451 Mse I

      5scr2                  588 base pairs         Unique Sites
```

Figure 3.21.I, Restriction map of 1.4Kb *PstI* sub-clone of NTG+5.

```
         |   10     |   20     |   30     |   40      |   50     |   60     |   70     |   80
  1 CTCAGAGAGC ATGGGCTGTT TGAAAGGCAA TTTCTAGAGT  CTCAGTTTAG ACACAGACTC GTTCCTAAAG AAAGCCTTTT 80
 81 CTCCTAGTTT ACAAGCAGAA GGCCAGTTCT TGTCTACTCC  CCTCCCCCAG AGTCCTCCAG GAAGCAAAAC ATCCATTGGA 160
161 ATACTGGTGT AGGGCAAGAA ACCTGATTCC AATGATTGGG  AGCAAGTGCT TGTACAAACC ACAGGGCTTC CAATTCTTTG 240
241 CTTTCAACAA AATTGAAGAA GGGCTCACAT TGTCAGCTGA  TAAATTATAC AAAATTTTTT GGACGCTAGG TCTGTATGCA 320
321 TGTTTTGGCA TATAATATGA AAGAGTTCAG ATAAGCGAGT  AACTGAAATA TAACCAGACA CTTGCCATTT CTATCCACTT 400
401 GTCTATACCA ACCAATATTT ATCAGGGACT TCTGCCCATA  CCAAATTCTC TTAAAAATAA AATTGATGCC GAATCCTGTT 480
481 CTCTTTCAAT CCCATTCGTA AATAATACTT ACCCATGAAT  ACATGAAAAA CAGTCCCATT TTTCCCAGTA AGAAATTATA 560
561 GATTATTTCC CTCCAGTTTT ACTTCTTC                                                          588
         |   10     |   20     |   30     |   40      |   50     |   60     |   70     |   80
```

Figure 3.21.II, DNA sequence for 1.4Kb *PstI* sub-clone of NTG+5.

<u>NTG+8</u>

- Insert size approximately 956bp.
- Sequenced initially using RV-UP figure 3.24, (see figure 3.23.I & II for predicted restriction map and sequence).
- Homology search using the Daresbury database shows the following:

```
8r
gb_pr:hsmhcapg

LOCUS       HSMHCAPG    66109 bp    DNA             PRI       16-NOV-1994
DEFINITION  H sapiens genes TAP1, TAP2, LMP2, LMP7 and DOB.
ACCESSION   X66401 S57528
KEYWORDS    DOB gene, LMP2 gene; LMP7 gene;
            major histocompatibility complex class II; tap1 gene; tap2 gene.
SOURCE      human. . . .

SCORES          Init1:     72 Initn:    154 Opt:     74
            79.3% identity in 29 bp overlap
```

Figure 3.22.I,

```
gb_pr:hstcrbv

LOCUS       HSTCRBV     77743 bp    DNA             PRI       20-APR-1994
DEFINITION  Human V beta T-cell receptor (TCRBV) gene locus.
ACCESSION   U03115
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens . . .

SCORES          Init1:     71 Initn:    206 Opt:    144
            57.3% identity in 239 bp overlap
```

Figure 3.22.II,

```
8r
gb_pr:humvcama

LOCUS       HUMVCAMA     2396 bp    DNA             PRI       14-JAN-1995
DEFINITION  Human vascular cell adhesion molecule-1 (VCAM1) gene, exon 1.
ACCESSION   M92431
KEYWORDS    vascular cell adhesion molecule-1.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens . . .

SCORES          Init1:     62 Initn:    138 Opt:     92
            61.5% identity in 117 bp overlap
```

Figure 3.22.III,

```
gb_pr:hsrb2p130

LOCUS       HSRB2P130    3249 bp    RNA              PRI       14-MAR-1994
DEFINITION  H.sapiens mRNA for Rb2/p130 protein.
ACCESSION   X74594
NID         g397147
KEYWORDS    E1A binding protein; retinoblastoma-associated protein.
SOURCE      human. . . .


SCORES           Init1:    68 Initn:    111 Opt:    112
            55.7% identity in 219 bp overlap
```

Figure 3.22.IV,

```
gb_pr:hsu04636

LOCUS       HSU04636     9453 bp    DNA              PRI       21-DEC-1994
DEFINITION  Human cyclooxygenase-2 (hCox-2) gene, complete cds.
ACCESSION   U04636
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens . . .


SCORES           Init1:    88 Initn:    146 Opt:    120
            54.1% identity in 442 bp overlap
```

Figure 3.22.V,

```
gb_pr:humvitdbp

LOCUS       HUMVITDBP    55136 bp   DNA              PRI       27-JAN-1995
DEFINITION  Human vitamin D-binding protein (GC) gene, complete cds.
ACCESSION   L10641 L10642
KEYWORDS    group-specific component; vitamin D-binding protein.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens . . .


SCORES           Init1:    71 Initn:    143 Opt:    94
            61.9% identity in 105 bp overlap
```

Figure 3.22.VI,

```
8r
gb_pr:humretblas

LOCUS       HUMRETBLAS 180388 bp    DNA              PRI       23-NOV-1994
DEFINITION  Human retinoblastoma susceptibility gene exons 1-27, complete cds.
ACCESSION   L11910
KEYWORDS    nuclear protein; recessive oncogene; retinoblastoma gene;
            retinoblastoma protein; retinoblastoma susceptibility;
            tumor supressor gene. . . .

SCORES           Init1:     76 Initn:    161 Opt:    117
            52.6% identity in 441 bp overlap
```

Figure 3.22.VII,

```
gb_pr:humptgs2

LOCUS       HUMPTGS2     10997 bp    DNA              PRI       21-JUL-1994
DEFINITION  Human PTGS2 gene for prostaglandin endoperoxide synthase-2.
ACCESSION   D28235
KEYWORDS    prostaglandin endoperoxide synthase-2.
SOURCE      Homo sapiens (library: EMBL3) Japanese peripheral blood DNA.
  ORGANISM  Homo sapiens . . .

SCORES           Init1:     88 Initn:    146 Opt:    120
            54.1% identity in 442 bp overlap
```

Figure 3.22.VIII,

```
gb_pr:humfglbtk

LOCUS       HUMFGLBTK    69363 bp    DNA              PRI       07-JUN-1995
DEFINITION  Human ftp-3 gene, exon 1; alpha-D-galactosidase A (GLA) gene, exons
            1-7; L44L gene, exons 1-5; Bruton's tyrosine kinase (BTK) gene,
            exons 1-18.
ACCESSION   L35265
KEYWORDS    Bruton's tyrosine kinase; alpha-D-galactosidase A. . . .

SCORES           Init1:     60 Initn:    132 Opt:     82
            54.6% identity in 183 bp overlap
```

Figure 3.22.IX,

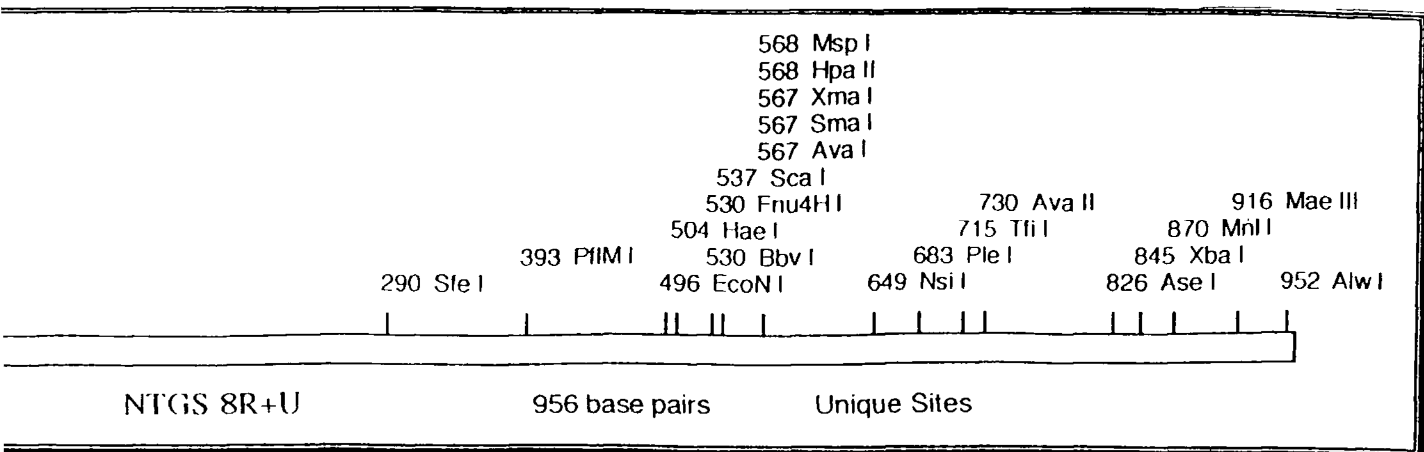


Figure 3.23.I, Restriction map of NTG+8.

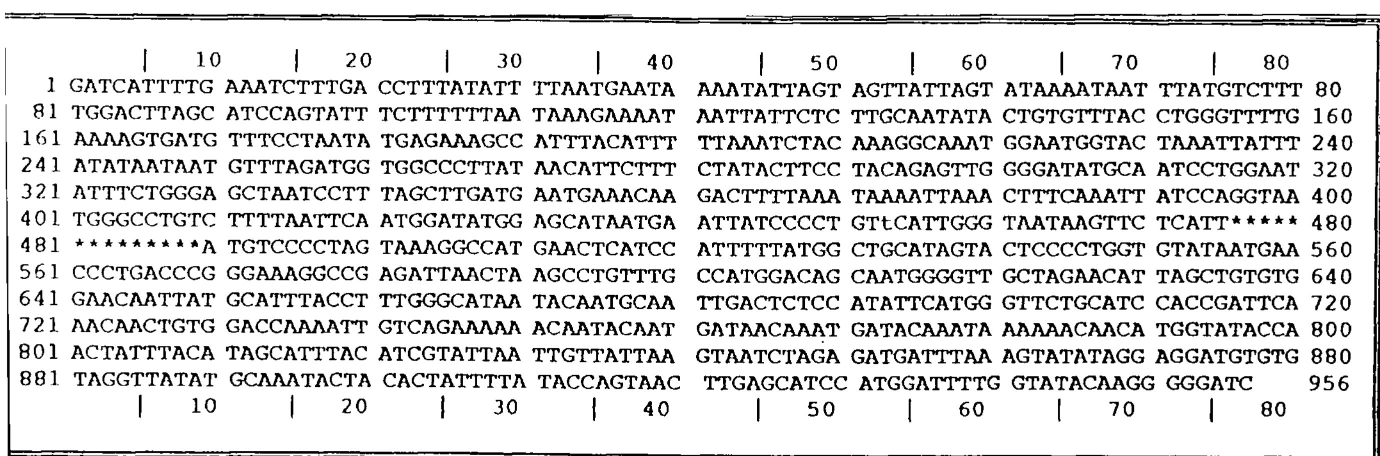


Figure 3.23.II, DNA sequence of NTG+8.

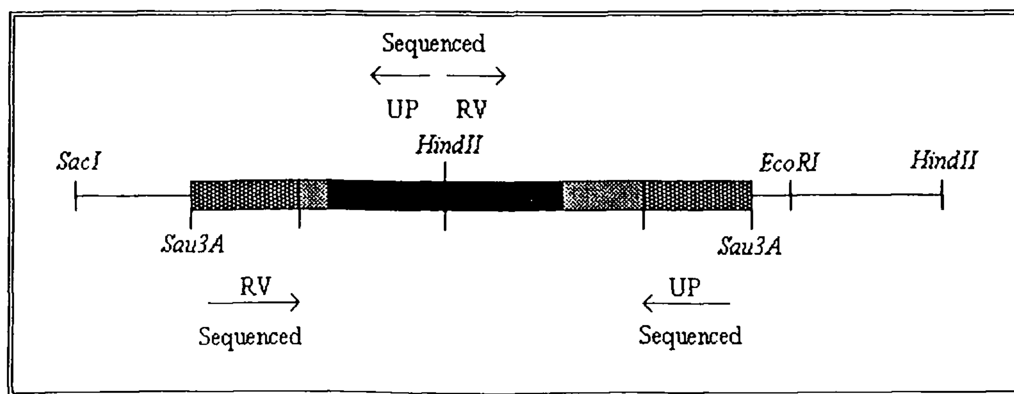


Figure 3.24, Sequenced region of NTG+8.

- NTG+8 was digested using the *SacI/EcoRI* double digest (see method 2.2.2.8), to check if it possesses *SacI* or *EcoRI* sites.

  SacI ⇒ 5.0 Kb

  EcoRI ⇒ 5.0 Kb

  SacI/EcoRI ⇒ 2.9Kb (vector) + 2.8Kb (insert)

Summary

The NTG8 insert does not contain *SacI* or *EcoRI* sites, the only sites present in NTG+8 are those in the polylinker.

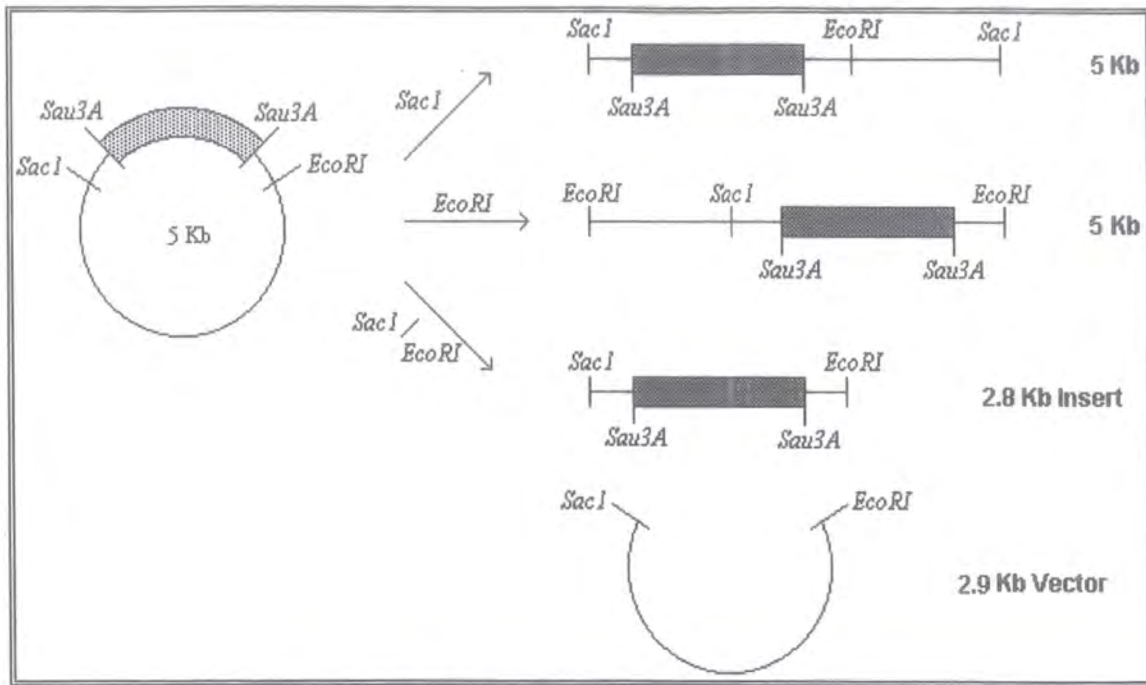


Figure 3.25, *EcoRI/SacI* digestion of NTG+8.

- 10 different enzymes were used for digesting NTG+8 to identify presence of restriction sites (see method 2.2.2.8), (see figure 3.26).

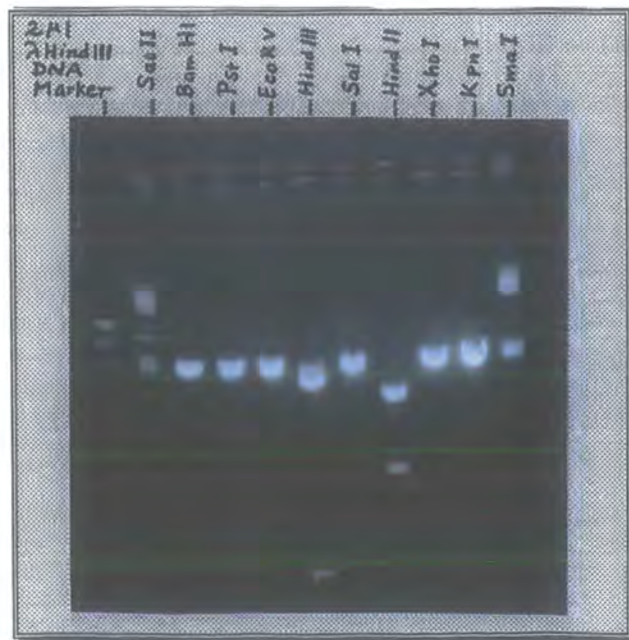


Figure 3.26, 2µl of each of the 10 different digest of NTG+8 on 1.2% Agarose gel.

HindII excises a fragment of 1.4Kb.

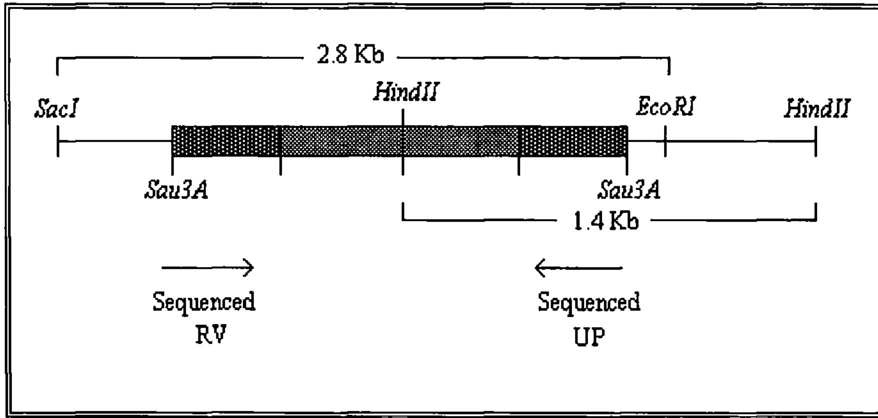


Figure 3.27, Fragments of NTG+8.

Summary - *HindII*

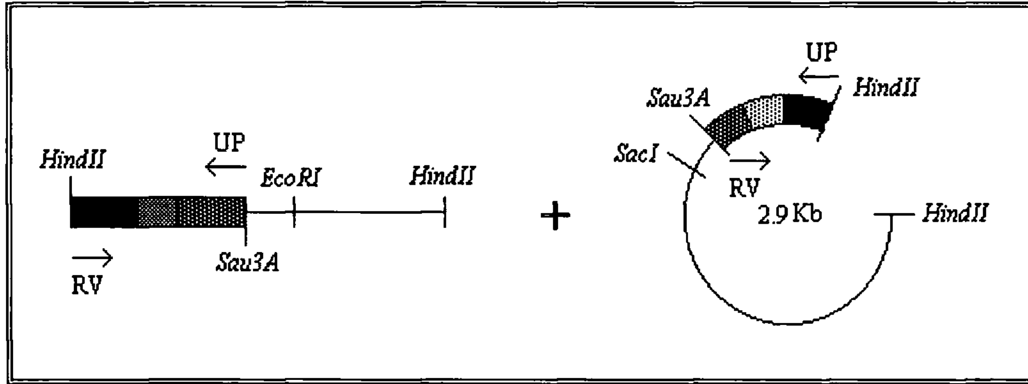


Figure 3.28, Products of *HindII* digestion of NTG+8.

- In order to make sub-clones of NTG+8 for sequencing the following was done:

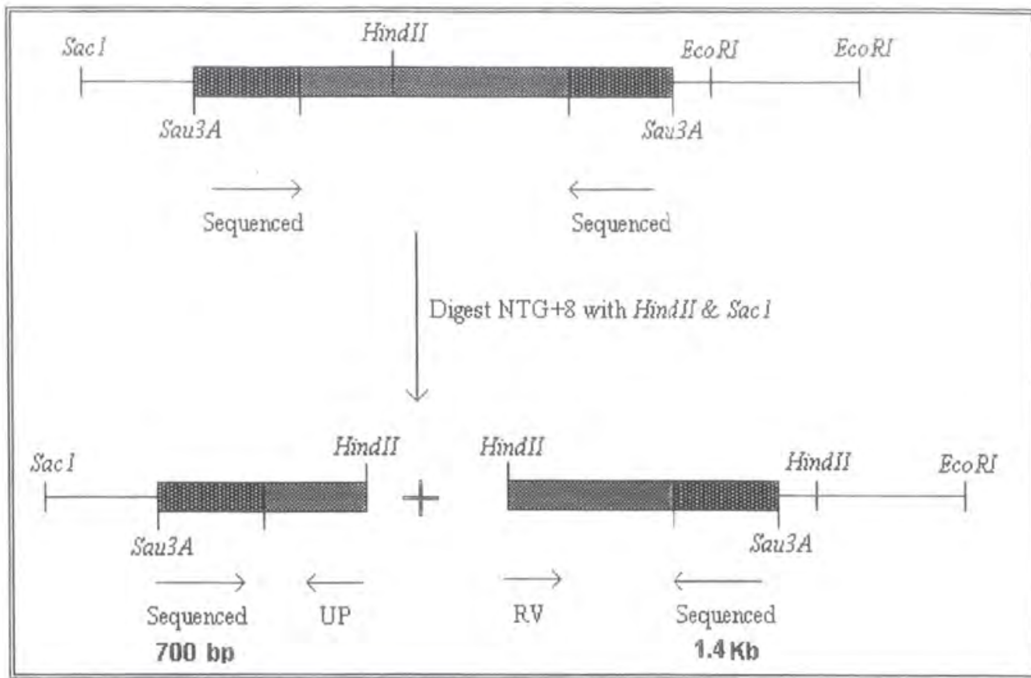


Figure 3.29, Sub-cloning strategy for NTG+8.

⇒ The 1.4Kb fragment and 700bp fragment were isolated from agarose gels using method 2.2.2.11. (see figure 3.30)

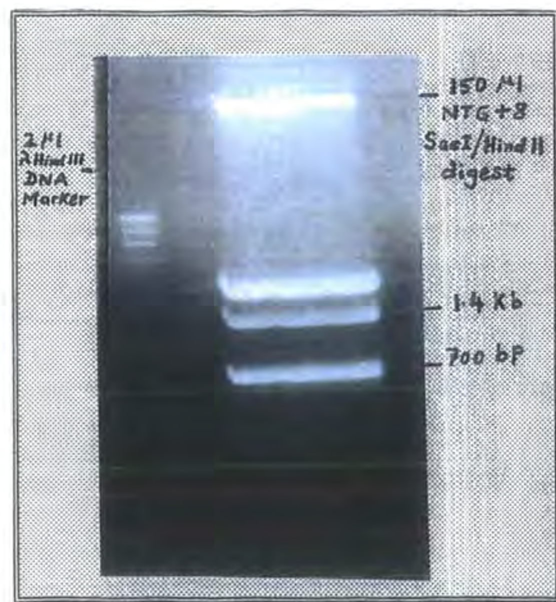


Figure 3.30, 150µl of NTG+8 *SacI/HindII* digest on 2% agarose gel.

⇒ The 700bp fragment was ligated into *SacI/HindII* cut pBluescript:

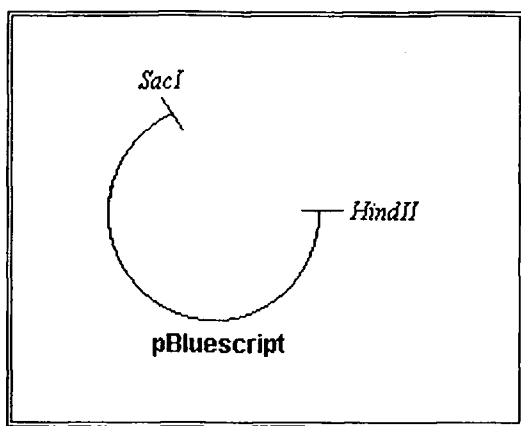


Figure 3.31, *SacI/HindII* cut pBluescript.

⇒ Sequenced using the UP. Homology search using the Daresbury database showed the following (see figure 3.33.I & II for predicted restriction map and sequence):

```
gb_pr:humretblas

LOCUS       HUMRETBLAS 180388 bp    DNA            PRI       23-NOV-1994
DEFINITION  Human retinoblastoma susceptibility gene exons 1-27, complete cds.
ACCESSION   L11910
KEYWORDS    nuclear protein; recessive oncogene; retinoblastoma gene;
            retinoblastoma protein; retinoblastoma susceptibility;
            tumor supressor gene. . . .

SCORES           Init1:   118 Initn:   192 Opt:   174

            86.4% identity in 66 bp overlap
```

Figure 3.32.I,

```
gb_pr:humtcradcv

LOCUS       HUMTCRADCV  97634 bp    DNA            PRI       13-JAN-1995
DEFINITION  Human Tcr-C-delta gene, exons 1-4; Tcr-V-delta gene, exons 1-2;
            T-cell receptor alpha (Tcr-alpha) gene, J1-J61 segments; and
            Tcr-C-alpha gene, exons 1-4.
ACCESSION   M94081
KEYWORDS    T-cell receptor C-alpha; T-cell receptor C-delta; . . .

SCORES           Init1:   110 Initn:   148 Opt:   138
            84.9% identity in 53 bp overlap
```

```
gb_pr:hsp53g

LOCUS       HSP53G       20303 bp    DNA              PRI       23-APR-1992
DEFINITION  Human p53 gene for transformation related protein p53 (also calle
            transformation-associated protein p53, cellular tumor antigen p53
            and non-viral tumour antigen p53).
ACCESSION   X54156
KEYWORDS    anti-oncogene; cell cycle control; growth suppressor; . . .


SCORES           Initl:    111 Initn:    148 Opt:    232
            66.0% identity in 203 bp overlap
```

Figure 3.32.III,

```
gb_pr:hsx11g

LOCUS       HSX11G       2864 bp    DNA              PRI       08-OCT-1994
DEFINITION  H.sapiens x11 gene, promoter region.
ACCESSION   Z32676
KEYWORDS    x11 gene; x11 protein.
SOURCE      human.
  ORGANISM  Homo sapiens . . .


SCORES           Initl:    204 Initn:    332 Opt:    419
            74.7% identity in 221 bp overlap
```

Figure 3.32.IV,

```
gb_pr:humtpa

LOCUS       HUMTPA       36594 bp    DNA              PRI       14-JAN-1995
DEFINITION  Human tissue plasminogen activator (t-PA) gene, complete cds.
ACCESSION   K03021
KEYWORDS    Alu repeat; KpnI repetitive sequence; repeat region;
            tissue plasminogen activator.
SOURCE      Human fetal liver, DNA library of T.Maniatis, multiple clones.


SCORES           Initl:    169 Initn:    323 Opt:    529
            76.9% identity in 242 bp overlap
```

Figure 3.32.V,

```
gb_pr:humafp

LOCUS        HUMAFP        27553 bp    DNA              PRI        26-MAY-1995
DEFINITION   Human alpha-fetoprotein gene, complete cds.
ACCESSION    M16110
KEYWORDS     alpha-fetoprotein.
SOURCE       human.
  ORGANISM   Homo sapiens . . .

SCORES            Init1:    101 Initn:    138 Opt:    184
             55.5% identity in 483 bp overlap
```

Figure 3.32.VI,

```
gb_pr:humtcrb-2

Continuation of HUMTCRB from base 350001 (L36092 Homo sapiens (clones: K41A,
 K35, K26, K56, X21B, G54, H137, H18, H18/G15gap, G15, X1A, A27, A212, A14,

 H7.1, H12.18, H130, A16, C215, G1, C68, C21, X11, X6A, CBG1, CBG1/C29gap, and
 C29) germline T-cel

SCORES            Init1:    292 Initn:    472 Opt:    509
           71.1% identity in 284 bp overlap
```

Figure 3.32:VII,

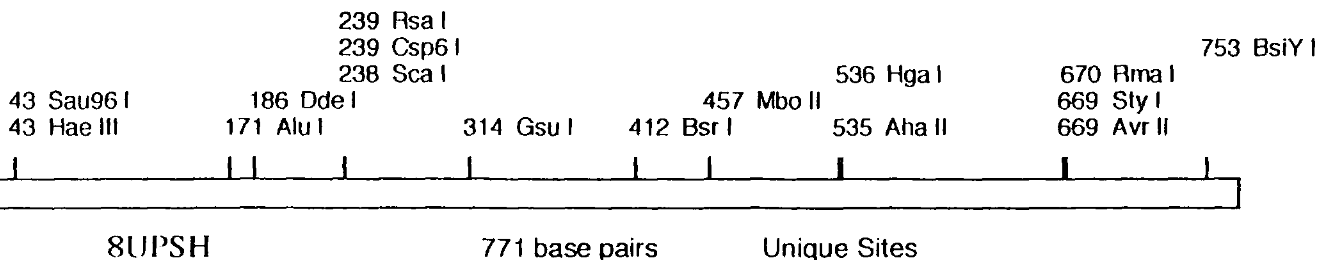


Figure 3.33.I, Restriction map for 700bp *SacI/HindII* sub-clone (NTG+8).

```
          |  10      |  20      |  30      |  40        |  50      |  60      |  70      |  80
  1 CTACAAAAAA AGCCAAACTG ACAGTATTAC TGAATAAGGA  AAGGCCCAAA GAGACAAAAT ACTTTTTATT TTGTAACCTC 80
 81 GGTATGACAC AACTTACCCT AACTATAAAG ACCCTAAATT  ACCAAGATGG GTGCTTATAA TATGGAGAGT TACAAAGTCA 160
161 TTTCACTTTT AGCTTTTTTA TTTCTCTCAG AATAAAAAGT  GTATAAGGAG TTGATAAAGA AGTTGATACT ATAAGTTAGT 240
241 ACTACAATGA CAGCACTTTT CAAGAAAAGA CTTTTTTCTC  TCTTACAAAT ATCATGTTAG CAGTATTTGT TTTCTCCAGA 320
321 AATAATGAGT AAATAAAAAC ATAAGTATGT GGGTAATTAG  TGTAGTTTCT TAAATAAATG AGTTAGGCAA CAGGCTAATA 400
401 ATGTATATTT CACTGGCTTT TCAATGCCAA CAATCATATT  CTTTATAAGG CACAGAGAAG ATTTTTCTAA AGAATAAGTA 480
481 TGTGAACCTG AAAAGTAATC ACCACTTGGT AGTGACAATA  TGGATAGGGT GAAGGGCGTC ACCACGAAGC AATGAAAAGA 560
561 TACATTTGCA GTTAAATTTG AAAACCCATG ATGTTTAATA  CATATAGTAA TAAAGAATAC TTCCTCCTGT CTCAAAATTA 640
641 TTTTAGAATT TAAGATAGAA CTAAAATACC TAGGGATAAT  GATATGACTA CCAAAAATTA AAAATTAAAG GATATTTGGA 720
721 GTATTATAAA TTAAGAATGA AAACTTATTA CCCCATGAAC  AGGGGATAAT T                                771
          |  10      |  20      |  30      |  40        |  50      |  60      |  70      |  80
```

Figure 3.33.II, DNA sequence for 700bp *SacI/HindII* sub-clone (NTG+8).

- The 1.4 Kb fragment was ligated into *HindII* cut into pBluescript:

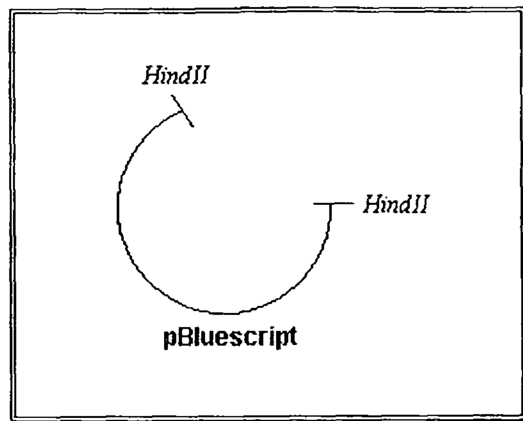


Figure 3.34, *HindII* cut pBluescript.

And sequenced using RV (see figure 3.27).

The same procedure was used for the following NTGs:

<u>NTG+13</u>

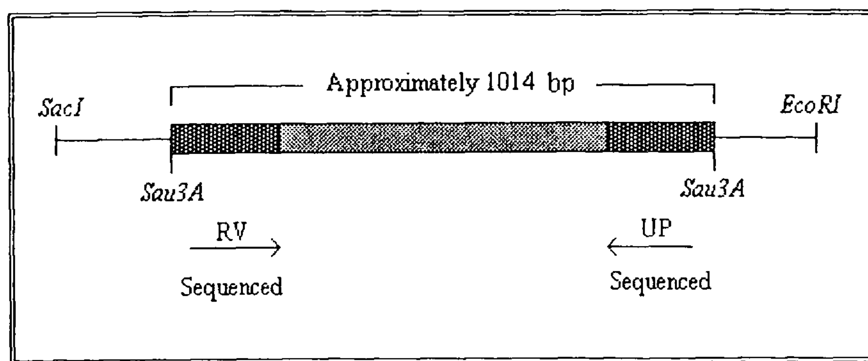


Figure 3.35, NTG+13 sequenced region.

⇒ Homology search shows the following (see figure 3.37.I & II for predicted restriction map and sequence):

```
13r /rev
gb_pr:humrag1

LOCUS       HUMRAG1      6545 bp    mRNA            PRI       08-JAN-1995
DEFINITION  Human recombination activating protein (RAG-1) gene, complete cds.
ACCESSION   M29474
KEYWORDS    recombination activating protein.
SOURCE      Human pre-B cell, line NALM6, cDNA to mRNA, clone H36.
  ORGANISM  Homo sapiens . . .

SCORES          Init1:    64 Initn:    106 Opt:     75
            55.0% identity in 249 bp overlap
```

Figure 3.36.I,

```
gb_pr:humds

LOCUS       HUMDS        22573 bp    DNA            PRI       22-MAR-1995
DEFINITION  Human gene for dihydrolipoamide succinyltransferase, complete cds
            (exon 1-15).
ACCESSION   D26535
KEYWORDS    dihydrolipoamide succinyltransferase.
SOURCE      Homo sapiens (library: lambda EMBL3) peripheral blood cells DNA,

SCORES          Init1:   373 Initn:   514 Opt:   514
            68.5% identity in 352 bp overlap
```

Figure 3.36.II,

```
gb_pr:hscsf1po

LOCUS       HSCSF1PO     35100 bp    DNA            PRI       24-APR-1995
DEFINITION  Human c-fms proto-oncogene for CSF-1 receptor.
ACCESSION   X14720
KEYWORDS    c-fms proto-oncogene; CSF-1 receptor; glycoprotein;
            growth factor receptor; proto-oncogene;
            transmembrane glycoprotein; tyrosine kinase. . . .

SCORES          Initl:    54 Initn:    125 Opt:     59
            85.0% identity in 20 bp overlap
```

Figure 3.36.III,

```
LOCUS       HUMCD4      13133 bp    DNA          PRI       01-NOV-1994
DEFINITION  Human recognition/surface antigen (CD4) gene, 5' end.
ACCESSION   M86525
KEYWORDS    glycoprotein; immunoglobulin super gene family;
            recognition antigen; surface antigen.
SOURCE      Homo sapiens DNA. . . .

SCORES          Initl:   340 Initn:   491 Opt:   532
            75.3% identity in 247 bp overlap
```

Figure 3.36.IV,

```
gb_pr:humgpp3a21

LOCUS       HUMGPP3A21   6829 bp    DNA          PRI        26-MAY-1995
DEFINITION  Human platelet glycoprotein IIIa, exon 15.
ACCESSION   M32686 J05427
KEYWORDS    Alu repeat; integrin; platelet fibrinogen receptor;
            platelet glycoprotein.
SEGMENT     21 of 21 . . .

SCORES          Initl:   310 Initn:   494 Opt:   485
            71.8% identity in 255 bp overlap
```

Figure 3.36.V,

```
gb_pr:hum24dc99z

LOCUS       HUM24DC99Z   3678 bp    DNA           PRI       06-JUL-1995
DEFINITION  Homo sapiens (subclone 9_a8 from P1 H24) DNA sequence.
ACCESSION   L43410
KEYWORDS    Interleukin growth hormone cluster on chromosome 5 (5q31).
SOURCE      Homo sapiens (tissue library: Subclones in pJT2 from P1 clone H24
            DNA. . . .

SCORES          Initl:   338 Initn:   498 Opt:   570
            66.9% identity in 402 bp overlap
```

Figure 3.36.VI,

```
gb_pr:humbfxiii

LOCUS       HUMBFXIII    33290 bp    DNA               PRI       31-OCT-1994
DEFINITION  Human factor XIII b subunit gene, complete cds.
ACCESSION   M64554 J05294
KEYWORDS    blood coagulation factor; factor XIII; factor XIIIb; zymogen.
SOURCE      Human DNA.
  ORGANISM  Homo sapiens . . .

SCORES           Init1:    76 Initn:   114 Opt:    86
            55.9% identity in 145 bp overlap
```

Figure 3.36.VII,

```
gb_pr:humretblas

LOCUS       HUMRETBLAS 180388 bp    DNA               PRI       23-NOV-1994
DEFINITION  Human retinoblastoma susceptibility gene exons 1-27, complete cds
ACCESSION   L11910
KEYWORDS    nuclear protein; recessive oncogene; retinoblastoma gene;
            retinoblastoma protein; retinoblastoma susceptibility;
            tumor supressor gene. . . .

SCORES           Init1:    76 Initn:   161 Opt:   117
            52.6% identity in 441 bp overlap
```

Figure 3.36.VIII,

```
gb_pr:hstcrbv

LOCUS       HSTCRBV     77743 bp    DNA               PRI       20-APR-1994
DEFINITION  Human V beta T-cell receptor (TCRBV) gene locus.
ACCESSION   U03115
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens . . .

SCORES           Init1:    71 Initn:   206 Opt:   144
            57.3% identity in 239 bp overlap
```

Figure 3.36.IX,

```
LOCUS       HUMMHCC6A     3453 bp    mRNA              PRI       07-JAN-1995
DEFINITION  Human MHC class III complement component C6 mRNA, complete cds.
ACCESSION   J05024
KEYWORDS    class III gene; complement component C2; complement component C6;
            complement system protein; major histocompatibility complex.
SOURCE      Human liver, cDNA to mRNA, clones lambda-C6/[5@-18:111,21:211]. .

SCORES          Init1:    62 Initn:    138 Opt:    61
            63.9% identity in 61 bp overlap
```

Figure 3.36.X,

```
gb_pr:humafp

LOCUS       HUMAFP       27553 bp    DNA              PRI        26-MAY-1995
DEFINITION  Human alpha-fetoprotein gene, complete cds.
ACCESSION   M16110
KEYWORDS    alpha-fetoprotein.
SOURCE      human.
  ORGANISM  Homo sapiens . . .

SCORES          Init1:   101 Initn:   138 Opt:   184
            55.5% identity in 483 bp overlap
```

Figure 3.36.XI,

```
gb_pr:hsat3

LOCUS       HSAT3        14206 bp    DNA              PRI        03-NOV-1994
DEFINITION  H.sapiens gene for antithrombin III.
ACCESSION   X68793 S52236 S52240
KEYWORDS    antithrombin; antithrombin III gene; AT3 gene; plasma protein
            serine proteinase inhibitor; serpin.
SOURCE      human. . . .

SCORES          Init1:    53 Initn:   125 Opt:    68
            55.8% identity in 163 bp overlap
```

Figure 3.36.XII,

```
gb_pr:hscuznso

LOCUS        HSCUZNSO     4093 bp    DNA            PRI        12-AUG-1994
DEFINITION   H.sapiens gene for Cu/Zn-superoxide cismutase.
ACCESSION    Z29336
KEYWORDS     Cu/Zn-superoxide dismutase.
SOURCE       human.
  ORGANISM   homo sapiens . . .

SOUPES           Initl:   352 Initn:   545 Opt:   515
             75.9% identity in 232 bp overlap
```

Figure 3.36.XIII,

```
gb_pr:humfmr1s

LOCUS        HUMFMR1S   152351 bp    DNA          PRI        18-JAN-1995
DEFINITION   Homo sapiens fragile X mental retardation protein (FMR-1) gene (o
             alternative splices), complete cds.
ACCESSION    L29074 L38501
KEYWORDS     fragile X mental retardation syndrome protein;
             fragile X syndrome; repeat region. . . .

SOURES           Initl:   69 Initn:   106 Opt:   71
             55.3% identity in 114 bp overlap
```

Figure 3.36.XIV,

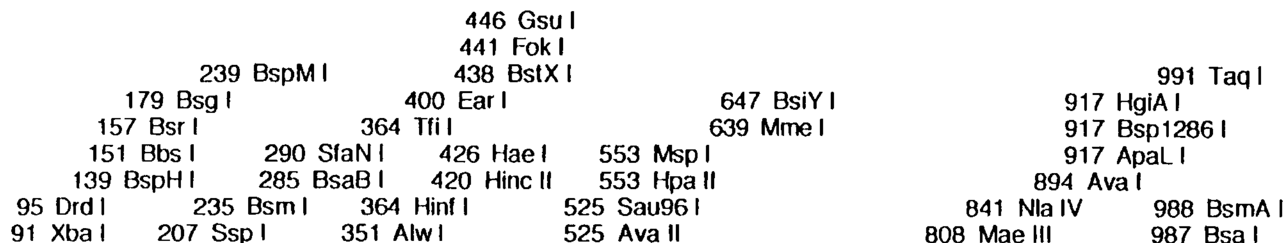


Figure 3.37.I, Restriction map of NTG+13.

```
        |    10     |    20     |    30     |    40       |    50     |    60     |    70     |    80
   1 GATCATGCAA TAGTAAGAAG TTCTAAACTA TACATATTCC  CACGCACCAA AATGTTTCCT TTTTCATTTG TCAGGACTTG 80
  81 CCCAAATCTA TCTAGACAGA GTTGTCATCA CAGAGTATTG  AATAGTGTAT AATTTGCTTC ATGAACATCT GTCTTCCCAG 160
 161 TATATCGTTC TTTATTGGCT GCACAATCAC AATTGGGCTG  GTGGATAATA TTTACTGACA TACTTATTTT GTAAGAATGC 240
 241 AGGTTGTTTC CACTTTGTAC AACGGAATTA TCAGATAACA  CCACGATCAG CATCTTAGTA TGCAGACTTC TGGCTTACGA 320
 321 TCAATCAATT TACTTGGTAT AAAATCCCAA GGATCAGGGA  CTTGAATCTG ACATTGTTCA TAATGCTTGC CACCTATGCG 400
 401 AAGAGATTAC CACCAGCTAG TTGACAGGCC AAACATTCCA  CATCCCTGGA GATGTATTTT GCCCAAACTT ACTAAGGCAT 480
 481 AAGACAGGCT TACTGAAAAT TTTACTTCCA TTTA*****G  ATTTGGACCA GGATAATATG CTGCATTGAA CTCCGGCGAA 560
 561 GAATACCATT TGTGAGGCGA CAACAGGTAG AGAAGGGAAG  CAATTGTAGA AGAAGGTTGT CAAGAAGGCC GAACAAGCTC 640
 641 CAACTGCCTT GGAGGGGGAA GTAAATCCCC AAAGATAAAC  TTTGGTGAAG GTGAAATAAT TTAGCAGATT TCAAGTAGCT 720
 721 GTTGATTTTC CCTGAAAGTG AAGTTTAACT CAGTTTTTCG  TTCTCTCTGG TATATTATCA TAGTATTATT ATTTATAGTG 800
 801 TCGTTCTGTC ACCCAGCCTG AAGTACAGTG GTACAAAGTT  GGCTCCCTGC AACTTCTTCC TCTGGGGTTC AAGTCATTCT 880
 881 TGTGCCTCAG CCTCCCGAGT AGCTGGGATT ACAGGCGTGC  ACCACCACGC CTGACTTAAT TTTGTATTCT TCAGTACAGA 960
 961 TGGGGTATCA CCATGTTGTC CAGGTTGGTC TCGAACTGCT  GCCCTCAACT GATC                             1014
        |    10     |    20     |    30     |    40       |    50     |    60     |    70     |    80
```

Figure 3.37.II, DNA sequence of NTG+13.

NTG+33

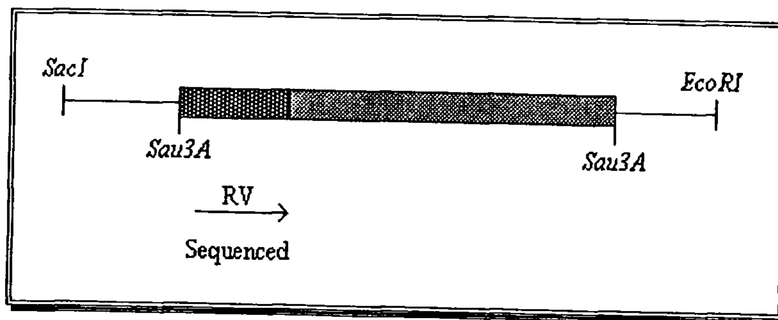


Figure 3.38, NTG+33 sequenced regions.

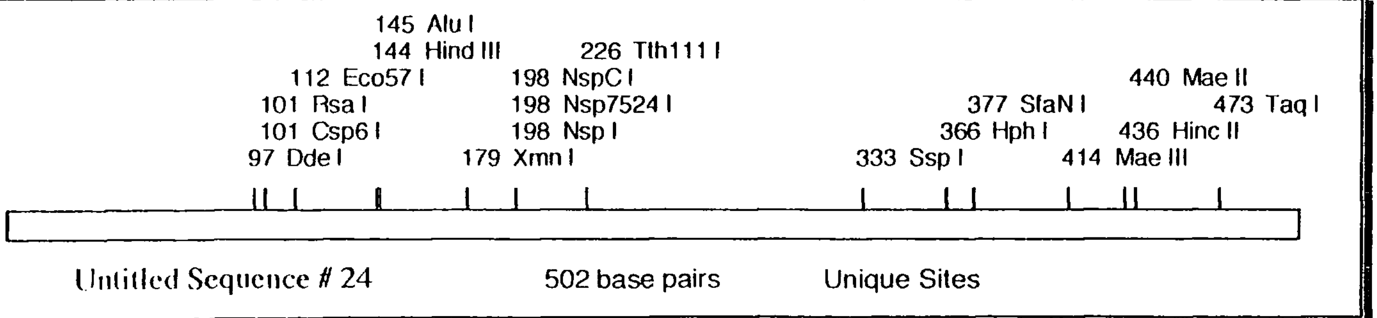


Figure 3.39.I, Restriction map of NTG+33.

```
          |  10      |  20      |  30      |  40        |  50      |  60      |  70      |  80
  1 CGGTAAGATA TTTTTAAAAA ATTCTTACCC TATGTTTAAT  GTTTTCAAAA TTCAATAGAA AAAAATATAA GGTCCTTATA 80
 81 CAGTGGCAAA TACATTCTCA GTACCACTAC ACTGAAGTAG  AATTCTGAAT TCCTATTTTT TCAAAGCTTT TTTTGTTTAA 160
161 ATTAGTCAAT ATGTAATAGA AGTATTTCAA GGTTATAACA  TGCAATGCTA GTTTATATAT TGATTGACAT TGTCTAATAA 240
241 GATAAAACTT TTGTTTCTTA AAGACTAGAC TGTTTAATCT  TATTTTTCTA CTGGATAAAT TATTTCAGAA TTTCTAATTA 320
321 TAACATGGTA ATAATATTAT TAAAATCCCA CCTTTCAAAA  AGTGAGGTGA TAATTCGCAT CTTTTGTCTT TCAATTGTGC 400
401 TGGGTCCCAG TTTGTGACTG CCTAAAAGAA AACCCGTTGA  CGTTTAGAAT AACTTTTTAA AAACTCCCTC TCTCGAAGTT 480
481 CAAATTATTT CCCTTGTGTT TA                                                                502
          |  10      |  20      |  30      |  40        |  50      |  60      |  70      |  80
```

Figure 3.39.II, DNA sequence of NTG+33.

3-2- Protein analysis

- Standard curve results:

Protein conc for std. (μg)	OD_{595}	OD_{595} duplicate	Mean
0	0	0	0
5	0.153	0.159	0.156
10	0.307	0.295	0.301
17.5	0.448	0.480	0.464
25	0.602	0.594	0.598
35	0.751	0.757	0.574
45	0.940	0.934	0.937
55	1.17	1.20	1.185

Table 3.3.

- From each cell line, [SK-N-SH] and [SK-N-Be2c], 10μl of nuclear extract was added to 2.5ml of diluted dye and read OD_{595} (see chapter 2, protein analysis).

- The value of protein (μg) was read from the standard curve and calculated the value of μ g protein / ml.

Cell line	OD_{595}	Protein µg / µl
SK-N-SH	0.19	0.9
SK-N-Be2c	0.40	1.6

Table 3.4.

- Protein extracts from SK-N-SH and SK-N-Be2c were separated by SDS-PAGE, to compare their composition. As shown in figure 3.40, few significant differences were noted.

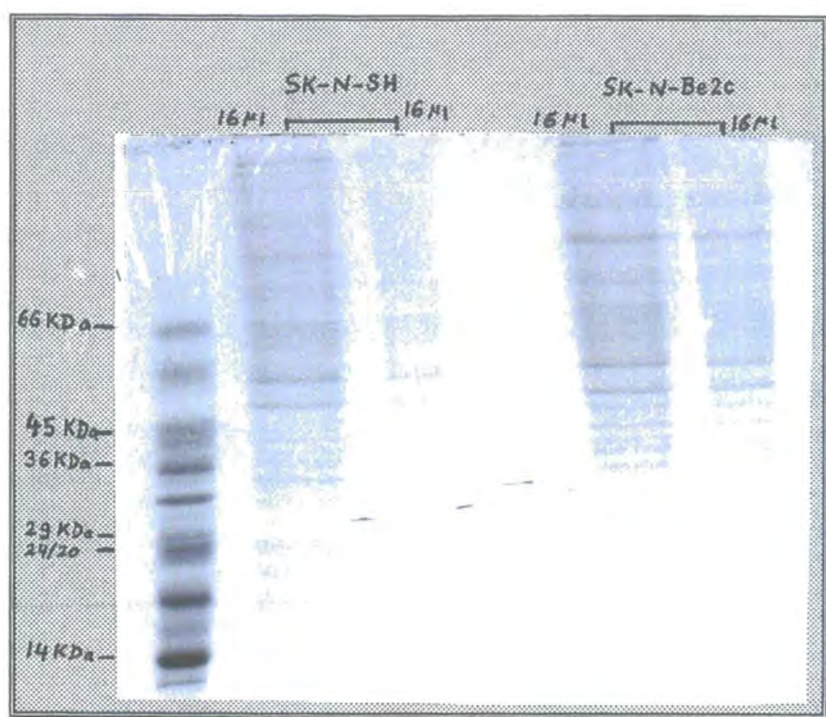


Figure 3.40, SDS PAGE of protein extracts from SK-N-SH and SK-N-Be2c.

3.3. Colony Hybridization

To ascertain whether any of the clones isolated had previously been identified as NTG clones from SK-N-Be2c, the clone NTG-21 from that cell line was used as a probe against colony blots of the NTG library from SK-N-SH cells. As shown in figure 3.41 it would appear that no significant hybridization to NTG-21 was shown by any of the SK-N-SH NTG clones.

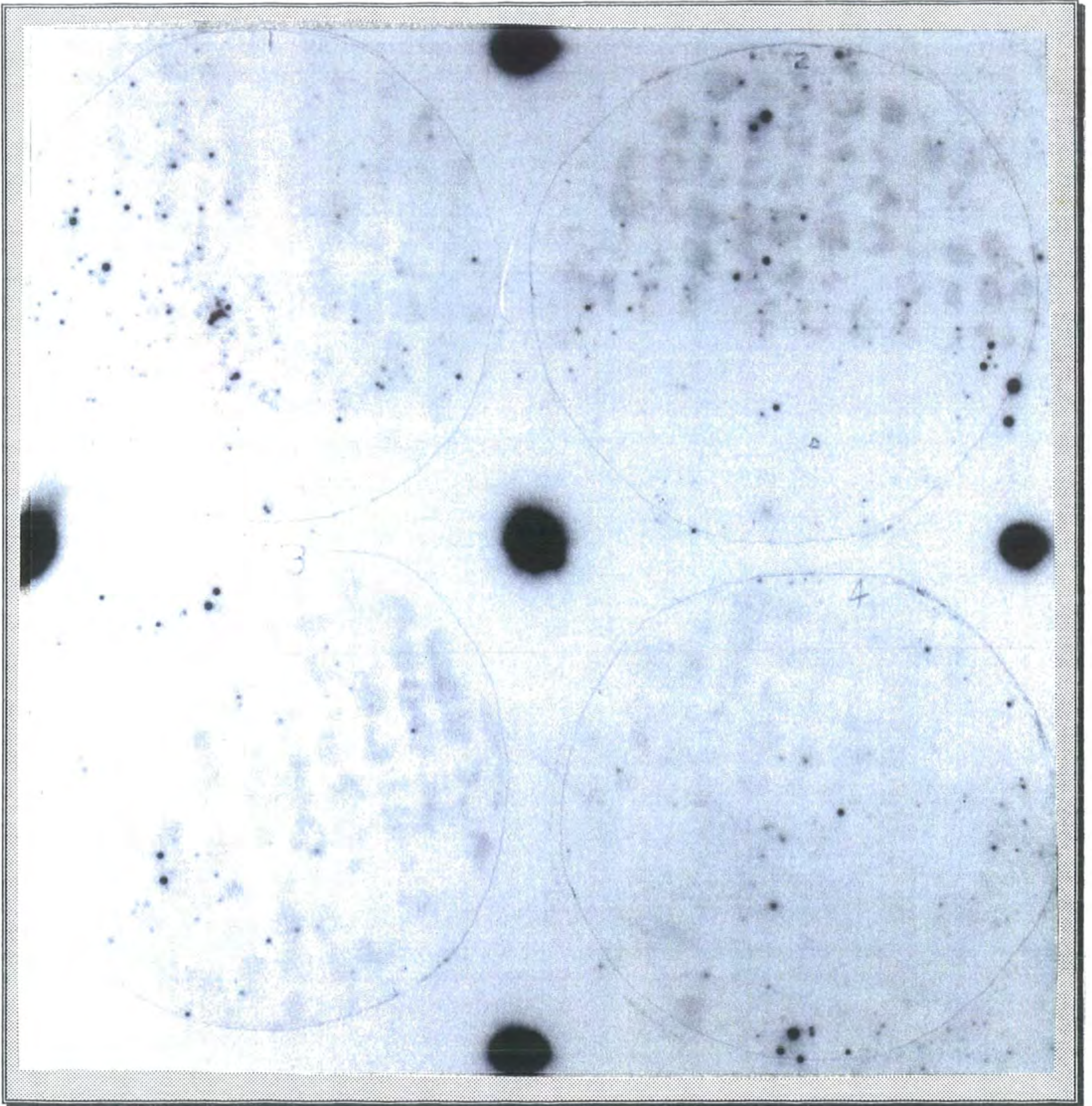


Figure 3.41, Colony blots probed with NTG-21 (see method 2.2.2.10 & 2.2.2.11).

4. Discussion

The project was carried out in order to identify potential target genes for N-*myc* using the cell line SK-N-SH.
The method used to generate genomic library has led to approximately 400 clones being isolated, some sequenced, and sent to be compared with known DNA sequenced using "Daresbury" data bases. This procedure can potentially identify genes present in the NTG library.
Out of the 400 clones, 114 NTGs had their insert size calculated, four partially sequenced, one had a low affinity N-Myc binding site "CAT GTG", plus C-Myc binding sites CAC GTG, and TCT CTTA (see table 3.2. page 44). However the final resolution of which clones are true in *vivo* targets for regulation by N-myc will await the cloning these putative regulatory sequences into reporter plasmids and monitoring activity in N-myc expressing neuroblastoma cells.

Colony hybridization was done for the 400 clones to check if any had the NTG-21 insert which was isolated previously in the same lab from the cell line SK-N-Be2c. The result was negative.
The study of the outcome of the homology search was based on high similarity over potential extended length, compared with the actual length of the sequenced part, and if the homologue is of interest by looking at the known neuroblastoma literature (if the gene plays a role in Neuroblastoma cells)

From the obtained sequence, the comparison data and considering the role of every gene; a substantial amount of information was gathered, some, in a way or another, play a role in neuroblastoma cells and a lot of homology had no role at all, and most were intron genes. Some of the genes picked out by sequence comparison include:

- Human pre-B cell, cDNA to mRNA. Human recombination activating protein (RAG-1) gene [66.0% identity in 203bp overlap to NTG+8].

- B-cell, also called B lymphocytes give rise to plasma cells, which produce antibodies. This defence system by B-cells is called antibody-mediated immunity or humoral immunity because these antibodies present in the blood stream.
During development, antibody genes are re-arranged and N-*myc* expression is switched off, and some studies have revealed that down-regulation of N-*myc* expression in the later stages of B-cell development is mediated primarily at the level of transcriptional initiation and that dominant, trans-acting factors present in more mature B-lineage cell lines act to down-regulate the transcription of N-*myc*.

- Human DNA sequence from cosmid HW2, Huntington's disease region [62.5% identity in 120 overlap to NTG+5].

- Huntington's disease formerly Huntington's chorea, is a hereditary disease of the central nervous system caused by a mutated gene on chromosome 4. (Neuroblastoma tumour arise from neural crest cells which form part of the sympathetic nervous system).

- Human retinoblastoma susceptibility gene [52.6% identity in 441 overlap to NTG+8].

- Retinoblastoma is a hereditary malignant tumour of the retina (delicate, multilayered, light-sensitive membrane lining the inner eyeball and connected by the optic nerve to the brain), transmitted as a dominant trait and occurring chiefly among infants.
Subsequent studies revealed amplification of N-*myc* in retinoblastoma and indirect evidence has indicated that the product of the retinoblastoma gene (pRb) may be necessary for the beta type transforming growth factor (TGF-beta) suppression of N-*myc* expression and not for TGF-beta1 inhibition of branching morphogenesis. Therefore suppression of N-*myc* is not necessary for inhibition of branching morphogenesis by TGF-beta1, (TGF-betas are potent inhibitors of epithelial cell proliferation). [Hannon G, Demntrick G, Beach D; Isolation of the Rb-related p130 through its interaction with CDK2 and cyclins: Genes Dev. 1993, 7 (12A), 2378-2391] & [Friend S. H, Bernards R , et al. ; A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma: Nature 1986, 323 (6089), 643-646.].

- Human vascular cell adhesion molecule-1 (VCAM1) gene [61.5% identity in 117 overlap to NTG+8] , [70.5% identity in 661 overlap to NTG+5].

- Vascular cell adhesion molecule-1 was first identified as a protein that appears on the surface of endothelial cells after exposure to inflammatory cytokines.
Through interaction with it's integrin counter receptor VLA-4, VCAM-1 mediates cell-cell interactions important for immune function. The recruitment of immune cells to the endothelium is a complex process involving the interaction of ligands or counter receptors on the surface of endothelial cells with receptors on immune cells, and in addition to endothelial cells, VCAM-1 is also expressed on lymphoid dendritic cells, and stromal fibroblasts in bone marrow, and on some tissue macrophages.
Endothelial cell surface ligands may be involved in metastasis of tumour cells, as it has been proposed that the interaction of tumour cells with activated endotheliun is a mechanism that facilitates the migration of circulating tumour cells into tissues. Also small cell lung cancers (where N-*myc* is amplified) express LFA-1 and MAC-1 integrin which binds to ICAM-1 (a member of the endothelial cell ligands). [Iademarco M. F, McQuillan J. J, et al.; Characterization of the promoter for vasculear cell adhesion molecule-1 (VCAM-1): J. Biol. Chem. 1992, 267 (23), 16323-16329]

- Human p53 gene - associated protein p53, cellular tumour antigen p53 [66.0% identity in 203 overlap to NTG+8].

- It was thought in several reports on p53 that this genotype is found to have virtually exclusive wild-type status in primary neuroblastoma, and that p53 plays no role in the development of neuroblastoma, but recently it was reported that the vast majority of undifferentiated neuroblastoma's exhibit abnormal cytoplasmic sequestration of wild-type p53. This inability of p53 to translocate to the nucleus presumably prevents the protein from functioning as a suppressor.

The loss of p53 function in neuroblastoma seems to play a major role in the tumorigenesis of undifferentiated neuroblastoma (neuroblastoma might abrogate the transactivating function of p53 inhibiting it's access to the nucleus, rather than by gene mutation). [Zauberman A, Barak Y, et al.; Sequence-specific DNA binding by p53: identification of target sites and lack of binding to p53-MDM2 complexes: EMBO 1993, vol. 12 no. 7 pp. 2799-2808].

- H . sapiens genes TAP1, TAP2, DOB, DOB gene, LMP2 gene, LMP7 gene, Major histocompatibility complex class II [79.3% identity in 29 overlap to NTG+8].

- Some studies are being carried out to determine whether a antitumour immune response would be generated by introducing Xenogeneic class II MHC genes into tumour cells, thus some approaches have been made to use retroviral-mediated transduction of class to MHC genes into human tumour cells for an effective alternative to current cancer treatment. [Glynne R, Kerr L. A, et al.; The major histocompatibility complex-encoded proteasome component LMP7, alternative first exons and post-translational processing: Euro. J. Immunol 1993, 23 (4), 860-866]

- H . sapiens genes TAP1, TAP2, DOB, DOB gene, LMP2 gene, LMP7 gene.

- The LMP7 gene maps to human MHC class II region within 15Kb of LMP2 and between TAP1 and TAP2.

TAP1 and TAP2 have homology to the ATP binding cassette (ABC) superfamily of transporters and have been shown to play a role in the processing and presentation of endogenous antigens through MHC class1 (neuroblastoma has a decreased expression of the MHC class 1 antigen which correlates with high level of expression of the N-*myc* gene, thus expression of MHC class 1 antigens may result in neuroblastoma cells being more resistant to T-cell mediated immune surveillance). [Beck S, Kelly A, et al.; DNA Sequence analysis of 66kb of the human MHC Class II region encoding a cluster of genes for antigen processing: J. Mol. Biol. (1992) 228, 433-441].

- H . sapiens mRNA for Rb2/p130 protein [55.7% identity in 219 overlap to NTG+5].

- Rbr-2 = retinoblastoma susceptibility gene.

- Rb2/p130 is a member of the Rb family (Retinoblastoma).

pRb and Rb2/p130 share considerable homology in their pocket domains, they share two additional regions of significant identity near their amino termini, they not only share primary sequence homology but also functional similarity, each of these proteins combine specifically to the viral oncoproteins, SV40 antigen and adenvirous EIA.

The structural and functional similarities between Rb and Rb2/p130 gives two possibilities for the role of Rb2/p130 either it may act as a inhibitor of cellular proliferation or as a tumour suppressor gene (Rb is the retinoblastoma susceptibility tumour suppressor gene, pRb is the product of this gene which has been shown to play a key role in the regulation of cell division. [Mayol X, Grana X, et al.; Cloning of a new member of the retinoblastoma gene family (pRb2) which binds to the E1A transforming domain: Oncogene (1993) 8, 2561-2566].

Other genes picked out by sequence comparison include:

- Human gene for dihydrolipoamide succiny 1 transferase.

- Human thymidine Kinase gene.

- Human recognition surface antigen (CD4) gene glycoprotein, immunoglobulin super gene family. [Edwards M. C, Gibbs R. A; Human dinorphism resulting from loss of an Alu: Genomics 1992, 14 (3), 590-597]

- H . Sapiens 3'flanking region for estrogene receptor gene. [Keaveney M, Parker M. G, et al.; Identification of a functional role of the 3' region of the human oestrogene receptor gene: J. Mol. Endocrinol 1993, 10 (2) 143-152]

- Human c-fms proto-oncogene for csf-1 receptor.

- Human alpha-fetoprotein gene. [Stocker N. G, Cheah K. S, et al.; A highly polymorphic region 3' to the human type II collagen gene: Nucleic Acids Res 1985, 13 (13) 4613-4622]

- Human mRNA for second protein of Inter-alpha-trypsin inhibitor complex.

- Human ATP synthase beta subunit (ATPSB) gene.

- Human acetylated low density lipoprotein (ACLDL).

- Homo Sapiens p-selection CD62 (GRMP) gene, 5' end and promoter region. [Pan J, McEver R. P; Characterization of the promoter for the human P-selectin gene: J. Biol. Chem. (1993) 268 (30), 22600-22608].

- Human antigen CD36 (Alu repeat).

- Human breakpoint cluster region (BCR) gene. [Gu Y, Alder H, et al.; Sequence analysis of the breakpoint cluster region in the ALL-1 gene involved in acute leukemia: Cancer Res. (1994) 54, 2327-2330].

- Human fragile X mental retardation gene. [Nicola N. A, Metcalf D; Subunit promiscuity among hemopoietic growth factor receptors: Cell (1991) 67 (1) 1-4].

- Human insulin receptor gene. [Elbein S. C; Molecular and clinical characterization of an insertional polymorphism of the insulin-receptor gene: Diabetes (1989) 38 (6), 737-743].

- Human T-cell receptor beta-chain. [Slightom J. L, Siemieniak D. R, et al.; Nucletide sequence analysis of 77.7Kb of the human V beta T-cell receptor gene locus, direct primer-walking using cosmid template DNAs: Genomics (1994) 20, 149-168].

- Human vitamin D binding protein gene.

- Human interferon-beta-2 gene.

- Human interleukin-6 receptor (ILb) gene. [Bowcock A. M, Kidd J. R, et al.; The human 'interferon-beta 2/hepatocyte stimulating factor/ interleukin-6' gene, DNA polymorphism studies and localization to chromosome 7p21: Genomics (1988) 3 (1) 8-16].

However the most intriguing and exciting of these potential N-myc targets was found by my lab. Colleague Dr. D. Pearson in clone NTG-21 using her cell line SK-N-Be2c (same procedure was used).
NTG-21 contains a sequence conserved with one found in intron one of the human DOPA decarboxylase gene. The degree of sequence identity suggests that they are not one and the same, but represent similar genes.m DOPA decarboxylase is involved in the diosynthesis of dopamine, epinephrine, norepinephrine and serotonine neuronal tissue, it shows high activity in both neuroblastoma and small cell lung carcinomas, another tissue with N-myc amplification. As N-myc and catecholamine levels both decline after retinoic acid induction, it was speculated that NTG-21 is part of a gene closely related to DOPA decarboxylase possible encoding an isoenzyme, further analysis of NTG-21 is bing carried out by Dr. Pearson to isolate the gene, obtain it's true identity, and find out whither or not it is under the control of N-myc.
Other research work is being carried out at Newcastle University to try to identify genes that are coamplified with N-myc, and determine how they affect prognosis by looking for genes which map close to N-myc such as DEAD box (DDXI) which may have some role in prognosis and response, ornithine decarboxylase (ODC), ribonucleotide reductase (RDR).

4.1. Conclusion:

Although no final conclusion was obtained with respect to the effect of N-*myc* upon other genes in neuroblastoma cells, the result of my findings would play a conducive part towards future work in that field, as more time is needed and more work needs to be carried out on the rest of the 400 clones, for example completing the sequencing, computer analysis to identify sequence homologies and binding sites. Once any gene of interest has been identified a Northern blot will be done to determine whether or not it is under control of N-*myc*, using three different cells as the levels of gene expression may depend on the concentration of N-Myc relative to other b-HLH-Z proteins:

⇒ SK-N-Be2c Neuroblastoma cell with amplified N-*myc*.
⇒ SK-N-SH Neuroblastoma cell with a single copy of N-*myc*.
⇒ Normal cell.

Experiments would be performed under normal culture conditions and after cells have been induced to differentiate with retinoic acid, as it has been shown to be one of the most potent chemical inducers of differentiation in human neuroblastoma. Until then the question whether N-*myc* has a control upon other genes remains unproven.

5. References

5.1. Journals:

1. Suzuki T, Yokota J, Mugishima H, et al.; Frequent loss of heterozygosity on chromosome 14q in neuroblastoma.
2. Sriratsan E.S., Ying K.L., Seeger; Deletion of chromosome 11 and of 14q sequences in neuroblastoma: Genes chrom cancer 1993, 7, 32-37.
3. Glass D.J., Yancopoulos G.D.; The neurotrophins and their receptors: Trends cell Biol 1993, 3, 262-268.
4. Levi-Montalcini R; The nerve growth factor, thirty-five years later: Science 1987, 237,1154-1164.
5. Gibbs R.B., Pfaff D.W.; Insitu hybridization detection of trk A mRNA in brain, distribution, colocalization with p75NGFR and upregulation by nerve growth factor: J Comp neurol 1994, 341 (3), 324-339.
6. Hempstead B.L., et al.; High-affinity NGF binding requires coexpression of the trk proto-oncogene and the low-affinity NGF receptor: Nature 1991, 350, 678-683.
7. Nakagawara A, et al; Inverse relationship between trk expression and N-myc amplification in human neuroblastomas: Cancer Res. 1992, 52, 1364-1368.
8. Alitala K, Schwab M; Oncogene amplification in tumour cells: Adv cancer Res. 1986, 47, 235-281.
9. Bishop J.M.; The molecular genetics of cancer: Science 1987, 235, 305-311.
10. Brodeur G.M., et al.; Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage: Science 1984 224, 1121-1124.
11. Kohl N.E., et al.; Transposition and amplification of oncogene related sequences in human neuroblastomas: Cell 1983, 35, 359-367.
12. Schwab M, et al.; Chromosome localization in normal human cells and neuroblastoma of gene related to C-myc: Nature 1984, 308, 288-291.
13. Cowell J.K.; Gene amplification in mammalian cells: Annu. Rev. Genet 1982, 16,21-59.
14. Squire J. A, Thorner P. S, et al.; Co-amplification of *MYCN* and a DEAD box gene (*DDX1*) in primary neuroblastoma: Oncogene 1995, 10, 1417-1422.
15. Corvi R, Savelyeva L, et al.; Cytogenetic evolution of *MYCN* and *MDM2* amplification in the neuroblastoma LS tumour and its cell line: European J. of cancer 1995, (31A) 4, 520-523.
16. Jones N; Transcriptional regulation by dimerization, two sides to an incestuous relationship: Cell 1990, 61, 9-11.
17. Kato G, et al.; An amino terminal C-Myc domain required for neoplastic transformation activates transcription: Cell Biol. 1990, 10, 5941-5920.

18. Blackwood E.M., et al; Max: a helix-loop- helix zipper protein that forms a sequence specific DNA biding comlex with myc: Science 1991, 251, 1211-1217.

19. Prendergast G.C., et al; Association of Myc, the murine homolog of Max which C-Myc stimulate methylation sensitive DNA binding and Ras cotransformation: Cell 1991, 65, 395-407.

20. Blackwell T.K., et al; Sequence-specific DNA-binding of C-Myc protein: Science 1990, 250, 1149-1151.

21. Blackwood E.M., et al; Myc and Max associate in vivo: Genes 1992, 6, 71-81.

22. Anderson K, et al; DNA-bound fos proteins activate transcription in yeast: Cell, 1988, 52, 179-184.
23. Kato G.J., et al.; An amino-terminal C-myc domain required for neoplastic transformation activates transcription: Mol. Cell, 1990, 13 id, 10, 5914-5920.
24. Blackwood E.M., Eisenman R.N.; A helix-loop-helix zipper protein that forms a sequence-spacific DNA- binding complex with myc: Science, 1991, 251, 1211-1217.
25. Kretznev L, Blackwood E.M.; The Myc and Max proteins possess distinct transcriptional activities: Nature, 1992, 359, 426-429.
26. Freytag S.O., Dang C.V., Lee W.M.; Definition of the activities and properties of C-myc required to inhibit cell differentiation: Cell growth differ, 1990, 1, 339.
27. Evan G.I, Amati B, et al.; The c-Myc protein induces cell cycle progression and apoptosis through dimerization with Max: EMBO J. 1993, 12, 5083-5087.
28. Zauberman A, Oren M, et al.; Sequence-specific DNA binding by p53, identification of target sites and lack of binding to p53-MDM2 compleexes: EMBO J. 1993, 7, 2799-2808.

5.2. Other bibliographic sources

29. Schwab M, et al.; Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell line and a neuroblastoma tumour: Nature, 1983, 305, 245-305.
30. Harold E, et al; Human N-myc gene contributes to neoplastic transformation of mammalian cell in culture: Nature, 1985, 316, 160-162.
31. Blackwell T.K., et al.; Sequence-Specific DNA binding by the c-Myc protein: Science, 1990, 1149-1151.

32. Fernandez C.B., et al; The ornithine decarboxylase gene is a transcriptional target of c-Myc: Proc. Natl. Sci., 1993, 7804-7808.
33. Zauberman A, et al; Sequence-specific DNA binding by p53: identification of target sites and lack of binding to p53-MDM2 complexes: EMBO J, 1993, 12, 2799-2808.
34. Amati B, Littlewood T.D.; The c-Myc protein cell cycle progression and apoptosis through dimerization with Max: EMBO J, 1993, 12, 5083-5087.
35. Antonis S, et al.; Mxi1- a protein that specifically interacts with Max to bind Myc-Max recognition sites: Cell, 1993, 72, 223-232.
36. Lewis C, et al; Oncogene and signal transduction: Cell, 1991, 64, 281-302.
37. Findeis M.A.; Genes to the rescue: Technology review, 1994, 97, 46-53.
38. National cancer institute; Questions and answers about gene therapy: CancerNet, 1995, Internet.
39. Christiansen H, et al.; Neuroblastoma-inverse relationship between expresion of N-myc and NGF-r: Oncogene, 1990, 5, 437-440.
40. Le Van Thai A, Costa E; Identification of a neuron-specific promoter of human aromatic L-amino acid decarboxylase gene: Mol. Brain Res., 1993, 17, 227-238.

5.3. Books

41. Franks L.M., Teich N.M.; Introduction to the cellular and molecular biology of cancer, Oxford Univ. press, 2nd ed., 1988.
42. Evans C.W.; The metastatic cell-Behaviour and Biochemistry: Chapman and Hall, 1st ed., 1991.
43. Brown T.A.; Gene cloning - an introduction: Chapman and Hall, 3rd ed., 1995.
44. Lewin B; Genes V: Oxford University press, 1994.

```
5u
gb_pr:hummna3c7b

LOCUS       HUMMHA3C78    8892 bp    DNA             PRI        31-MAR-1995
DEFINITION  Homo sapiens MHC class I alpha chain gene, leader exon, alpha 1 and
            2 exons (HLA A3, C7, B7, BfS, C4A3, C4B1, DR2, DQ6).
ACCESSION   L29411
KEYWORDS    cell surface antigen; cell surface glycoprotein; class I gene;
            integral membrane protein; major histocompatibility complex. . . .

SCORES           Init1:   431 Initn:   733 Opt:   592
            70.9% identity in 316 bp overlap

            130       140       150       160       170       180
5u        TTTTTAAAAAATACATAAGCAAATCCTATATTACAAT-TACTAGAACCTGGCCAGTCATG
                                        ::: ::: :  : ::   : :   : ::
hummna    GTTTGTCTTTTTGATTCTGGCTTATTTCACTTATAATATTTTTGAGGTTCGGTGGGCACA
          5820      5830      5840      5850      5860      5870

            190       200       210       220       230       240
5u        GTGGCTCATGCCTATAATCCCAGCACTTTGGGAGGCTGGAGTGGGTTAATTGCTTGAGCC
          :::::::: ::::  : : ::::::::::::::::::: ::  :::  ::  : ::::
hummna    GTGGCTCACGCCTGGATTTCCAGCACTTTGGGAGGCTGAAGCAGGTGGATCACCTGAGTT
          5880      5890      5900      5910      5920      5930

            250       260       270       280       290       300
5u        TATGAACATGAGACGAGCCTGGTCAACATAGCAAAACCCCATCTTTAGAAAAAAATACTA
          :  ::    :: :: ::::::: :::::: :  :: :::::::: ::  :::::  :  :
hummna    TCGGAGTTCGAAACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATAATAA
          5940      5950      5960      5970      5980      5990

            310       320       330       340       350       360
5u        AAATTACCCAGCCATGGTGGCACATGCCCATAGTCCCAAATACTCAGGAGGCTGAGGCGG
          :: ::: :: : : :: ::::   ::::  :: :::::: ::::  :::::::::::: :
hummna    AAGTTAGCCGGGCGTGATGGCGGGTGCCTGTAATCCCAACTACTTGGGAGGCTGAGGCAG
          6000      6010      6020      6030      6040      6050

            370       380       390       400       410       420
5u        GAGGATTGCTTGAGCCTGGAAGGTGAAGGTTGCAGTGAGG--CAATCACACCACTGCACT
          ::: :: ::::::  : :: : ::: :: ::::::::::     ::::  ::::::::::
hummna    GAGAATCGCTTGAATCCGGGAAGTGGAGCTTGCAGTGAGCTGAGATCAGGCCACTGCACT
          6060      6070      6080      6090      6100      6110

            430       440       450       460       470       480
5u        CCAGCCTGACTGAC-AGAATAAGACCTCGTCTCGAAAAAAAATTATTGGAAGTTTTATAT
          :::::::: :  :: ::: : : :   : :::: ::::::::  ::
hummna    CCAGCCTGGGCAACAAGAGTGAAATTCCATCTCCAAAAAAAAAAATAAAACAATAATAAT
          6120      6130      6140      6150      6160      6170

            490       500       510       520       530       540
5u        TATTCTTCCCTCTAGTTGGACTTGCATTCCACTCCACTCGGCCAGCAGAACTTATCCTAT

hummna    AATAATATTTTCGAGGTTCATCCAAGTTGTAGTATGGGTCAGAATTTCATTCCTTTTAAG
          6180      6190      6200      6210      6220      6230
```

NTG +5

```
LOCUS       HSRB2P130     3249 bp    RNA             PRI       14-MAR-1994
DEFINITION  H.sapiens mRNA for Rb2/p130 protein.
ACCESSION   X74594
NID         g397147
KEYWORDS    E1A binding protein; retinoblastoma-associated protein.
SOURCE      human. . . .

SCORES          Init1:    63 Initn:    111 Opt:    112
            55.7% identity in 219 bp overlap

                                  579       569       559       549
5sc3                        GAAGAAGTAAAACTGGAGGGAAATAATCTATAATTTCTTACT
                                          ::::::::::: ::::          : : :
hsrb2p  AGCTACCGCAGCATGAGCGAAAGCTACACGCTGGAGGGAAAGATCT--------TCATT
            40        50        60        70        80

            539       529       519       509       499       489
5sc3    GGGAAAAATGGGACTGTTTTTCATGTATTCATGGGTAAGTATTATTTACGAATGGGATTG
        ::  :  :::    ::  ::  ::::    : ::    ::  :  :  :  :  :  ::  :  :
hsrb2p  GGTTAGCATG----TGCCTTATATGT-GGCTTGCAGAA--AATCTGTTCCAACTGTAAGC
          90        100       110       120       130

            479       469       459       449       439       430
5sc3    AAAGAGAACAGGATTCGGCATCAATTTTATTTTTAAGAGAAT-TTGGTATGGGCAGA--A
        :::: : ::::      ::  : : :: :    ::: :  :::::  ::  :::  :::: :
nsrb2p  AAAG-GGACAGTGGAAGGAAACTATGTATCTTTAACTAGAATCCTGAAATGTTCAGAGCA
           140       150       160       170       180       190

       429       419       409       399       389       379
5sc3    GTCCCTGAT--AAATATTGGTTGGTATAGACAAGTGGATAGAAATGGCAAGTGTCTGGTT
        :  : : ::  :: : ::    :    :: : :::::: ::: ::::::: : :
nsrb2p  GAGCTTAATCGAATTTTTTAATAAGATGAAGAAGTGGGAAGACATGGCAAATCTACCCCC
          200       210       220       230       240       250



        369       359       349       339       329       319
5sc3    ATATTTCAGTTACTCGCTTATCTGAACTCTTTCATATTATATGCCAAAACATGCATACAG
        : :::::::
nsrb2p  ACATTTCAGAGAACGTACTGAGAGATTAGAAAGAAACTTCACTGTTTCTGCTGTAATTTT
          260       270       280       290       300       310

LOCUS       HUMCD36A      2561 bp    DNA            PRI        03-MAR-1995
DEFINITION  Human antigen CD36 gene, alternate exons 2a and 2b and exon 3.
ACCESSION   L06849
NID         g691762
KEYWORDS    Alu repeat; antigen CD36.
SOURCE      Homo sapiens (tissue library: EMBL 4) DNA. . . .

SCORES          Init1:    69 Initn:    112 Opt:    69
            94.7% identity in 19 bp overlap

          410       420       430       440       450       460
5sc3    CCAACCAATATTTATCAGGGACTTCTGCCCATACCAAATTCTCTTAAAAATAAAATTGAT
                                      ::::::::::: ::::::
humcd3  AACATGACTCCATTGCTGTCTTAAATATAAATACCAAATTCTATTAAAAGCTGTCTACAG
           2070      2080      2090      2100      2110      2120

          470       480       490       500       510       520
5sc3    GCCGAATCCTGTTCTCTTTCAATCCCATTCGTAAATAATACTTACCCATGAATACATGAA

humcd3  GTATGCATGTTAGTAGAAATAATTGTTTTAAGTTATGTCCAAAGAGCATGTTGGCATGCT
           2130      2140      2150      2160      2170      2180
```

NTG +5

```
gb_pr:humretblas

LOCUS       HUMRETBLAS 180388 bp    DNA             PRI       23-NOV-1994
DEFINITION  Human retinoblastoma susceptibility gene exons 1-27, complete cds
ACCESSION   L11910
KEYWORDS    nuclear protein; recessive oncogene; retinoblastoma gene;
            retinoblastoma protein; retinoblastoma susceptibility;
            tumor supressor gene. . . .

 SCORES          Init1:    76 Initn:    161 Opt:    117
            52.6% identity in 441 bp overlap

                  10        20        30        40            50
8r          GATCATTTTGAAATCTTTGACCTTTATATTTTAATGAATAAAATATT-----AGT
                                     ::  : :  :::::::::::::        :
humret TATGATTAGTAGTTTTAATTCAAAATAAAAATGGTATGTTGAATAAAATATTACATAAAC
          107790    107800    107810    107820    107830    107840

            60        70        80           90         100
8r     AGTTATTAGTATAAAATAATTTATGTCTTTT---GGACTTAGCAT--CCAGTATTTCTTT
       :  :::::::::::::::::: ::    ::      : :::  ::     :: : :::
humret ATAAATTAGTATAAAATAATTTA-GTAAATTAAATCATTTAATATAAAAAGGAATTCAAA
          107850    107860    107870    107880    107890    107900

          110           120       130       140       150       160
8r     TTTAATAAAGA----AAATAATTATTCTCTTGCAATATACTGTGTTTACCTGGGTTTTGA
        : :: :::       :  :  :: : :    :: :  ::: :         ::    :
humret GACATTACAGATTGGCACCACATAATTTAAGTCAGT-CACTTTTAGACACTAAATAACTT
           107910    107920    107930    107940    107950    107960

              170       180       190       200       210       220
8r     AAAGTGATGTTTCCTAATATGAGAAAGCCATTTACATTTTTAAATCTACAAAGGCAAATG
        ::: ::: :::  ::  ::::   : : : : : :     :   ::    :::  :: :::
humret TAAGAGATTTTTTTTA--ATGAAGGAACAAATCAAA---ATGGCTCAGAAAAATCAGATG
            107970      107980    107990      108000    108010

               230       240       250       260       270      279
8r     GAATGG-TAC-TAAATTATTTATATAATAATGTTTAGATGGTGGCCCTTATAACATTCTT
       :: ::: :::  :::: :  :: ::  ::::: :::          :  : :  :::   ::
humret GAGTGGATACACAAATAAAATACATGTTAATGCTTA-----ACACATTGAATACAAATTT
     108020    108030    108040    108050          108060    108070

    280       290       300       310       320       330       339
8r     TCTATACTTCCTACAGAGTTGGGGATATGCAATCCTGGAATATTTCTGGGAGCTAATCCT
       ::: ::  : ::: ::: ::      :: : : : ::      ::: : : : :::: :
humret TCTTTA--TACTAAAGACTTTAAAATGT-CCATGTGTTAATTTCTTTTGGAGGTGGAAAA
               108080    108090    108100    108110    108120

    340       350       360        370       380          390
8r     TTAGCTTGATGAATGAAACAAGACTTTT-AAATAAAATTAAACTTTCAAAT--TATCCAG
        ::: :::    :    :::  : ::::::: : :     :    :::::  :: :  : :::
humret ATAGTTTGTCCA---AAAAGACACTTTTCACAGTTGAAGGAACTTGAAAGTTCTGTCCCA
    108130    108140         108150    108160    108170    108180

          400       410        420       430       440       450
8r     GTAATGGGCCT-GTCTTTTAATTCAATGGATATGGAGCATAATGAATTATCCCCTGTTCA
       :: :   : :::  :  ::: :::  : : :     :    :::  : :: ::::
humret GTGA--GTCCTAATGGTTTTATTTCAGGCAGCAGATTCATTGTCAAATATCTTACTTTTT
          108190    108200    108210    108220    108230    108240

          460       470
8r     TTGGGTAATAAGTTCTCATT

humret AAGGTCTGTAGGTTATGCTGAATAAAATTCTCTGCACCATGAACTTCAGAGAATCTGAAG
          108250    108260    108270    108280    108290    108300
```

NTG +8

```
gb_pr:humretblas

LOCUS       HUMRETBLAS 180388 bp    DNA             PRI       23-NOV-1994
DEFINITION  Human retinoblastoma susceptibility gene exons 1-27, complete cds.
ACCESSION   L11910
KEYWORDS    nuclear protein; recessive oncogene; retinoblastoma gene;
            retinoblastoma protein; retinoblastoma susceptibility;
            tumor supressor gene. . . .


SCORES          Init1:   118 Initn:   192 Opt:   174


            86.4% identity in 66 bp overlap

             380       390       400       410       420       430
3u      AATCTCGGCCTTTCCCGGGTCAGGGTTCATTATACACCAGGGGAGTACTATGCAGCCATA
                                     ::::::::: ::: :::::::::::::::
humret TGATAGACTGGATTAAGAAAATGTGGCACATATACACCA-TGGAATACTATGCAGCCATA
   168180    168190    168200    168210    168220    168230

             440       450       460
3u      AAAATGGATGAGTTCATGGCCTTTACTAGGGGACAT
        :::: :::::::::::: ::::: :: :::::::
humret AAAAATGATGAGTTCATGTCCTTT-GTA-GGGACATGGATGAAATTGGAAACCATCATTC
    168240    168250    168260    168270    168280    168290

humret TCAGTAAACTATCGCAAGAACAAAAAACCAAACACCGCATATTCTCACTCATAGGTGGGA
      168300    168310    168320    168330    168340    168350


     gb_pr:humvcama

     LOCUS       HUMVCAMA      2396 bp     DNA            PRI       14-JAN-1995
     DEFINITION  Human vascular cell adhesion molecule-1 (VCAM1) gene, exon 1.
     ACCESSION   M92431
     KEYWORDS    vascular cell adhesion molecule-1.
     SOURCE      Homo sapiens DNA.
       ORGANISM  Homo sapiens . . .


     SCORES          Init1:    62 Initn:   138 Opt:    92
                 61.5% identity in 117 bp overlap

                          10        20        30        40        50
     3r          GATCATTTTGAAATCTTTGACCTTTATATTTTAATGAATAAA-ATA----TTAGTA
                                          :::::: ::::::::: :::    : :: :
     humvca CAAAAAGTTGATGTTTGTTGCTAAAAGAACTATTTTTATGAATAAATATAAAACTAAGAA
                1610      1620      1630      1640      1650      1660

                 60        70        80        90       100
     3r     GTTAT-TAGT--ATAAAATAATTTATGTCTTTTGGACTTAGCATCCAGTATTTCTTTT--
            :::::  ::    :    :::   ::: ::  :  :    :: :    : : :::: :::
     humvca GTTATGGTGTCCCTTTTTTAAAAAATGACTCATCAA--AAGAAATAACTTTTTCCTTTCT
                1670      1680      1690         1700      1710

             110       120       130       140       150       160
     3r     -TTAATAAAGAAAATAATTATTCTCTTGCAATATACTGTGTTTACCTGGGTTTTGAAAAG
             ::   : :::::: ::::: ::::::
     humvca CTTGTAAGAGAAAAAAATTAATCTCTTTTAGAATTGCAAACATATTTCCTTGATGGAGAA
          1720      1730      1740      1750      1760      1770
```

NTG +8

```
13u
gb_pr:humtkra

LOCUS       HUMTKRA      13500 bp    DNA             PRI       14-JAN-1995
DEFINITION  Human thymidine kinase gene, complete cds, with clustered Alu
            repeats in the introns.
ACCESSION   M15205 M15206
KEYWORDS    Alu repeat; repeat region; thymidine kinase.
SOURCE      Human DNA (library of Y.-F.Lau), clone lambda-tk46 [1]. . . .


SCORES          Init1:   415 Initn:   573 Opt:   594
            78.8% identity in 240 bp overlap

                                                  10        20        30
13u                                     GATCAGTTGAGGGCAGCAGTTCGAGACCAA
                                        ::::: :::::: ::: :::: :::::::
humtkr  TCCCAGCACTTTGGAAGGCCTAGGCGGGCGGATCACTTGAGGTCAGGAGTTTGAGACCAG
          9380      9390      9400      9410      9420      9430

              40        50        60        70        80        90
13u     CCTGGACAACATGGTGATACCCCATCTGTACTGAAGAATACAAAATTAAGTCAGGCGTGG
        :: :: ::::::::::: ::::::::: :::: :: ::::::::: : :: :: ::::::
humtkr  CCAGGCCAACATGGTGAAACCCCATCTCTACT-AAAAATACAAAAATTAGCCAAGCGTGG
          9440      9450      9460      9470      9480      9490

             100       110       120       130       140       150
13u     TGGTGCACGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCACAAGAATGACTTGAACCC
        :::  ::: :::::::: ::::::::: ::::::: :::::::::::::  ::::::::
humtkr  TGGCACACACCTGTAATTCCAGCTACTTGGGAGGCAGAGGCACAAGAATTGCTTGAACCT
           9500      9510      9520      9530      9540      9550

             160       170       180       190       200       210
13u     CAGAGGAAGAAGTTGCAGGGAGCCAACTTTGTACCACTGTACTTCAGGCTGGGTGACAGA
          :::: ::: ::::::: :::::::   :: : :::::: ::: ::: :::: :::::::
humtkr  GGGAGGCAGAGGTTGCAGTGAGCCAAGATTATGCCACTGCACTCCAGCCTGGATGACAGA
           9560      9570      9580      9590      9600      9610

             220       230       240       250       260       270
13u     ACGACACTATAAATAATAATACTATGATAATATACCAGAGAGAACGAAAAACTGAGTTAA
         ::: ::: :    : : :: :      :: ::
humtkr  GCGAGACTCTGTCTCAAAAAAAATAGACAAAGCCAGGCGCAGTGGCTCATGCCTGTAATC
           9620      9630      9640      9650      9660      9670
```

NTG +13

```
13u /rev
go_pr:hshw2

LOCUS       HSHW2        22677 bp    DNA              PRI       12-MAY-1995
DEFINITION  Human DNA sequence from cosmid HW2, Huntington's Disease Region
            chromosome 4p16.3.
ACCESSION   Z49250
KEYWORDS    4p16.3; CpG island; helix-loop-helix; transcription repressor.
SOURCE      human. . . .



SCORES           Init1:   506 Initn:   506 Opt:   547
            79.4% identity in 218 bp overlap

               239       229       219       209       199       189
13u     CTGGTATATTATCATAGTATTATTATTTATAGTGTCGTTCTGTCACCCAGCCTGAAGTAC
                                      ::: : : :::::: ::::: ::  ::: :
hsnw2   GCTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTTGCTCTGTCGCCCAGGCTAGAGTGC
              9570      9580      9590      9600      9610      9620

               179       169       159       149       139       129
13u     AGTGGTACAAAGTTGGCTCCCTGCAACTTCTTCCTCTGGGGTTCAAGTCATTCTTGTGCC
        :::::    :  : ::::: ::::::: ::  ::::  ::::::::   :::::  ::::
hsnw2   AGTGGCGTGATCTCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAACGATTCTCCTGCC
              9630      9640      9650      9650      9670      9680

               119       109        99        89        79        69
13u     TCAGCCTCCCGAGTAGCTGGGATTACAGGCGTGCACCACCACGCCTGACTTAATTTTGTA
        :::::::::::::: ::::::::::::::::  ::::::: :::: : :: : :::::::
hsnw2   TCAGCCTCCCGAGTCGCTGGGATTACAGGCGCCCACCACCGCGCCCGGCTAATTTTTGTA
              9690      9700      9710      9720      9730      9740

                59        49        39        29        19         9
13u     TTCTTCAGTACAGATGGGGTATCACCATGTTGTCCAGGTTGGTCTCGAACTGCTGCCCTC
        :: :: :::: ::::::::: ::::::: ::: ::::: :::::::::::: ::: ::::
hsnw2   TTTTTTAGTAGAGATGGGGTTTCACCATCTTGGCCAGGCTGGTCTCGAACTCCTGACCTC
              9750      9760      9770      9780      9790      9800

13u     AACTGATC
        :  :::::
hsnw2   A--TGATCCCCCCCCCCCCCCCCCCCCCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAG
              9810      9820      9830      9840      9850
```