Transcriptional and Post-Transcriptional Control of Therapeutic Gene Expression during Disease Activity

By

Hodan Hassan Ahmed Mohamed

Bone and Joint Research Unit

William Harvey Research Institute

Barts and the London School of Medicine

Queen Mary, University of London, UK

A thesis submitted to the University of London

for the degree of Doctor of Philosophy

August 2014

Declaration

I confirm that all experiments were performed by myself (unless stated otherwise in the text), under the supervision of Dr David Gould and Professor Yuti Chernajovsky. To the best of my knowledge, the thesis comprises only my original work towards the PhD and due acknowledgment has been made in the text to all other materials used. I confirm that no part of this thesis has been submitted for any degree or qualification of any other university or institute of learning.

Hodan Hassan A. Mohamed

August 2014

Acknowledgements

First, I would like to sincerely thank the Medical Research Council (MRC) for funding my MRes and PhD studies during this four year MRes/PhD Studentship. I extend my gratitude to the British Pharmacological Society (BPS) for their additional funding.

Second, I am truly fortunate to have been supervised by Dr David Gould and Professor Yuti Chernajovsky. This work would not have been possible without their continuous support, advice, guidance, constructive criticism and encouragement.

Third, I am very lucky to have been a member of the Bone and Joint Research Unit (BJRU). Many thanks are due to past and present BJRU members for their kindness, advice and moral support. I am sure future BJRU members will have the same pleasant experience as I have had.

Fourth, a special thanks is due to Gill Adams, our lab manager, for proof-reading this thesis and for her endless help and encouragement during my studies.

Fifth, I would like to thank Dr Martin Carrier, who played an integral role during my MRes degree. Without his invaluable advice and guidance, my progression to do the PhD would not have been possible.

Last but not least, I am most grateful to my mother and sisters for their endless support, encouragement and extraordinary patience. My success is your success.

This thesis is dedicated to my late father, Dr Hassan A. Mohamed, who taught my sisters and I, that a good education lasts a lifetime.

Abstract

Rheumatoid arthritis (RA) is a chronic inflammatory autoimmune disease which predominantly affects the synovial joints. Local gene therapy represents an approach to produce therapeutic molecules (i.e. soluble TNF receptor (sTNFR)-Fc and interleukin-1 receptor antagonist (IL-1Ra)) directly in arthritic joints. Gene therapy could be designed to link the level of therapeutic gene expression directly to disease activity, through the use of transcriptional and post-transcriptional regulatory elements. The experiments in this thesis describe the construction of multi-responsive, composite synthetic promoters, comprised of the binding sites for an array of transcription factors activated in arthritic joints. Optimal spatial arrangements of binding sites in relation to each other and to the TATA box were determined by Assembly PCR cloning and the functionality of the resulting synthetic promoters revealed additive or synergistic induction of luciferase reporter gene expression in response to combined stimulation. Candidate synthetic promoters were cloned into a lentiviral vector between insulator elements and displayed significantly enhanced induction, in excess of 1,500 fold in response to combined stimulation. Inflammation-specific activation of lentiviral synthetic promoters was confirmed in a carrageenan-induced paw inflammation mouse model, which demonstrated the strong correlation between local luciferase gene expression and paw inflammation.

Post-transcriptional gene regulation was also investigated by exploiting the differential expression of endogenous miR-23b during inflammation. Insertion of miR-23b target sites into the 3'UTR of the luciferase gene subjected luciferase mRNA to regulation by miR-23b. Experiments demonstrated that high basal gene expression driven by constitutive and inducible promoters was significantly downregulated by miR-23b without significantly impairing high gene expression upon stimulation. Overall, the experiments in this thesis have confirmed the induction of inflammation-specific gene expression, regulated by inflammation-responsive endogenous transcriptional and post-transcriptional elements.

Table of Contents

Chapter 6: Post-Transcriptional Regulation of Luciferase Gene Expression by Inflammation-responsive miR-23b   284-312

List of Figures

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 6

Abbreviations

α	Alpha
β	Beta
2'-OMe	2'-O-methyl
3'UTR	3'-untranslated region
5'-cHS4	5'-end of the chicken β-globin locus
AAV	Adeno-associated virus
ACPA	Anti-citrullinated protein antibody
ACR	American College of Rheumatology
ADAMTS	A disintegrin and metalloproteinase with thrombospondin motifs
Ago	Argonaute
AIA	Antigen-induced arthritis
AIDS	Acquired Immune Deficiency Syndrome
AMO	Anti-miRNA oligonucleotide
ANOVA	Analysis of variation
Anti-CCP	Anticyclic citrullinated peptide
AP-1	Activator protein-1
APC	Antigen presenting cell
ARD	Ankyrin repeat domain
AS	Annealing sequence
Bcl-2	B-cell CLL/lymphoma 2
bHLH	Basic helix-loop-helix
BLAST	Basic local alignment search tool
BLP	Bacterial lipopeptide Pam_3CSK_4
bp	Base pair
BSA	Bovine serum albumin
bZIP	Basic leucine zipper

C.elegans	*Caenorhabditis elegans*
C/EBP	CCAAT/enhancer-binding proteins
C3	Complement factor 3
CA	Capsid
CCCD	Cooled charged-coupled device camera
CDK2	Cyclin-dependent kinase 2
cDNA	Complimentary DNA
CIA	Collagen-induced arthritis
CIP	Calf intestinal phosphatase
CMV	Cytomegalovirus
COMP	Cartilage oligomeric matrix protein
COX	Cyclooxygenase
COXIB	COX-2 inhibitor
CpG	Cytosine and guanine separated by a phosphate
cPPT	Central polypurine tract
CRP	C-reactive protein
Ct	Cycle threshold
CTS	Central termination sequence
DAMP	Damage-associated molecular pattern
DAS28	Disease Activity Score in 28 joints
DGCR8	DiGeorge Syndrome Critical Region Gene 8
dH_2O	Distilled H_2O
DMARD	Disease modifying anti-rheumatic drug
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethylsulfoxide
Dox	Doxycycline
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA

dsTNFR	Dimeric soluble TNF receptor
E.coli	*Escherichia coli*
EAE	Experimental autoimmune encephalomyelitis
eGFP	Enhanced green fluorescent protein
Egr-1	Early growth response-1
EPO	Erythropoietin
ERK	Extracellular signal-regulated kinase
ESEL	E-selectin
ESR	Erythrocyte sedimentation rate
Ets	E26 transformation-specific
FADD	Fas-associated death domain
Fc	Fragment crystallisable region
FLS	Fibroblast-like synoviocyte
GM-CSF	Granulocyte-macrophage colony stimulating factor
gRNA	Genomic RNA
HEK	Human embryonic kidney
HGF	Hepatocyte growth factor
HIF-1α	Hypoxia inducible factor-1α
Hi-FBS	Heat-inactivated foetal bovine serum
hIL-17A	Human IL-17A
HIV-1	Human immunodeficiency virus-1
HLA	Human leukocyte antigen
HOMECs	Human omental microvascular endothelial cells
HRE	Hypoxia-responsive element
hsa	Homo sapien
HSCs	Haematopoietic stem cells
HSP70	Heat-shock protein 70
HSV	Herpes simplex virus

HUVECs	Human umbilical vein endothelial cells
HVJ	Hemagglutinating virus of Japan
i.a.	Intra-articular
i.m	Intramuscular
ICAM-1	Intercellular adhesion molecule-1
IFU	Infectious units
Ig	Immunoglobulin
iGEM	International genetically engineered machine
IKKα	IκB-kinase α
IL	Interleukin
IL-1Ra	Interleukin-1 receptor antagonist
IN	Integrase
IRAK1	IL-1 receptor-associated kinase 1
IκB	Inhibitor of κB
JDP	Jun dimerisation partners
LAP	Liver-enriched transcriptional activating protein
LB	Luria bertani
LFA-1	Leukocyte function-associated antigen-1
Lin-4	Lineage-4
LIP	Liver-enriched transcriptional inhibitory protein
LNA	Locked nucleic acid
LP	Lentiviral particle
LTR	Long terminal repeat
Luc$^+$	Luciferase gene
LV	Lentiviral vector
MAPK	Mitogen-activated protein kinase
MAR	Matrix attachment region
mCMV	Minimal CMV

MCP-1	Monocyte chemotactic protein-1
M-CSF	Macrophage colony stimulating factor
MHC	Major histocompatibility complex
mIL-17A	Mouse IL-17A
miRNA	MicroRNA
MMP	Matrix metalloproteinase
mRNA	Messenger RNA
MTX	Methotrexate
NAB	NGFI-A binding protein
NEB	New England Biolab
Nef	Negative regulatory factor
NEMO	NFκB essential modulator
NFAT	Nuclear factor of activated T cells
NF-IL6	Nuclear factor for IL-6 expression
NGF	Neuronal growth factor
NICE	National Institute for Health and Care Excellence
NIK	NFκB inducing kinase
NPC	Nuclear pore complex
NSAID	Non-steroidal anti-inflammatory drugs
ODD	Oxygen-dependent degradation domain
ODN	Oligodeoxynucleotide
Ori	Origin of replication
PAMP	Pathogen-associated molecular pattern
PAS	Per/ARNT/Sim
Pasha	Partner of Drosha
PBMC	Peripheral blood mononuclear cell
PBS	Primer binding site
PCA	Polymerase chain assembly

P-CAF	p300/CREB-binding protein-associated factor
PCR	Polymerase chain reaction
PGE$_2$	Prostaglandin E2
PHD	Prolyl hydroxylase enzyme
PIC	Preinitiation complex
PIC	Preintegration complex
piRNA	Piwi-interacting RNA
PMA	Phorbol-12-myristate-13-acetate
PPT	Polypurine track
PR	Protease
pre-miRNA	Precursor miRNA
pri-miRNA	Primary miRNA
pVHL	Von hippel-lindau protein
RA	Rheumatoid arthritis
RANKL	Receptor activator of nuclear factor B ligand
RASF	RA synovial fibroblast
RER	Rough endoplasmic reticulum
Rev	Regulator of virion protein expression
RF	Rheumatoid factor
RHD	Rel homology domain
RISC	RNA-induced silencing complex
RLU	Relative light unit
RNAi	RNA interference
ROI	Regions of interest
RT	Reverse transcriptase
RTC	Reverse transcription complex
rtTA	Reverse Tet-repressor
Saa3	*Serum amyloid A3*

SAPK	Stress-activated protein kinase
SCW	Streptococcal cell wall
SEM	Standard error of the mean
SFFV	Spleen focus forming virus
shRNA	Short hairpin RNA
SIN	Self-inactivating
siRNA	Small interfering RNA
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism
sTNFR	Soluble TNF receptor
stRNA	Short temporal RNA
TAB	TAK1-binding protein
TAD	Transactivation domain
TAF	TBP-associated factor
TAR	Transactivation-responsive region
Tat	Transactivator of transcription
TBP	TATA-binding protein
TCR	T-cell receptor
Tet	Tetracycline
TetO	Tet operator
TetR	Tet repressor protein
TetR-KRAB	Tetracycline repressor-kruppel associated box
TF	Transcription factor
TFBS	Transcription factor binding site
TIMP	Tissue inhibitors of metalloproteinase
TK	Thymidine kinase
TLR	Toll-like receptor
TNF	Tumour necrosis factor

TPA	12-O-tetradecanoylphorbol-13-acetate
TRAF6	TNF receptor-associated factor 6
TRBP	Transactivating response RNA-binding protein
TSS	Transcriptional start site
tTA	Tetracycline-controlled transactivator
TU	Transducing units
u-Pa	Urokinase-type plasminogen activator
v/v	Volume per volume
VCAM-1	Vascular adhesion molecule
VEGF	Vascular endothelial growth factor
Vif	Virion infectivity factor
Vpr	Viral protein R
Vpu	Viral protein U
VSV-G	Vesicular stomatitis virus glycoprotein
w/v	Weight per volume
w/w	Weight per weight
WPRE	Woodchuck post-transcriptional regulatory element
WT	Wild type
YFP	Yellow fluorescence protein

CHAPTER 1:

Introduction

1.1. Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a multifactorial, autoimmune and systemic disease characterised by chronic and progressive inflammation. The characteristic features of RA include inflammatory cell infiltration into the joint and hyperplasia of the synovial membrane, resulting in structural damage to the bone, cartilage and ligaments within the synovial joints (Iwamoto *et al.,* 2008).

1.1.1. Epidemiology

Over the past two decades, population-based studies conducted in various geographically and ethnically diverse populations have consistently estimated the prevalence of RA in the adult population to be approximately 1% worldwide (Gabriel and Michaud, 2009). Using the 1987 American College of Rheumatology (ACR) classification criteria for RA, which predominantly identifies patients with active RA (Arnett *et al.,* 1988), Symmons and colleagues estimated a 0.8% prevalence of RA in a Norfolk population study, equating to approximately 400,000 people in the UK suffering from the disease (Symmons *et al.,* 2002, Rheumatoid Arthritis; National Clinical Guideline for Management and Treatment in Adults, UK, 2009). Epidemiological studies have generally demonstrated the RA incidence to be low, but importantly these studies highlight another feature of RA, which is the striking imbalance between the sexes, as women are typically three times more likely to develop RA than men. The reasons for this overrepresentation of women are not clear however, hormonal and X-linked genetic factors are likely to be involved (van VollenHoven, 2009).

In 2006, Alamanos and colleagues conducted a systematic review of incidence and prevalence studies of RA from January 1988 to December 2005 across northern American countries, north and south European and developing countries. Although there were substantial variations in the incidence and prevalence across the various studies and time periods (potentially due to methodological limitations), the considerable decline in RA

incidence, with a prominent shift toward an elderly age of onset, was a consistent observation (Alamanos *et al.,* 2006). Similarly, the peak age of incidence in the UK for both genders is ~70 years, although all ages can develop RA. Despite extensive research, the aetiology of RA remains obscure, however, the contributions of genetic and environmental risk factors in the development of RA have been highlighted by epidemiological studies.

1.1.1.2. Genetic factors

Although the prevalence of RA has consistently been estimated at 1% worldwide, the variation in the prevalence of RA among different ethnic groups has supported a genetic role in disease risk. Genetic and epidemiological data have implicated the major histocompatibility complex (MHC) class II molecules in the pathogenesis of numerous autoimmune diseases. The human leukocyte antigen (HLA), located on the short arm of chromosome 6, is a gene-rich area encoding for MHC where a large number of these genes are related to immune functions in humans. Namely, the MHC class II molecules, located on the surface of antigen-presenting cells, are central to the adaptive immune system as these molecules display various peptides for recognition by the T-cell receptors of CD4$^+$ T helper cells (Jones *et al.,* 2006). The association of certain MHC class II HLA-DRB1 alleles with RA has long been established (Gregersen *et al.,* 1987) which has since been consistently identified as a RA susceptibility gene in many populations around the world (The Wellcome Trust Case Control Consortium, 2007). The HLA-DRB1 alleles code a five amino acid sequence motif (QKRAA) on the HLA-DRβ chain, termed the 'shared epitope', which is carried by most patients with RA and is thought to affect the antigen presentation of specific peptides to T-cell receptors. Therefore, the presentation of arthritis-related peptides by disease-associated HLA-DRB1 alleles may result in the expansion of autoantigen-specific T-cells in the joints and lymph nodes of RA patients (Choy, 2012).

Data from twin studies have reported that only ~60% of the genetic contribution to RA can be implicated by genetic factors (MacGregor *et al.,* 2000), which encouraged the search for non-

MHC genes and highlighted the potential influence of non-genetic factors. Genome-wide association studies using single nucleotide polymorphisms (SNPs) have also indicated that in addition to genes in the HLA region, numerous other non-MHC genes may be involved in disease susceptibility, particularly those encoding immune regulatory factors. For example, the PTPN22 gene encodes a protein tyrosine phosphatase which is involved in the negative regulation of T and B-cell activation via the T cell receptors and B cell antigen receptors, and has been reproducibly associated with RA. Other putative RA susceptibility genes have been associated with cytokine signalling such as the IL-2 gene, which plays an important role for T-cell homeostasis and survival and also the TNFAIP3 gene, which is a negative regulator of NFκB and inhibitor of the effects of TNF-receptor mediated signalling (Ruyssen-Witrand *et al.,* 2012). Although major advances have been made in identifying RA susceptibility genes both within and outside of the MHC, epidemiological data suggests an important role of non-genetic factors in RA aetiology.

1.1.1.3. Non-genetic factors

Population studies conducted within a country have been useful in separating the contribution of genetic and non-genetic factors in patients. For example, the high prevalence of disease in population subgroups typically suggests genetic associations whereas higher rates of disease in specific geographical areas are indicative of environmental influences. These observations were supported by epidemiological data reporting very high RA prevalence in Native American Pima Indians (5.3%) and in the Chippewa Indians (6.8%) which suggested that although a predominant genetic association can be assumed, the influence of social habits, living conditions and other environmental factors may also play a role in RA development (Silman and Pearson, 2002). Increased urbanisation has also been associated with increased RA prevalence in populations. For example, the Xhosa tribe of South Africa living in urban rather than rural environments demonstrated a higher prevalence of RA (Solomon *et al.,* 1975). Similar observations were also described in populations in urban, suburban and rural areas of Taiwan (Chou *et al.,* 1994).

Circulating RA associated autoantibodies such as rheumatoid factor (RF) and anticyclic citrullinated peptide (anti-CCP) have been detected more than 10 years before the onset of clinical disease (Nielen *et al.,* 2004) and there is increasing evidence suggesting that early environmental factors may influence the onset of RA before clinical disease and symptoms become apparent. Conversely, much indirect evidence has implicated exposure to infectious agents in the development of RA in adult life, including Epstein-Barr virus, Parvovirus and bacteria such as *Proteus* and *Mycoplasma*, however, no agent has been conclusively shown to be causative (Silman and Pearson, 2002).

Incidence cohorts have consistently demonstrated a high prevalence of RA in women compared to men which strongly suggests that sex hormonal factors are likely to modulate the susceptibility and course of RA (van VollenHoven, 2009). Further evidence includes the reduced risk of developing RA in women who regularly take the oral contraceptive pill (Brennan *et al.,* 1997) and the clinical fluctuations in RA symptoms during the menstrual cycle and pregnancy (Cutolo and Lahita, 2005). Interestingly, gene-environment studies have established compelling links between cigarette smoking, the shared epitope and anti-citrullinated protein antibody (ACPA) production, with an increased risk of RA (Silman and Pearson, 2002).

1.1.2. Clinical features

RA can affect any synovial joint but characteristically affects the small synovial joints of the hands (Fig 1.1) and feet (metacarpophalangeal, proximal interphalangeal and metatarsophalangeal joints) and usually progresses to the larger joints such as the knee, hip, elbow and shoulder, in a symmetrical fashion (Suresh, 2004).

Figure 1.1. A hand affected by rheumatoid arthritis. RA patient with joint swelling in the hands. Photograph courtesy of James Heilman, MD, via Wikimedia Commons. (This photograph is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported licence, which grants permission to copy this document under the terms of the GNU Free Documentation License, Version 1.2).

The classical manifestations of RA include pain (due to stretching of pain receptors in the tissues surrounding the joint), heat and sometimes redness (due to increased blood flow and inflammation of the joint), swelling (due to proliferation of the synovial membrane and increased synovial fluid) and joint stiffness (due to loss of muscle and increased pain and swelling which often results in loss of function). The degree of progressive damage is related to the intensity and duration of inflammation therefore, chronic inflammation consequently results in deformity, disability and multiple co-morbidities (Rheumatoid Arthritis; National Clinical Guideline for Management and Treatment in Adults, UK, 2009).

RA is a systemic disease and in addition to joint symptoms, the progression of RA from a self-limiting phase to a chronic phase can result in extra-articular or systemic manifestations or both (Hochberg *et al.,* 2008). Extra-articular manifestations include rheumatoid nodules, vasculitis, pericarditis, keratoconjunctivitis sicca, uveitis, subcutaneous and pulmonary nodule formation, while systemic manifestations include the production of acute-phase proteins, haematological abnormalities, cardiovascular disease, osteoporosis, fatigue and depression and subsequently a poorer quality of life (Choy, 2012). These extra-articular and systemic complications associated with chronic RA can lower the life expectancy of patients by 3-10 years (Amaya-Amaya *et al.,* 2013) as well as increasing the mortality rate, which is twice as high as the general population and appears to be increasing (McInnes and Schett, 2011; Gonzalez *et al.,* 2007).

The clinical diagnosis of RA is often initiated by a consultation regarding a history of joint swelling, early morning stiffness lasting more than 30 minutes and systemic symptoms such as tiredness, malaise fever, weight loss which is followed by examinations of the joints (Suresh, 2004). Although no blood test can conclusively diagnose RA, routine blood tests include the erythrocyte sedimentation rate (ESR) test which is based on the observation that red blood cells of patients with inflammatory conditions sink at a faster rate than normal red blood cells, the C-reactive protein (CRP) test which can indicate the presence of inflammation in the body based on raised CRP levels in the blood and also, serological tests which primarily detect the amount of the rheumatoid factor (RF) antibody present in the blood. The RF antibody is present in eight out of ten people with RA, however, this antibody has also been detected in one in twenty people who do not have RA, which can lead to false positive results. Early changes in the joint may precede the symptomatic onset of RA by many years, therefore, the combination of radiological scans, blood tests, consultations and examination of joints have enabled clinicians to efficiently diagnose RA in patients. However, diagnoses of early RA are not without challenges and common limitations include the lack of specific diagnostic tests, poor sensitivity in laboratory tests, 'normal' test results in patients with definite RA and general

inability to recognise early synovitis which often does not progress to RA (Rheumatoid Arthritis; National Clinical Guideline for Management and Treatment in Adults, UK, 2009).

Also, the morbidity and long-term disability associated with RA has imposed a considerable economic burden on both RA patients and the health services. The total UK costs including NHS expense, carer fees, nursing homes, private expenditure, sick leave and work-related disability are estimated to be approximately £3.8 - £4.75 billion per year (Rheumatoid Arthritis; National Clinical Guideline for Management and Treatment in Adults, UK, 2009). Therefore, early diagnosis of RA and administration of effective treatments are imperative in improving the quality of life of the patient and reducing the economic burden imposed on RA patients and the health services.

1.1.3. Pathogenesis of RA

The complex interaction between immune cells, cytokines, effector molecules and signalling pathways are fundamental to the development of the inflammatory process in the synovium, which is the primary site of pathology.

1.1.3.1. Anatomy of a healthy joint

Normal synovial joints are composed of two opposing bone surfaces covered in articular cartilage which is responsible for weight-bearing, shock absorbing and reducing friction. The articular surfaces are separated by a narrow joint cavity containing synovial fluid, an albumin and hyaluronic acid-rich fluid responsible for lubricating, nourishing and removing the waste from the articular cartilage. The joint capsule encloses the joint cavity and retains the synovial fluid and is comprised of the outer fibrous capsule, which is adjoined to the periosteum of the bones, and the inner synovial membrane. The synovial membrane is composed of an intimal lining, comprising one to three layers of cells loosely attached to the basement membrane, and the synovial sublining that is composed mainly of extracellular matrix, blood and lymphatic vessels, adipocytes and fibroblasts (Haywood and Walsh, 2001; Knedla *et al.,* 2007). The

synovial intimal lining consists of two mesenchymal-like cells; fibroblast-like type B synoviocytes (synovial fibroblasts), which produce matrix molecules and hyaluronan to increase the viscosity of the synovial fluid, and macrophage-like type A synoviocytes, which clear apoptotic neutrophils and other cells from the joint cavity (Shiozawa *et al.,* 2011).

1.1.3.2. Pre-articular phase of RA

As previously described, genetic predisposition and environmental (non-genetic) factors are believed to initiate the onset of RA by promoting the dysregulation of the immune system and breakdown of immune tolerance leading to autoimmunity, as indicated by the presence of autoantibodies against self-antigens (Song and Kang, 2010). The activation of the innate immune system is the earliest event in the pathogenesis of RA and occurs prior to the development of the clinical signs (Nielen *et al.,* 2004). This process initially involves the activation and maturation of dendritic cells in response to exposure to numerous triggers including bacterial and viral products, immune complexes, cytokines, multiple endogenous ligands and disruption of cell-cell contact (Hitchon and El-Gabalawy, 2011). The RA synovium contains abundant myeloid and immature plasmacytoid dendritic cells which are primarily located in the sublining tissue near T-cell-B-cell aggregates and in perivascular lymphocytic areas. These dendritic cells express cytokines (IL-12, -15, -18 and 23), HLA class II molecules, chemokine receptors and costimulatory molecules required for T-cell activation and antigen presentation (McInnes and Schett, 2011).

There is substantial evidence suggesting that invading T-cells, particularly CD4[+] helper cells are vital in the early immunological response. Antigen-presenting cells, including dendritic cells, macrophages and activated B-cells, present arthritis-associated antigens to T-cells leading to stimulation and expansion of autoantigen-specific T-cells in the joints and lymph nodes (Buch and Emery, 2002). In addition to antigen-presentation, B-cells which are primarily located in T-cell-B-cell aggregates, become activated by the autoantigen-specific T-cells to produce autoantibodies e.g. RF and anti-CCP autoantibodies. These autoantibodies can form

larger immune complexes resulting in complement fixation, neutrophil activation, and the production of pro-inflammatory cytokines which perpetuates inflammation (Hitchon and El-Gabalawy, 2011). Consequently, T-and B-cell activation results in increased production of pro-inflammatory cytokines and chemokines resulting in additional T-cell, macrophage and B-cell interactions. Specifically, the interaction between synovial macrophages and T-cells drives TNFα and IL-1β production and activated fibroblasts secrete IL-6, IL-8 and prostaglandin E2 (PGE$_2$). Furthermore, T-cells differentiate into Th17-cells which produce high levels of pro-inflammatory IL-17 cytokine in the synovium resulting in the perpetuation of inflammation and progression to the onset of clinical disease (Choy, 2012).

1.1.3.3. Articular phase of RA

The transition from the pre-articular to the articular phase of RA is a multi-step complex process. The key role of activated RA synovial fibroblasts (RASFs) in the pathogenesis of RA has become increasingly evident, as these cells significantly contribute to the destructive processes in RA (Ospelt *et al.,* 2004). During the pre-clinical phase, RASFs may become activated by infectious and non-infectious agents and their respective degradation products. Microbial fragments or endogenous ligands, such as RNA from necrotic cells within the synovium, can stimulate RASFs via Toll-like receptors (TLR-2, -3 and 4), found on the cell surface of RASFs (Seibl *et al.,* 2003). TLR signalling in RASFs results in upregulated expression of pro-inflammatory cytokines and chemokines leading to the attraction and accumulation of immune cells to the synovium and through a positive feedback loop, RASFs can initiate and perpetuate the disease process.

The characteristics of activated RASFs include changes to the normal spindle-shaped cytoskeleton, dense rough endoplasmic reticulum (RER) and large nucleus with prominent nucleoli indicating both active RNA metabolism and protein production (Huber *et al.,* 2006). These activated RASFs produce matrix metalloproteinases (MMPs) and various pro-inflammatory cytokines (IL-1β, IL-6, IL-7, IL-15, IL-16, LTβ, GM-CSF) and chemokines (MIP-

1α, MCP-1 and RANTES) that promote the recruitment and activation of other immune cells to the synovium (McInnes and Schett, 2007). The substantial influx and local activation of immune cells including T-cells, B-cells, plasma cells, dendritic cells, macrophages, mast cells and neutrophils into the synovium resulting in synovitis (Choy, 2012) and further infiltration of immune cells and the progression of synovitis is perpetuated by angiogenesis, which provides oxygen and nutrients to the expanding tissue (Paleolog, 2002). Additionally, the expanded endothelial cells contribute to local cytokine production and leukocyte recruitment, further exacerbating the prolonged inflammatory environment in the joint and persistence of the disease process (Szekanecz and Koch, 2000).

The first step of synovial invasion involves the attachment of RASFs to the articular cartilage, a process mediated by the upregulation of adhesion molecules on the surface of RASFs. Adhesion molecules, specifically integrins of the β1 subfamily, mediate the attachment of RASFs to fibronectin-rich sites of the articular cartilage, cartilage oligomeric matrix protein (COMP) and to collagens, namely collagen type II. Following adhesion of RASFs, integrins and other adhesion molecules such as vascular adhesion molecule (VCAM-1) activate signalling pathways within the cell involved in the regulation of the early cell cycle genes, such as c-Fos/AP-1 and c-myc, and activate the expression of MMPs, thereby promoting hyperplasia of the synovium and ultimately, cartilage degradation (Shiozawa and Tsumiyama, 2009). Synovial hyperplasia is a characteristic feature of RA resulting from the combined effect of RASF hyperproliferation and impaired apoptosis. The upregulated transcription factor c-Fos/AP-1 activates Wee1 kinase, which inhibits mitotic cell division by phosphorylating cdc2, resulting in arrested mitotic cell division and cellular proliferation (Kawasaki et al., 2003). RASFs also display altered characteristics in cell death pathways and a low rate of apoptosis. The resistance to apoptosis potentially results from increased expression of anti-apoptotic proteins such as FLIP and Bcl-2 and survival proteins such as heat-shock protein 70 (HSP70), sentrin-1 (also known as SUMO-1) and sumoylated proteins which prolong survival of RASFs (Huber et al., 2006). Consequently, the hyperplastic synovium develops into the characteristic

synovial overgrowth, commonly known as 'pannus'. This 'tumour-like' granulation tissue exhibits features of tumours such as hyperproliferation, evasion of apoptosis, anchorage-independence and loss of contact inhibition, all of which contribute to the overgrowth of the pannus and its gradual invasion into the adjacent articular cartilage and the underlying bone, resulting in cartilage degradation and bone erosion (Shiozawa *et al.,* 2011).

Articular cartilage is composed of a non-mineralised surface layer and a mineralised layer adjacent to the bone and both layers contain chondrocytes which are involved in cartilage metabolism (McInnes and Schett, 2007). The destruction of articular cartilage is a multi-step process mediated by the release of matrix-degrading enzymes such as aggrecanases 1 and 2 (also known as A Disintegrin and Metalloproteinase with Thrombospondin Motifs (ADAMTS-)-4 and 5) and MMPs. In healthy joints, MMP activity is balanced by endogenously produced tissue inhibitors of metalloproteinases (TIMPs) which inhibit MMPs by non-covalently binding to its enzymatic active site with a 1:1 stoichiometry. However in RA, pro-inflammatory cytokines such as TNFα and IL-1β, growth factors and matrix molecules induce the expression of MMPs via transcriptional activation from RASFs resulting in the amount of MMPs exceeding approximately x44 that of TIMPs (Shiozawa *et al.,* 2011). Further cartilage degradation is promoted by IL-1β and IL-17 which induce a switch in chondrocytes from an anabolic matrix-synthesising state to a catabolic state, characterised by the formation of ADAMTSs and MMPs. Chondrocytes also synthesise and respond to pro-inflammatory cytokines to accelerate the switch from an anabolic to a catabolic state (Otero and Goldring, 2007). Other proteinases, including urokinase-type plasminogen activator and the cathepsins B, L and D which contribute to cartilage destruction by degrading various cartilage matrix components and cathepsin K, are expressed by synovial fibroblasts, synovial macrophages, articular cartilage chondrocytes, osteoclasts and osteoblasts (Muller-Ladner *et al.,* 2007).

The primary mediators of bone destruction are osteoclasts, which are found at the interface between the pannus and the adjacent subchondral bone. The presence of M-CSF

(macrophage colony stimulating factor) and RANKL (receptor activator of nuclear factor B ligand) which are also expressed at the site of pannus invasion, enable the differentiation of osteoclasts from their precursors (McInnes and Schett, 2011). The osteoclastic activity of osteoclasts is initiated by the binding of RANKL, produced from activated lymphocytes and osteoblasts, to the cognate RANK receptor on their cell surface, and the activated osteoclasts along with cathepsin K, promote the local activation of bone resorption with destruction of the mineralised bone matrix, resulting in bone erosion (Gravallese, 2002).

Pro-inflammatory cytokines are involved in each phase of the pathogenesis of RA, from promoting autoimmunity (including during the pre-articular phase), to maintaining chronic inflammatory synovitis and also during the destruction of adjacent joint tissue (McInnes and Schett, 2007). The main cytokines involved in integrating the immune-regulatory and tissue destruction events in RA are summarised in Figure 1.2.
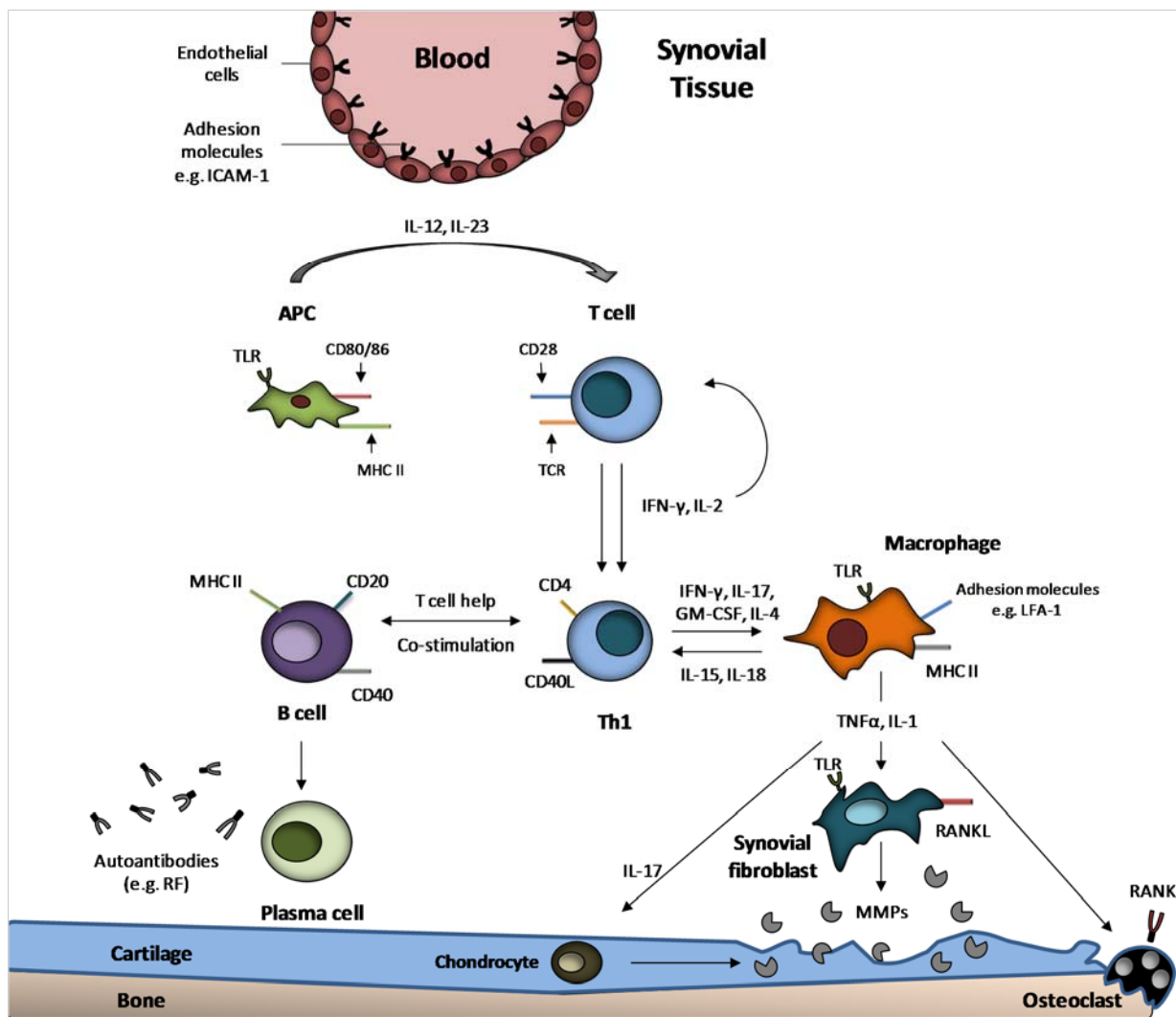
Figure 1.2. Schematic diagram depicting the events occurring in the RA joint. Cell surface adhesion molecules such as the intercellular adhesion molecule-1 (ICAM-1), present on endothelial cells, facilitate the trafficking of cells i.e. leukocytes, expressing leukocyte function-associated antigen-1 (LFA-1) between the blood and the synovial tissue. In the synovial tissue, pro-inflammatory cytokines such as IL-12 and IL-23 are produced by antigen-presenting cells (APC) that also possess co-stimulatory molecules (CD80/86) and the major histocompatibility complex (MHC) class II molecule which interact with CD28 and T-cell receptor (TCR), respectively, on the cell surface of T cells. Activated T cells secrete IFN-γ, IL-17, GM-CSF and IL-4. B-cells produce autoantibodies such as Rheumatoid Factor (RF). Macrophages secrete IL-15 and IL-18 and also produce TNFα and IL-1, which activate synovial fibroblasts and chondrocytes to secrete matrix metalloproteinases (MMPs) that contribute to cartilage degradation and activate osteoclasts involved in bone destruction. Figure redrawn, with modifications, from the image by Smolen and Steiner (2003).

1.1.4. Animal models of RA

Animal models of RA have proven to be very useful research tools for the study of the pathogenic pathways involved in the disease and also, for the development and preclinical evaluation of anti-arthritic therapies (Bevaart *et al.,* 2010).

Numerous mouse models of RA and/or paw inflammation have been established including collagen-induced arthritis (CIA) (Courtenay *et al.,* 1980), antigen-induced arthritis (AIA) (Brackertz *et al.,* 1977), carrageenan-induced paw oedema (Levy, 1969), as well as transgenic spontaneous models of arthritis such as the K/BxN (Kouskoff *et al.,* 1996) and TNFα mouse models (Butler *et al.,* 1997). Although no animal model perfectly duplicates the pathogenic features of human RA, most have some pathological similarities to the human disease. However, animal models often have greater bone resorption and bone formation in response to joint inflammation, as well as a more rapid disease progression than observed in human RA (Bendele, 2001).

In addition to the selection of an animal model which has similar pathology to that of human disease, other important selection criteria include the predictability of the disease process, reproducibility of data, appropriate duration of disease and the relative ease of performing the model. The aforementioned rodent models of 'induced' arthritis and/or paw inflammation have been extensively used to develop a greater understanding of the pathogenesis of the disease process, to identify new therapeutic targets and to investigate the therapeutic efficacy of new drugs (Kannan *et al.,* 2005).

1.1.4.1. Collagen-induced arthritis (CIA)

The CIA mouse model is an extensively used model of RA which shares both immunological and pathological features of human RA. Arthritis is elicited in genetically susceptible strains of mice such as DBA/1, B10.Q, and B10.RIII, following immunisation with heterologous type II

collagen emulsified in complete Freund's adjuvant (intra-dermal injection) which results in the breakdown of immune tolerance and the induction of an autoimmune-mediated attack on the joints (Kollias *et al.,* 2011).

The immunopathogenesis of CIA is predominantly mediated by T-cell and B-cell specific responses to type II collagen, resulting in the production of collagen-specific T-cells and anti-collagen type II antibodies by activated B-cells. During the early stages of disease, collagen-specific antibodies bind to the collagen type II in the joint, resulting in immune-complex formation, complement activation, induction of a local inflammatory response, recruitment of monocytes and T-cells to the joint and the production of proinflammatory cytokines and mediators (Bevaart *et al.,* 2010).

Chronic polyarthritis typically occurs 21-28 days after immunisation and peaks at approximately day 35 and gradually declines over the following weeks. Mouse CIA is characterised by cartilage degradation associated with immune complex deposition on articular surfaces, synovial hyperplasia, mononuclear cell infiltration, bone resorption, periosteal bone proliferation and inflammation, similar to human RA. However, there is little or no sex bias in CIA and importantly, these mice develop antibodies directed against collagen, which is not a consistent feature in patients with RA (Brand *et al.,* 2007). Nevertheless, this model has been extremely useful in evaluating the therapeutic efficacy of biological agents such as soluble TNF receptor (Bevaart *et al.,* 2010).

The CIA mouse model is quite reproducible and a disease incidence of 80% can be achieved for most susceptible strains. However, the kinetics of arthritis development for each mouse in the same experimental group can vary widely (Brand *et al.,* 2007), which can impact on the timing/setup of experiments. Also, disease can occur in any combination of paws or joints, therefore, other mouse models with more predictable arthritis and/or inflammation, may be better suited for early evaluation studies of the therapeutic efficacy of drugs, on single and defined inflamed joints.

1.1.4.2. Antigen-induced arthritis (AIA)

AIA in most strains of mice, rats and rabbits can be elicited by immunisation with an antigen (e.g. methylated BSA in complete Freund's adjuvant, administered subcutaneously or intradermally) followed by a subsequent intraarticular challenge with the same antigen at least 14 days later, in one joint and a control joint can be treated with vehicle alone. Methylated BSA binds to the cartilage and induces an immune-complex-mediated inflammatory process, complement activation and T-cell-mediated responses on the articular cartilage surface, resulting in the onset of acute inflammation, secretion of pro-inflammatory mediators, immune cell infiltration and cartilage destruction, 3-4 weeks after AIA (Asquith *et al.,* 2009).

Mouse, rat and rabbit AIA models are commonly used to investigate the role of specific cytokines in various pathogenic processes the arthritis (e.g. van de Loo *et al.,* 1995). Also, the ability to repeatedly induce inflammation by re-challenging mice with methylated BSA offers the prospect of monitoring the therapeutic efficacy of biological agents e.g. IL-1 receptor antagonist, since IL-1 plays a major role in AIA (Di Domizio *et al.,* 2013).

1.1.4.3. Carrageenan-induced paw oedema

Carrageenans are sulphated polygalactans derived from a number of red seaweeds, of the class *Rhodophyceae*, and are of three major types; lambda (λ), kappa (κ) and iota (ɩ). The induction of paw inflammation by an intraplantar injection of λ-carrageenan in rat paws was first demonstrated by Winter *et al.,* (1962), where the inflammatory process was described as acute, non-immune and highly reproducible. A similar observation in mice was reported by Levy (1969) and later research by Henriques *et al.,* (1987) revealed that λ-carrageenan injected into the mouse paw elicits biphasic oedema where the first phase peaks at ~5 hours post-carrageenan injection and is characterised by low-intensity oedema while the second phase develops after 24 hours and peaks at 72 hours in a dose-dependent manner and decreases thereafter.

The injection of λ-carrageenan into mice paws provokes an acute local inflammatory reaction which is characterised by the production of pro-inflammatory mediators such as histamine, serotonin, bradykinin, complement and prostaglandins which results in an increase in vascular permeability and cellular infiltration leading to oedema formation (Necas and Bartosikova, 2013). The infiltration and activation of leukocytes, mainly neutrophils, contribute to the inflammatory reaction by producing reactive oxygen and nitrogen species and hydroxyl radicals, which can also be produced at the injured site (Posadas *et al.,* 2004).

In the literature, there are over 400 papers reporting the use of mouse paw oedema and the carrageenan-induced paw oedema model has been increasingly used to study different mechanisms of the inflammatory process and to study the therapeutic effects of anti-inflammatory drugs e.g. NSAIDs, where it has become a popular model for localised inflammation (Posadas *et al.,* 2004).

All animal models of human diseases have inherent limitations and the choice of the most suitable model is influenced by the selection criteria of the specific experiment. Overall, the research on animal models of RA and paw inflammation has positively contributed to the growing understanding of the disease mechanisms and has been instrumental in facilitating the development and preclinical evaluation of novel therapeutics for the treatment of the disease.

1.1.5. Treatment of RA

Although the aetiology of RA is unclear, the pathogenesis of RA is better understood and therefore, has facilitated the innovative developments of effective RA therapies. The general aims of drug management in RA are to alleviate the symptoms of the disease, and to modify the pathological process to prevent or stop further disease progression in order to ultimately achieve RA remission i.e. having scores below certain levels on disease activity indices (e.g. Disease Activity Score in 28 joints (DAS28) <2.6) (Rheumatoid Arthritis; National Clinical Guideline for Management and Treatment in Adults, UK, 2009). The current therapies used to treat RA can be divided into three groups; non-steroidal anti-inflammatory drugs (NSAIDs), disease modifying anti-rheumatic drugs (DMARDs) and biological therapies.

1.1.5.1. NSAIDs

NSAIDs are effective anti-inflammatory and analgesic drugs, which inhibit the biosynthesis of prostaglandins to provide symptomatic relief. NSAIDs exert their actions by binding to the active site and inhibiting cyclooxygenase (COX) enzymes, which are key enzymes responsible for the formation of eicosanoids from arachidonic acid. Eicosanoids can be subdivided into prostanoids, comprising prostaglandins, prostacyclin and thromboxane A2, all of which play important roles in inflammation (Crofford, 2013).

COX exists in two isoforms; the constitutively expressed COX-1, which is present in most tissues including vascular endothelium, stomach mucosa and kidneys, and the inducible COX-2, which is induced in a number of cells by pro-inflammatory stimuli and is therefore the main target for the treatment of inflammatory diseases (Mitchell and Warner, 1999). The prostanoids produced by COX-1 are important in gastric protection, therefore, the broad inhibition of COX enzymes have detrimental effects on gastric function e.g. non-selective NSAIDs (COX-1 and COX-2 inhibitors) such as indomethacin, naproxen and diclofenac inhibit prostaglandin production which increases gastric motility and provokes severe gastric lesions. This

consequently led to the development of selective COX-2 inhibitors (COXIBs) which significantly reduced gastric side effects (Jackson *et al.,* 2000). However, COXIBs have been linked to increased risk of cardiovascular events and increased reports of deaths associated with COXIBs, particularly rofecoxib (Moodley, 2008). Also, COXIBs decrease the release of prostacyclins, which are inhibitors of platelet aggregation and are potent vasodilators, therefore, COXIBs may also promote increased prothrombotic activity and cardiovascular complications (Mukherjee *et al.,* 2001).

1.1.5.2. DMARDs

Conventional DMARDs, including methotrexate (MTX), sulphasalazine, hydroxychloroquine, leflunomide and gold injections, act to modify the progression of RA by reducing pain, inflammation and joint damage (Upchurch and Kay, 2012). MTX is the most frequent drug of choice for the early stage of RA and exerts anti-proliferative effects by competitively inhibiting folate-dependent enzymes such as dihydrofolate reductase to prevent *de novo* purine and pyrimidine synthesis, which is essential for DNA and RNA synthesis. Also, the anti-inflammatory effects of MTX are due to increased extracellular levels of 5-aminoimidazole-4-carboxamide ribonucleotide, that leads to inhibition of adenosine and AMP-deaminase and accumulation of extracellular adenosine and AMP that bind to the adenosine-A2 receptor to increase the production of anti-inflammatory IL-10 cytokine and inhibit NFκB activity. Consequently, treatment with MTX can result in decreased cell proliferation, increased cell apoptosis of immune cells and decreased pro-inflammatory cytokine production (Meier *et al.,* 2013). Disease activity is evaluated periodically and it has been observed that the effectiveness of DMARDs often decreases as the disease progresses or when patients experience adverse effects and therefore, DMARD therapy often fails to stop the progressive destruction of articular cartilage and bone. The decision to initiate biological drug treatment in these DMARD-inadequate responders is recommended to patients who have persistently elevated DAS28 >5.1 indicating severe RA and who have failed to respond to two DMARDS (including MTX, unless contraindicated) taken over a minimum of 6 months each. In such

cases, biological TNFα antagonists i.e. infliximab, are the first 'biologic' option (Kiely *et al.,* 2012).

1.1.5.3. Biological drugs

The substantial evidence of the integral role of pro-inflammatory cytokines and immune cells in RA pathology has led to the development of targeted biological drugs which modify the disease process by inhibiting specific pro-inflammatory cytokines or inflammatory cells implicated in inflammatory pathways (Table 1.1). These biological proteins have revolutionised the treatment of RA and improved the quality of life of millions of RA sufferers. However, these drugs are not without adverse effects which therefore limit their long-term use.

Agent	Structure	Method of action	Dosage and administration	Reported side effects
Etanercept	Fusion protein consisting of the extracellular domain of the p75-TNF receptor fused with human IgG_1	Binds to circulating TNF and lymphotoxin and membrane bound TNF, thereby blocking the interaction of TNF with TNF-receptors. May have effects that are mediated through neuro-endocrine pathways as wells	25 mg twice weekly for 3 months or 50mg once weekly, subcutaneously	-Injection site reactions (pain, swelling, erythema)

-Upper respiratory tract infections (reactivation of latent diseases)

-Oncogenic potential (particularly lymphoma)

-Infusion-related reactions (fever, chills, chest pain, shortness of breath)

-Serious infections e.g. pneumonia, cellulitis |
| Infliximab | Chimeric (human/mouse) humanised monoclonal antibody against $TNF\alpha$ | Binds to soluble and membrane-bound forms of $TNF\alpha$, thereby blocking the interaction of $TNF\alpha$ with TNF-receptors. Infliximab causes apoptosis of cells with cell surface TNF | 3-10 mg/kg over the course of 2-3 hours, intravenously. Administered every 4-8 weeks. Initial dosing may be at 0, 2, and 6 weeks | |
| Adalimumab | 'Fully' human recombinant IgG_1 monoclonal antibody with specificity for $TNF\alpha$ | Binds to soluble and membrane bound $TNF\alpha$ leading to a blockade of activity of $TNF\alpha$ causing apoptosis of cells with membrane-bound TNF | 40 mg every other week, subcutaneously | |
| Abatacept | Fusion protein consisting of the extracellular protein of human CTLA4 with a fragment of the Fc portion of human IgG_1 | Binds to CD80/86 on the cell surface of antigen-presenting cells to prevent co-stimulatory binding of CD28 on the surface of T-cells, thereby inhibiting T-cell activation | The dose of abatacept depends on body weight (500mg-1000mg range). Intravenous infusions at week 0, 2, 4 and then every 4 weeks thereafter | -Infections including sepsis and pneumonia
-Contraindicated in patients with severe, uncontrolled and/or opportunistic infections |
| Rituximab | Chimeric monoclonal antibody against CD20 antigen | Binds to CD20 on the cell surface of mature B cells, causing apoptosis. Targets and selectively depletes CD20+ B-cells without targeting stem cells or existing plasma cells. Cell lysis via either complement-dependent cytotoxicity and/or antibody-dependent cell-mediated cytotoxicity | Two 1000 mg intravenous infusions, given 2 weeks apart | -Contraindicated in patients allergic to the drug or its components, hepatitis B carriers, patients with cardiac arrhythmia, angina pectoris or active infections |

Table 1.1. Biological Agents: Their structure, method of action, doses, administration and side effects. Information in table modified from Callen, (2007) and National Institute for Health and Care Excellence (NICE) Technology Appraisal Guidance

Although >30% of patients receive biologics as a monotherapy, the combination of biologics with DMARDs such as MTX (unless contraindicated) dramatically increases the efficiency of the treatment. Generally, the targeted suppression of key inflammatory pathways involved in joint inflammation and destruction using biologics reduces the signs and symptoms of RA, inhibits the progression of structural damage and improves the physical function in patients with moderate to severe RA (Doan *et al.*, 2006). However, the treatment with biologics is based on continuous immunosuppresion irrespective of the patient's inflammatory status and their continued use raises many concerns and challenges, particularly the serious side effects such as the increase in the risk of infections and reactivation of latent diseases. As well as being very expensive, many patients develop inadequate responses to biological treatments (Kiely *et al.*, 2012). Also, the relatively short half-life of these proteins (few days) require frequent systemic injections at high doses in order to maintain efficacious levels of the therapeutic protein in the joint. Consequently, non-target organs are exposed to high concentrations of anti-inflammatory agents which can increase the risk of immunosuppression.

Due to the localised nature of the RA joint, the intra-articular delivery of therapeutic agents provides a safer and more cost-effective alternative to systemically administered drugs. However, the intra-articular delivery of therapeutic proteins is a challenge due to the rapid clearance of small molecules (via the synovial capillaries) and macromolecules (via the lymphatic system) and repeated intra-articular injections is not a feasible option for the treatment of RA (Evans *et al.*, 2013). An attractive alternative to protein therapy is to genetically modify the cells to synthesise and express the encoded therapeutic proteins within the joint, using gene therapy.

1.2. Gene therapy

Somatic gene therapy can be broadly defined as an experimental technique involving the transfer of genetic material into living cells (excluding germ cells) to cure, treat or prevent diseases (Chernajovsky *et al.,* 2004). The success of gene therapy approaches is primarily dependent on efficient gene transfer and expression of therapeutic genes, which has been achieved in various experimental systems and has brought gene therapy to the forefront of molecular medicine. Safer methods have been developed and extensively optimised to provide efficient gene delivery to RA joints, using *in vivo* and *ex vivo* strategies in conjunction with a variety of viral and non-viral vectors.

1.2.1. Mode of gene transfer

Systemic and local delivery of therapeutic genes can be achieved by either direct *in vivo* or indirect *ex vivo* gene delivery. The general concept of the *ex vivo* strategy involves the genetic modification of cells *in vitro* with the therapeutic DNA, followed by transplantation of the genetically modified cells either locally or systemically. *Ex vivo* gene delivery offers the advantage of characterising and then selecting and expanding the desired transduced cells for delivery. However, this strategy is limited by the accessibility and isolation of the cells and their ability to survive in culture for long periods of time without significantly changing their phenotype as well as the complexity and high cost of the approach (Adriaansen *et al.,* 2006).

In contrast, the *in vivo* strategy involves the direct delivery of genes, usually by viral vectors containing the therapeutic gene at the site where therapeutic protein expression is needed and involves genetic modification of the target cells (Robbins *et al.,* 2003; Chernajovsky *et al.,* 2004). In the context of *in vivo* gene delivery directly into RA joints, the transcriptional and translational machinery of the transduced cells can produce and secrete the therapeutic proteins into the joint space to provide sustained intra-articular expression of therapeutic

proteins and potential long-term amelioration of chronic joint disease after a single or a limited number of gene delivery procedures (Gouze *et al.,* 2007).

In vivo gene therapy has been studied with varying success in animal models of RA mainly using recombinant adenoviral, recombinant adeno-associated virus, retroviral and lentiviral vectors. One of the most difficult challenges in gene therapy is the development of vectors and delivery approaches that are safe, efficient and targeted. Substantial progress has been made in the application of gene therapy by acknowledging the fundamental differences between gene delivery vectors i.e. vector capacity, duration of transgene expression, the ability to target dividing or non-dividing cells, extrachromosomal or genomic integration, immunogenicity and safety (Gould and Favorov, 2003). Both viral and non-viral vectors have been explored as a means of transporting therapeutic genes into target cells.

1.2.2. Non-viral vectors

Non-viral vectors include any method of gene transfer that does not involve the production of a viral particle and can be divided into two classes: (1) RNA or DNA amplified in bacteria or eukaryotic cells, which do not require a viral particle for transfer into cells e.g. plasmid DNA or (2) oligonucleotides or related chemically synthesised molecules (Ponder, 2001). Plasmid DNA is commonly used for gene expression studies and the main advantages of plasmid DNA vectors are their large packaging capacities, cheap and simple production, their lack of accessory proteins and their generally non-immunogenic profile, which has been further improved by the removal of immunogenic unmethylated CpG motifs (Krieg *et al.,* 1995). Plasmid DNA lacks the inherent ability to enter cells and localise to the nucleus and therefore relies on physical or chemical methods to enter cells. Such delivery methods include electroporation, direct injection into tissue or the circulation and DNA complexing reagents such as cationic lipids which condense DNA thereby protecting it from degradation and facilitating DNA uptake by endocytosis (Gould and Favorov, 2003; Gould and Chernajovsky,

2007). Following cell entry, the plasmid DNA enters the nucleus and exists as extrachromosomal DNA which is transcribed and translated using the cell machinery. However, because the DNA can be degraded or is lost during mitosis, gene expression is transient.

There has been limited success with direct intra-articular injection of plasmid DNA encoding therapeutic genes for example, Bloquel *et al.,* (2007) who demonstrated decreased joint destruction in the ankles of mice with collagen-induced arthritis (CIA), following the intra-articular delivery with electroporation of TNF-receptor encoding plasmid DNA (Bloquel *et al.,* 2007). With the exception of skeletal muscle where gene expression can last for months to years (Wolff *et al.,* 1992), intra-articular plasmid DNA expression is generally low with poor efficacy and longevity which limits their use for local RA gene therapy.

1.2.3. Viral vectors

Viral vectors are derived from viruses with either RNA or DNA genomes and are either integrating or non-integrating vectors (Verma and Weitzman, 2005). The viral vectors employed for experimental gene delivery to inflamed joints are adenoviruses, adeno-associated viruses (AAVs), retroviruses, herpes simplex viruses (HSV) and lentiviruses. The genomes of these viruses have been significantly altered to create safer vehicles for gene delivery (Kay *et al.,* 2001) and have different advantages and disadvantages, as summarised in Table 1.2.

Vector	Genetic material	Immunogenicity	Vector yield (TU/ml)	Transgene capacity	Vector genome forms	Transgene expression	Transduction of quiescent cells	Genotoxicity
Retrovirus	RNA	Low	Moderate (1×10^{10})	Upto 8 kb	Integrated	Long term (years)	No	Integration might induce oncogenesis
Lentivirus	RNA	Low	Moderate (1×10^{10})	Upto 8 kb	Integrated	Long term (years)	Yes	Integration might induce oncogenesis
Herpes simplex virus (HSV)	dsDNA	High	High (1×10^{12})	~ 35 kb	Episomal	Transient	Yes	Low
Adenovirus (gutless)	dsDNA	High (but less than 1st and 2nd generation Ad	High (1×10^{12})	~ 35 kb	Episomal	Short term (weeks)	Yes	Low
Adeno-associated virus (AAV)	ssDNA	Low	High (1×10^{12})	~ 4 kb	Integrated (10%) Episomal (90%)	Medium to long term (year)	Yes	Needs more investigation

Table 1.2. The properties of viral vectors. Table adapted from Gould and Favorov (2003) and Bouard *et al.,* (2009).

Although adenoviruses are commonly used in experimental gene transfer, these vectors are highly immunogenic and exhibit short-term transgene expression, which undermines the efficacy of the therapy. Despite their low transgene capacity, AAVs have been identified as the vector of choice for human gene therapy clinical trials (Evans *et al.,* 2013). In contrast, the main advantages of retroviral and lentiviral vectors, over the other vectors, are their ability to permanently integrate into the host genome, thus conferring stable transmission and long-term gene expression in the target cells and subsequent progeny cells. Additionally, both vectors have relatively large packaging capacities that can accommodate upto 8 kb transgene cassettes and have reduced *in vivo* immunogenicity compared to other gene therapy viral vectors i.e. adenoviruses, which minimises the risk of inducing a host immune reaction against the viral vector and/or transduced cells (Buchschacher and Wong-Staal, 2000). However, retroviral vectors can only transduce dividing cells which limits their application for gene therapy to *ex vivo* gene transfer. In contrast, lentiviral vectors have low immunogenicity and have the unique ability to transduce both dividing and non-dividing cells making them attractive candidates as delivery vectors for *in vivo* gene therapy (Kay *et al.,* 2001; Sinn *et al.,* 2005),

1.2.3.1. Lentivirus life cycle

Early research into Acquired Immune Deficiency Syndrome (AIDS) significantly improved the understanding of the structure and biology of the causative Human Immunodeficiency Virus-1 (HIV-1), which also laid the foundation for the development of HIV-based lentiviral vectors for gene therapy. HIV-1 belongs to the genus *lentivirus*, which is part of the *Retroviridae* family. Both retroviruses and lentiviruses are RNA-based viruses that replicate through a DNA intermediate, however, lentiviruses have a more complex genome and mechanisms that control their stages of infection (Buchschacher and Wong-Staal, 2000). Figure 1.3 schematically depicts the structure and genome of the HIV-1 lentivirus.

[A

Env proteins
gp120
gp41

Gag proteins
p17 matrix
p24 capsid

Pro-pol proteins
Integrase
Protease
Reverse transcriptase
RNA
Lipid bilayer
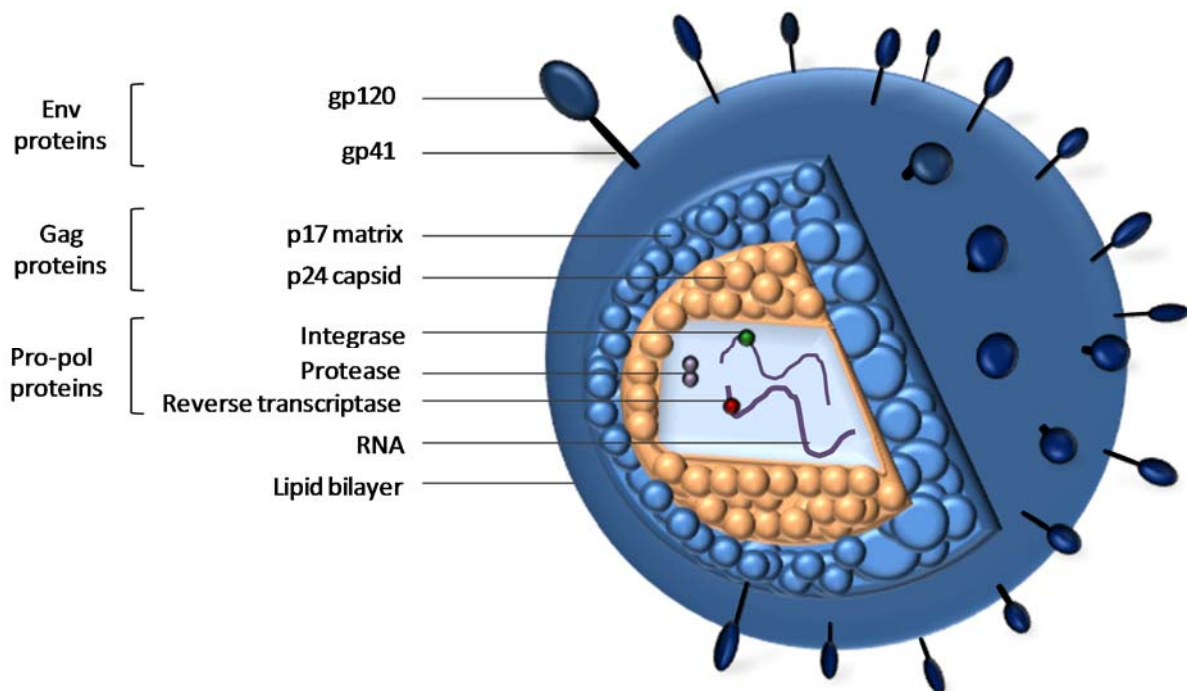
[B
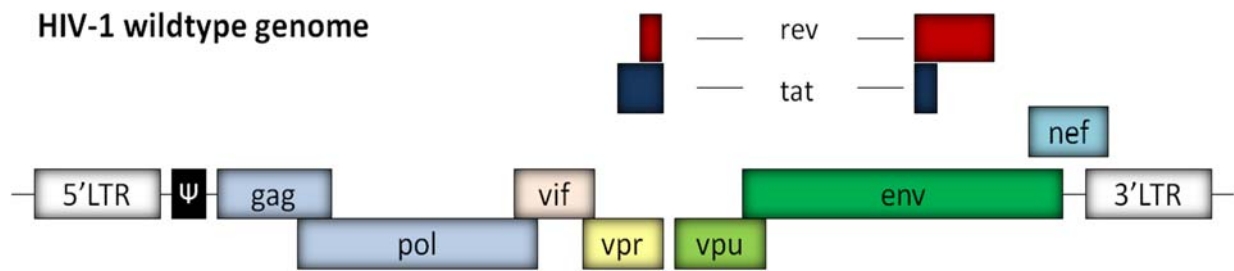
Figure 1.3. Schematic diagram of a HIV-1 mature virion and genome. The HIV-1 mature virion contains the RNA genome, envelope glycoproteins gp120 and gp41, matrix protein p17, capsid protein p24, as well as the integrase, protease and reverse transcriptase enzymes [A] (HIV-1 virion figure was redrawn, with modifications, from the image by Kumar *et al.* (2007) in Robbins Basic Pathology, 8th edition). The wildtype HIV-1 genome consists of the 5'- and 3'-long terminal repeats (LTR), RNA packaging sequence (Ψ), the gag, pol, and env genes which encode the capsid, polymerase and envelope proteins, respectively and the genes encoding the transactivator of transcription (tat), regulator of virion protein expression (rev), negative regulatory factor (nef), virion infectivity factor (vif), viral protein U (vpu) and viral protein R (vpr) [B].

The HIV-1 life cycle can be divided into two temporally distinct phases: (1) infection, which results in the entry of the viral genome into the cell and (2) replication, which includes the early and late phases of gene expression when viral regulatory products and structural genes are expressed and assembled into viral particles (Kay *et al.,* 2001).

1. Virion binding and entry

The initial step in the HIV-1 replication cycle (early phase of infection) is the binding of gp120 to the CD4 molecule on the surface of host $CD4^+$ T-cells and macrophages. This interaction, which is facilitated by the viral entry co-receptors CCR5 or CXCR4 results in a conformational change in the transmembrane gp41 protein, exposing the hydrophobic fusion domain at the N-terminus of gp-41, allowing the fusion between the viral and cell membranes and internalisation of the capsid (Melikyan, 2008).

2. Reverse transcription

In the cytoplasm, the virions undergo structural changes to form the reverse transcription complex (RTC), which is the site of viral DNA synthesis. Reverse transcription is initiated with the binding of the specific cellular tRNA ($tRNA_{Lys3}$) primer to the primer binding site (PBS), followed by extension of the primer until the end of the genomic RNA (gRNA) molecule at the 5'-end. The RNA portion within the resulting RNA/DNA duplex is degraded by the RNase H subunit of the reverse transcriptase (RT) enzyme to form the minus-strand DNA, which is subsequently transferred to the 3'-end of the gRNA in the complementary R region during the process of first 'minus-strand DNA' transfer. Minus-strand DNA synthesis is continued and the viral RNA is degraded by RNase H, except for the central polypurine tract (cPPT) and the 3'-polypurine tract (PPT), which serve as primers for the synthesis of the plus-strand DNA. The plus-strand DNA is extended from the PPT primer until the 3'-LTR and the PBS region which

is followed by the second (plus-strand) DNA transfer. Synthesis of the transferred plus-strand DNA continues upto the cPPT, after which the process is terminated, thereby creating a 99 bp central DNA flap between the cPPT and the central termination sequence (CTS), which is thought to facilitate nuclear entry. DNA synthesis of plus- and minus-DNA strands is continued to the ends of both templates until a double stranded viral DNA is formed. The resulting viral DNA subsequently binds to viral and cellular proteins, such as RT, IN, NC, Vpr and MA, to form the preintegration complex (PIC) which is actively transported into the nucleus via the nuclear pore complex (Pluta and Kacprzak, 2009; Sarafianos *et al.,* 2009; Hu and Hughes, 2012).

3. Integration

The integration of viral DNA into the host cell genome is catalysed by the lentiviral integrase (IN) enzyme which binds to the short terminal fragments called att in the U5 and U3 LTRs resulting in the bending of the DNA molecule. The first reaction, known as 3'-processing occurs in the cytoplasm within the PIC and involves the removal of two nucleotides from each 3'- end of the viral DNA which exposures the OH group. The second reaction called DNA strand transfer takes place in the nucleus at the site of integration where the IN enzyme uses the 3'-OH groups to cleave the phosphodiester backbone of the host chromosomal DNA resulting in the joining of the viral 3'-ends to the 5'-target DNA. In the final reaction called gap repair, the extra nucleotides from the 5'-ends of the viral cDNA are removed and joined to the host DNA 3'-ends with the aid of host cell DNA repair enzymes (Poeschla, 2008).

4. Gene Expression

Following integration, the provirus exploits the cellular transcriptional and translational machinery. During transcription, the LTR acts as the viral promoter where RNA polymerase II binds to the first nucleotide of the R region in the 5'-LTR, with low efficiency. The viral Tat transactivator protein binds to the transactivation-responsive region (TAR) in the 5'-end of the nascent mRNA transcript, which subsequently recruits the host elongation factor composed of cyclin CycT1 and kinase CDK9, to enhance the processitivity of RNA polymerase II. The p300/CREB-binding protein-associated factor (P-CAF) mediates the acetylation of Tat which liberates Tat from TAR. The free acetylated Tat protein recruits P-CAF to the phosphorylated RNA polymerase II to further stimulate transcript elongation (Karn and Stoltzfus, 2009; Pluta and Kacprzak, 2009).

The encoding sequences of the Gag and Pol proteins are present in different reading frames and are translated from a full-length unspliced RNA in the cytoplasm. The translation of the Gag-Pro-Pol polyprotein requires a single -1 frameshift whereas the Env gene is transcribed as a full length mRNA which encodes the viral envelope glycoproteins gp120 and gp41 and the resulting proteins are assembled at the plasma membrane (Bolinger and Boris-Lawrie, 2009).

5. Assembly, budding and maturation

The assembly process packages all of the components required for infectivity into the virion, which include two copies of the (+) sense genomic viral RNA, molecules for cDNA synthesis e.g. tRNALys, the viral Env protein, Gag polyprotein and the three viral enzymes, protease (PR), reverse transcriptase (RT) and integrase (IN). At the plasma membrane, the amino terminal Gag domain called MA, binds to the plasma membrane and recruits the viral Env glycoproteins. The protein-protein interactions required for immature virion assembly are mediated by the central domain of Gag called CA, which creates the capsid of the mature viral

core. The NC domain of Gag binds to the psi-packaging (Ψ) signal in the 5'-LTR of the viral genomic RNA during virus assembly. Following the trafficking of the Gag-RNA, Gag-Pro-Pol and Gag polyproteins and Env glycoproteins to the plasma membrane, the HIV-1 virus buds at the plasma membrane of the infected cells and is released as an immature virion. The virion undergoes maturation, mediated by the viral PR enzyme, which cleaves the Gag and Gag-Pro-Pol polyproteins at ten different sites to produce the structural proteins MA, CA, NC, p6 and the enzymatic proteins PR, RT and IN, resulting in a mature and fully infectious HIV-1 virion (Sundquist and Krausslich, 2012). The HIV-1 life cycle is schematically depicted in Figure 1.4.
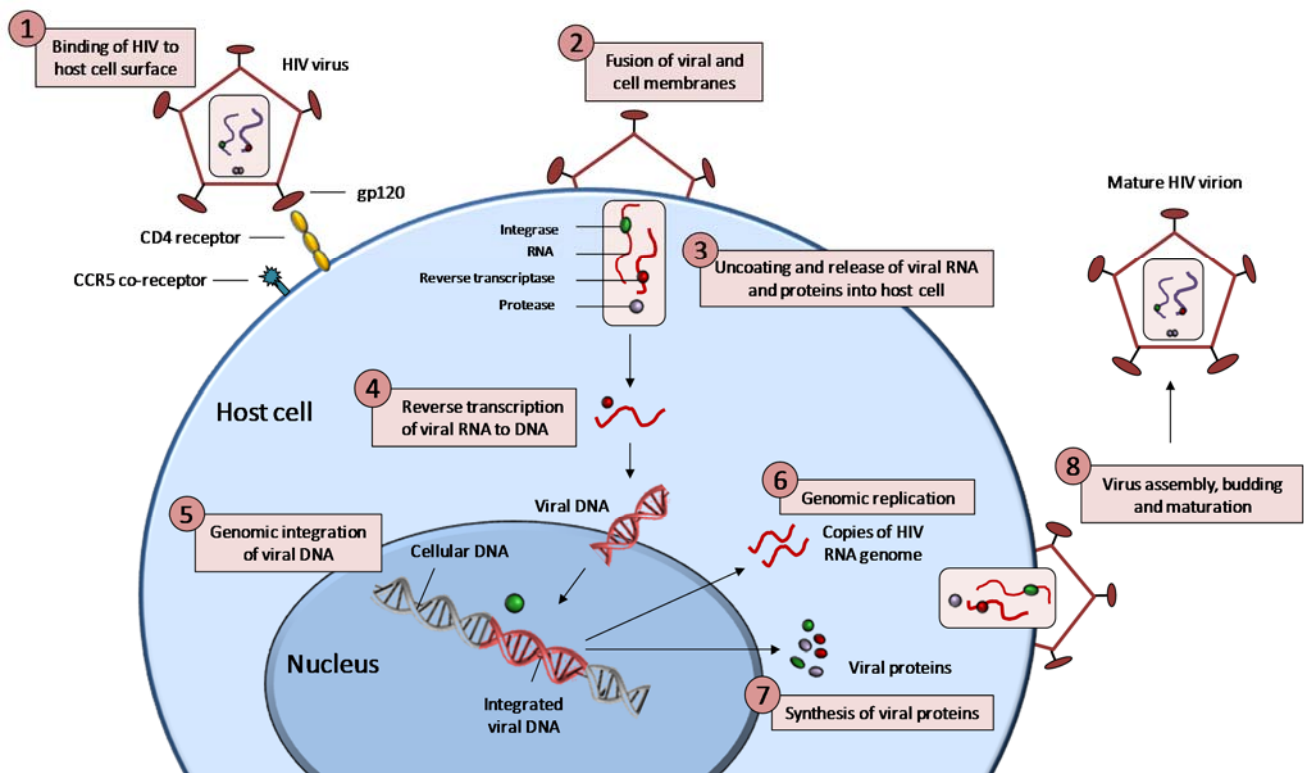
Figure 1.4. HIV-1 life cycle. The HIV-1 virus binds to the host cell surface receptor [1] allowing fusion between the viral and cell membranes [2]. The viral RNA and proteins are released into the host cell cytoplasm [3] and the viral RNA is reverse transcribed to form the viral DNA [4]. The viral DNA is integrated into the host cellular DNA in the nucleus [5]. Two copies of the HIV RNA genome is formed following genomic replication [6] and the viral proteins are synthesised [7]. The virus components are assembled and

1.2.4. Lentiviral vectors

Since the first demonstration of the ability of lentiviral vectors to transduce neurons *in vivo* (Naldini, *et al.,* 1996a; 1996b), lentiviral vectors have been extensively modified to achieve high levels of efficiency and biosafety. These vectors are engineered to be replication-defective, to permit infection and integration of the transgene into the target cells whilst preventing multiplication and spread of the virus to other cells and the emergence of replication-competent lentiviruses (Durand and Cimarelli, 2011). Early improvements in the development of lentiviral vectors were based on the concept of separating the cis-acting elements necessary for vector RNA synthesis, packaging, reverse transcription and cDNA integration from the *trans*-elements that encode viral enzymes, structural and accessory proteins, on two DNA strands that can be delivered in different plasmids. Vector-production systems typically consist of a packaging expression cassette (helper vector), a vector expression cassette containing the heterologous promoter driving expression of a transgene which is flanked by LTRs and *cis*-elements (transfer vector) and an envelope expression cassette (Pluta and Kacprzak, 2009).

Significant efforts have been made to minimise the risk of replication-competent lentiviruses by modifying the transfer vector or the packaging vector to reduce areas of overlap between the two to prevent the reconstitution of replication-competent lentiviruses. The first-generation lentiviral vectors were created by a three plasmid transfection into a packaging cell line, where the packaging vector comprised all HIV genes except for the Env envelope gene, which was encoded in the envelope vector (Naldini *et al.,* 1996a). This system was further improved in the second-generation lentiviral vectors where most of the accessory genes (*vif, vpr, vpu* and *nef*) encoding proteins that are likely to be virulence factors, were eliminated. However, the *gag*, *pol*, *tat* and *rev* were retained to enable transcriptional and posttranscriptional functions (Zufferey *et al.,* 1997). Safety tests have confirmed the absence of replication-competent lentiviruses with this system and due to their good safety record and high efficiency, second-

generation systems are the most frequently used systems for experimental purposes. Conversely, the packaging vector used to create third-generation lentiviral vectors have the *tat* gene eliminated and the *rev* gene is encoded by a fourth plasmid. Therefore, only three out of the nine genes are present in the genome; the *gag*, *pol* and *rev* genes, which further eliminates the possibility of reconstituting a wild-type virus through recombination, making this system more suitable for human gene therapy compared to the others (Dull *et al.,* 1998).

Following co-transfection of the plasmids into a packaging cell line i.e. Human Embryonic Kidney (HEK) 293T cells, the cells express the components of the vector and release the infectious lentiviral particles into the cell supernatant, which only permit a single round of infection. The lentivirus particles are then used to transduce the target cells, however, the tropism of the lentivirus is determined by the envelope protein. The vast majority of lentiviral vectors are pseudotyped with the Vesicular Stomatitis Virus glycoprotein (VSV-G) envelope and can target a broad range of cells due to the interaction of the VSV-G envelope with phospholipid components of the cell membrane of the majority of cell types. The VSV-G also provides structural stability to the lentivirus during ultracentrifugation, allowing concentration of the virus for further applications (Burns *et al.,* 1993).

Several modifications to the transfer vectors have contributed to increasing their biosafety and their performances of gene transfer. For example, a 400-nucleotide deletion in the promoter sequence of the 3'LTR generated 'self-inactivating' (SIN) vectors with abolished LTR promoter activity. During reverse transcription, the defective promoter in the 3'LTR is transferred to the 5'LTR of the proviral DNA and inactivates proviral transcription, thereby minimising the risk of generating replication-competent lentiviruses. Also, following integration of proviral DNA, the promoter activity of the 3'LTR can activate expression of adjacent host genes through an enhancer effect, however SIN vectors lack the 3'LTR promoter activity which decreases the likelihood of such events occurring and improves the performance of the vector (Zufferey *et al.,* 1998). Therefore, SIN vectors do not typically demonstrate insertional mutagenesis (*in*

vitro), which is frequently observed in lentiviral and gamma-retroviral vectors (Bokhoven *et al.,* 2009). Consequently, SIN transfer vectors consist of modified LTRs, a transgene driven by a suitable promoter of choice and the Ψ-packaging sequence. Further modifications to the transfer vector include the incorporation of insulator elements such as the chromatin insulator from the 5'-end of the chicken β-globin locus (5'-cHS4), positioned to flank the heterologous promoter and transgene, function to shield the promoter from the action of distal promoters in the host cell genome (enhancer blocking effects) and protect the transgene against gene silencing by preventing the advancement of adjacent inactive condensed chromatin (barrier effects) (Pikaart *et al.,* 1998). Also, the inclusion of the cPPT and the central termination sequence (CTS), has been shown to increase lentiviral vector transduction efficiency (Sirven *et al.,* 2000) and the incorporation of the woodchuck posttranscriptional regulatory element (WPRE) immediately downstream of the stop codon increases the viral titre and improves transgene expression by increasing mRNA stability (Zufferey *et al.,* 1999).
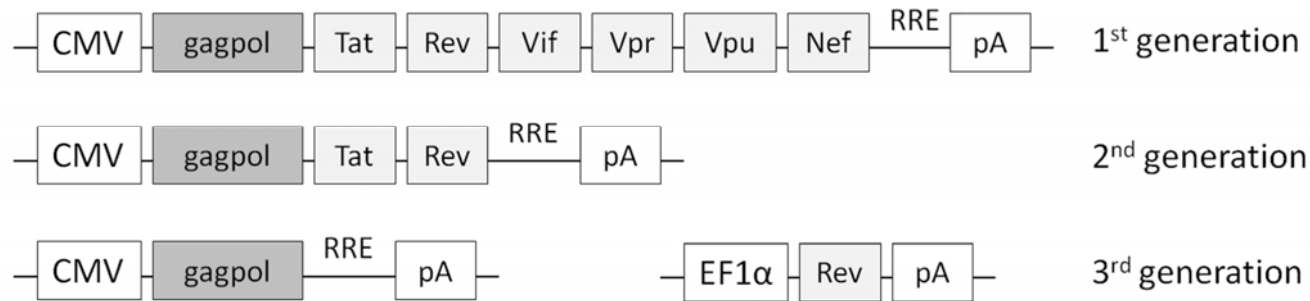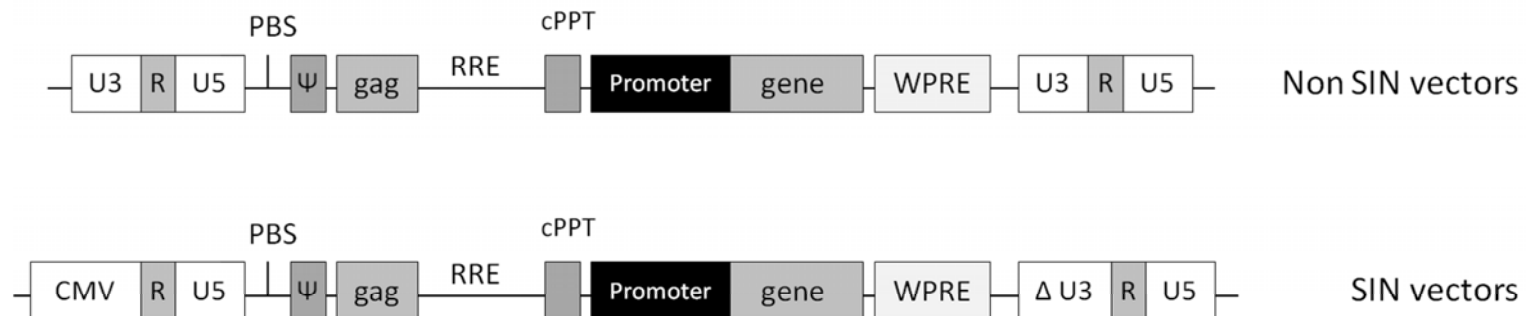
Figure 1.5 Lentiviral vector transfer, packaging and envelope constructs.

(A) First, second and third generation packaging constructs. (B) Non self-inactivating (SIN) and SIN lentiviral transfer vectors.

(C) Envelope Construct, Vesicular Stomatitis Virus Glycoprotein (VSV-G) gene driven by the CMV promoter.

1.2.5. Targets for RA Gene therapy

Logically, the success of both protein and gene therapy for the treatment of RA is dependent on the type of therapeutic agent used to target the activity of the inflammatory/destructive molecule. To date, anti-cytokine therapies e.g. anti-TNF and IL-1Ra, have been the most successful strategies to ameliorate inflammation and disease severity in both experimental arthritis and in humans. Notably, AAV-mediated delivery of genes encoding TNFR-Fc (etanercept) and IL-1Ra has entered clinical trials and these studies have reported variable success of the treatments (Evans *et al.,* 2013).

In addition to targeting pro-inflammatory cytokines, experimental (gene therapy) success has also been achieved through the use of vectors encoding therapeutic agents designed to inhibit gene expression of inflammatory/destructive molecules or overexpress natural therapeutic molecules. For example, Th2 cytokines such as IL-4, IL-10 and IL-13 are immunosuppressive and are known to inhibit the activities of IL-1β and TNFα. Overexpression of these cytokines by gene delivery has demonstrated beneficial effects on arthritis, particularly IL-4, which exerts bone and chondrocyte protective actions (van de Loo and van den Berg, 2002). Also, the inhibition of T-cell activation through the blockade of the CD28/CD80-86 co-stimulation pathway using adenoviral expressed CTLA-4-IgG$_1$ fusion protein, efficiently suppressed established CIA (Quattrocchi *et al.,* 2000).

The invasion of articular cartilage and bone by the synovial pannus is a hallmark feature of joint destruction in RA, therefore the induction of synovial cell apoptosis offers another target for RA therapy. Synoviocyte depletion has been achieved by *ex vivo* gene transfer of Fas ligand, which upon recognition of Fas antigen, induces apoptosis (Okamoto *et al.,*1998) and also *in vivo* adenoviral gene transfer of the Fas-associated death domain (FADD) which binds to the intracellular death domain of Fas to promote Fas-mediated apoptosis (Kobayashi *et al.,* 2000). An alternative approach to induce cell death is through the delivery of the HSV thymidine kinase (TK)-mediated suicide gene, as demonstrated by Goossens and colleagues

(1999), who reported a reduction in joint swelling and ablation of the synovial lining layer in the joints of rhesus monkeys with CIA, following the i.a. delivery of Ad-HSK-TK delivery and parental treatment with the prodrug, ganciclovir (Goossens *et al.,* 1999).

Other targets for RA therapy include the prevention of cartilage destruction, for example the retroviral transduction of MMP-1 specific ribozymes inhibited MMP-1 production in RASF and reduced the invasiveness of the cells in a SCID mouse model of RA (Rutkauskaite *et al.,* 2004). Conversely, a cartilage regeneration approach has been demonstrated using adenoviral delivered insulin-like growth factor, which led to stimulation of new proteoglycan synthesis in cartilage in rabbit joints with antigen-induced arthritis (AIA) (Mi *et al.,* 2000).

NFκB plays a major role in the pathogenesis of RA and is highly expressed in RA tissues. Several studies have shown that the delivery of NFκB inhibitors decreases arthritis and inflammatory cytokine production in animal models of RA, for example, the local adenoviral–mediated gene transfer of dominant-negative IKKβ exerted anti-arthritic effects by significantly decreasing NFκB DNA expression in the joints of rats with AIA (Tak *et al.,* 2001). However, a major disadvantage of targeting ubiquitously expressed transcription factors is that the systemic dissemination of transcription factor inhibitors can potentially result in detrimental side effects, due to the vital role of NFκB in a wide range of physiological processes.

Overall, a common feature of the gene therapy systems described above and numerous others, is the use of constitutive promoters e.g. CMV. The disadvantage of uncontrolled expression of therapeutic genes by these promoters is the potential side effects associated with continuous therapeutic gene expression, especially in non-target tissues. The incorporation of regulatable gene expression systems are an attractive alternative to constitutive gene expression, particularly for RA therapy, as they can improve the safety and efficacy of the therapy.

1.2.6. Regulated gene expression

Due to the relapsing nature of RA, therapeutic gene expression could be engineered to follow the intermittent course of disease within the joints so that therapeutic molecules are maximally expressed during disease flare and marginally expressed during disease remission. This regulated approach can potentially enhance the therapeutic effects, reduce side effects and lower the amount of therapeutic vector required (Adriaansen *et al.,* 2006). Many innovative gene regulation strategies have been developed to circumvent unwanted therapeutic gene expression and the approaches relevant to RA gene therapy typically employ the use of 'transcriptionally targeted' vectors. Transcriptionally-regulated therapeutic gene expression for RA can be achieved by constructing pharmacological or physiological-regulated synthetic promoters, which are typically comprised of the binding site of an inducible transcriptional factor coupled to a core promoter motif e.g. TATA box (Khoury *et al.,* 2007). Therefore, an understanding of the role of core promoters and transcription factors in eukaryotic transcription is particularly relevant to the rationale design and construction of synthetic promoters.

1.2.6.1. Eukaryotic transcription

The expression of protein-coding genes is primarily mediated by RNA polymerase II (RNA pol II) and a complex network of general transcription factors; TFIID, TFIIB, TFIIF, TFIIE and TFIIH. The core promoter fundamentally contributes to the transcriptional process by serving as a platform for the assembly of the transcriptional preinitiation complex (PIC), comprised of RNA pol II and its basal transcription factors, which functions to cooperatively specify the transcriptional start site (TSS) to initiate transcription (Butler and Kadonaga, 2002).

Many polymerase II promoters have a TATA box (consensus sequence; TATAA/TA) which is a core promoter motif located -25 bp upstream of the transcriptional start site. During transcription initiation, this sequence is recognised by the TATA-binding protein (TBP) subunit of the general transcription factor TFIID, which also consists of multiple TBP-associated

factors (TAFs) that regulate the DNA binding activity of TBP. TFIIB subsequently binds to TBP to form a TBP-TFIIB complex, which is bound by RNA pol II in association with TFIIF. Finally, TFIIE and TFIIH associate with the complex to form the PIC. Following PIC formation, TFIIH, which contains DNA helicase activity, uses ATP to unwind the DNA at the TSS, thereby exposing the template strand. TFIIH also phosphorylates the C-terminal domain of RNA pol II, inducing a conformational change which allows RNA pol II to be released from the general transcription factors and begin the elongation phase of transcription (Thomas and Chiang, 2006).

The formation of this PIC-promoter complex is sufficient for a basal level of transcription, however, activated transcription is dependent on the recruitment of gene-, stimuli- or tissue-specific transcription factors to their cognate binding sites in the promoter regions of target genes. These transcription factors can influence the recruitment of general transcription factors and the assembly of the PIC to manipulate the activity of the basal machinery as well as regulating gene expression by assisting the processitivity of RNA pol II during elongation. This regulatory action of transcription factors allows cells to control the levels of specific gene transcription e.g. in response to various stimuli (Papadakis *et al.,* 2004).

Overall, the core promoter motifs (i.e. TATA box) and transcription factor binding sites are imperative for activated transcription and the coupling of these components in synthetic promoters represents an attractive strategy to harness the cellular transcriptional machinery and exploit the endogenous activity of pharmacologically- or physiologically-inducible transcription factors for transcriptionally-regulated gene therapy.

1.2.6.2. Pharmacologically-regulated gene expression

In pharmacological regulated systems, the synthetic promoter contains the binding sites of transcription factors responsive to antibiotics such as tetracycline where gene expression is reversibly turned 'on' or 'off' in the absence or presence of tetracycline (Tet) or doxycycline

(Dox; a Tet derivative). The Tet-regulated system is analogous to the lac operon in prokaryotes where the transcription of bacterial resistance genes is negatively regulated by the Tet repressor protein (TetR) binding to the Tet operator (*TetO*) DNA sequence. In the presence of Tet, Tet binds to TetR and prevents TetR from binding to the promoter region of the gene, thereby allowing transcription to proceed. The components of the system were adapted into a regulatory system suitable for eukaryotic cells by fusing the TetR to the C-terminal domain of VP16 from HSV to create the synthetic transcription factor known as the tetracycline-controlled transactivator (tTA), that activates transcription from Tet-responsive synthetic promoters (P_{tet}; 7 *TetO* repeats cloned upstream of the mCMV promoter), in the absence of Tet (Gossen and Bujard, 1992). An inducible Tet-On system was subsequently developed by mutating the TetR to generate the rtTA (reverse Tet-repressor), which displays the reverse phenotype and induces gene expression in the presence of the Tet analogue, Dox (Gossen *et al.,* 1995). Both Tet-systems permit stringent control of gene expression in response to varying concentrations of Tet or Dox, with no immunogenicity in mouse and primate models and has been the pharmacological-inducible system of choice in both *in vitro* assays and *in vivo* gene therapy systems.

Using the Tet-On system, therapeutic viral IL-10 expression from AAV vectors was successfully regulated and significantly reduced arthritis, following the intramuscular (i.m) delivery of the construct prior to disease onset in CIA mice (Apparailly *et al.,* 2002). Using the same mouse model, Gould *et al.,* (2004) also utilised the Tet-On promoter to regulate expression of the human dimeric soluble TNFR (dsTNFR) from plasmid DNA, following i.m. delivery by electroporation and disease progression was inhibited by constitutive and Tet-regulated dsTNFR in mice with mild arthritis (Gould *et al.,* 2004).

The main disadvantage of the Tet-On system is the significant basal activity of the synthetic promoter in the absence of antibiotic, which undermines the efficacy of the system. This was addressed by developing the tetR-KRAB (tetracycline repressor-Kruppel associated box) which binds to the P_{tet} and reduces the basal activity (Forster *et al.,* 1999). Also, an improved

transactivator rtTA2^S-M2 was developed, which exhibits greater stability than rtTA, enhanced responsiveness to Dox and enhanced *TetO* binding (Urlinger *et al.,* 2000). By incorporating these improved components Gould *et al.,* (2007) demonstrated the inhibition of disease progression by dTNFR in CIA mice, following the i.m delivery of the improved Dox-regulated plasmid DNA system (Gould *et al.,* 2007). However, another drawback associated with drug-inducible synthetic promoters is the requirement to constantly monitor the disease in order to achieve optimal efficacy in RA. This is also made challenging by the variable and relapsing clinical course of RA, and therefore highlights the need for self-regulating inducible systems e.g. disease-inducible synthetic promoters.

1.2.6.3. Physiologically-regulated gene expression

Physiologically-regulated synthetic promoters are an attractive alternative to pharmacologically-regulated synthetic promoters, as they offer the advantage of autoregulation by harnessing the pathophysiological characteristics of the disease to regulate gene expression (Adriaansen *et al.,* 2006). For example, inflammation-inducible promoters are ideal systems for RA gene therapy, as they can potentially induce high levels of anti-inflammatory agents during disease flare and lower levels during disease remission, to mirror the clinical course and severity of the patients' inflammatory response in the joint.

Pro-inflammatory cytokines, heat shock proteins, acute phase proteins, some transcription factors and hypoxia-responsive genes are all associated with the pathogenesis of RA and their promoters or binding sites can be harnessed for disease-inducible gene expression. The pioneering research of Varley *et al.,* (1997) reported the development of a two-component adenoviral expression system which was responsive to inflammatory stimuli *in vivo*. In this system, the complement factor 3 (C3) promoter regulates the transcription of the HIV transactivator of transcription (Tat) protein, and Tat then stimulates the HIV LTR promoter for downstream transgene expression in the same vector. Both the C3 and HIV LTR promoters

have NFκB binding sites and are therefore NFκB-responsive and inducible by inflammation. The feasible use of the C3-Tat/HIV promoter as an inflammation-inducible promoter in RA was demonstrated in different experimental arthritis models using adenoviral-vector delivered hIL-1Ra (Bakker *et al.,* 2002) and hIL-10 (Miagkov *et al.,* 2002) and both systems resulted in diminished disease activity and decreased paw swelling. Although this system can achieve autoregulated gene expression in response to inflammation, the foreign Tat protein has a high immunogenic potential which renders this system unsuitable for RA gene therapy.

As an alternative approach, van de Loo *et al.,* (2004) developed a human IL-1β/IL-6 synthetic promoter, comprised of the human IL-1β enhancer sequence upstream of the human IL-6 promoter, and demonstrated low basal activity and high reporter gene expression during the inflammatory response (van de Loo *et al.,* 2004). The therapeutic efficacy of regulated expression of locally delivered adenoviral IL-4 (Geurts *et al.,* 2007), and lentiviral IL-10 (Henningsson *et al.,* 2012) under the control of the human IL-1β/IL-6 synthetic promoter was demonstrated in CIA mice, and confirmed the feasibility of inflammation-inducible synthetic promoters for RA gene therapy.

Other innovative systems have exploited the upregulated expression of transcription factors in RA such as NFκB. These NFκB-responsive promoters contain NFκB binding sites and during inflammation, NFκB binds to its cognate binding sites within the synthetic promoter to induce therapeutic gene expression, as demonstrated by Khoury *et al.,* (2007). In this study, the local intra-articular (i.a.) administration of AAV-anti-TNF agents, under the control of an NFκB-responsive promoter, delayed the disease onset and decreased the incidence and severity of joint damage in CIA mice and therapeutic gene expression was strongly correlated with joint inflammation (Khoury *et al.,* 2007). Similarly, hypoxia-responsive promoters containing hypoxia-responsive elements (HRE) have demonstrated hypoxia-inducible gene expression, particularly in cancer gene therapy, following their delivery into hypoxic cancerous cells (Shibata *et al.,* 2000). With focus on RA, Subang *et al.,* (2012) recently constructed a novel hybrid promoter containing HRE and *TetO* in the same synthetic promoter, which was

inducible by hypoxia and doxycycline. This innovative system provided an additional layer of regulation as gene expression can be terminated by tetR-KRAB even in the presence of the pathophysiological stimuli, which cannot be achieved with standard transcription-factor responsive promoters (Subang *et al.,* 2012).

Overall, the suitability of transcription-factor responsive promoters and similar systems is highly dependent on the expression profile of the transcription factors during the pathophysiological conditions in the RA joint. The endogenous activity of numerous TFs can be harnessed to achieve disease-specific gene expression (i.e. inflammation or hypoxia-inducible) due to their upregulated expression in response to the characteristic pathophysiological stimuli in RA, described below.

1.3. Transcription factors in RA

Pathological conditions such as hypoxia and inflammation activate a range of transcription factors in RA which respond by initiating the transcription of target genes, many of which are involved in perpetuating the disease process. Therefore, transcription factors play a fundamental role in the pathophysiology of RA.

1.3.1. Activator Protein-1 (AP-1)

The activator protein-1 (AP-1) transcription factors are homodimers or heterodimers composed of members of the Jun (c-Jun, JunB and JunD), Fos (c-Fos, FosB, Fra-1 and Fra-2), Jun dimerisation partners (JDP1 and JDP2) and related activating transcription factors (ATF2, LRF1/ATF3 and B-ATF) protein families. In contrast to the Fos proteins which can only form heterodimers with members of the Jun family, the Jun proteins can both homodimerise with each other or heterodimerise with Fos members to create transcriptionally active complexes (Shaulian and Karin, 2001). These AP-1 subunit proteins belong to the basic leucine zipper (bZIP) family of transcription factors which share a DNA-binding domain rich in basic amino acids and a leucine zipper structure required for protein-protein interactions. Despite these structural similarities, the various AP-1 dimers exert significant differences in their DNA binding affinity and specificity. For example, the AP-1 dimers Jun-Jun, Jun-Fos and Jun-Fra-2 usually bind to the heptamer consensus sequence 5'-TGA(C/G)TCA-3', which is also known as the TREs (TPA (12-O-tetradecanoylphorbol-13-acetate) / PMA (phorbol-12-myristate-13-acetate) response elements) based on their ability to induce transcription in response to the phorbol ester TPA/PMA (Angel *et al.,* 1987).

AP-1 activity is induced in response to various physiological and pathological stimuli including cytokines, stress signals, growth factors, bacterial and viral infections and also oncogenic stimuli. These signals are transferred by the stress-activated protein kinases (SAPK) p38 and JNK, and also to the extracellular signal-regulated kinase (ERK) which belong to the family of

mitogen-activated protein kinase (MAPK). The MAPK signalling pathways regulate the transcription of AP-1 sub-proteins that form the functional AP-1 transcription factor, which in turn binds to the AP-1 binding sites in the promoter regions of target genes involved in physiological and pathological processes including development, cell proliferation, organogenesis, bone cell biology, apoptosis and tumour development, as indicated by studies using cells and mice deficient in individual AP-1 proteins (Jochum *et al.,* 2001).

Many pro-inflammatory cytokines (i.e. TNFα and IL-1β), collagenase and MMPs are under the direct regulation of AP-1 via promoter binding, indicating the involvement of AP-1 in inflammatory responses and over the years, substantial evidence has identified AP-1 (Fos/Jun) as an integral transcription factor involved in the pathogenesis of RA. Earlier studies by Asahara *et al.,* (1997) showed that AP-1 had a high DNA-binding activity in the synovial tissues of RA patients compared to little or no activity in OA patients and this AP-1 activity was mainly detected in adherent cells, particularly in synovial cells and macrophages. Also, the AP-1 activity was significantly higher in mononuclear cells infiltrating into the RA synovium compared to RA peripheral blood mononuclear cells. Interestingly, the high DNA-binding activity of AP-1 in RA cells and tissues correlated with the increased expression of c-fos and c-jun mRNA *in situ* and also the disease activity. Furthermore, the constitutive upregulation of AP-1 in the RA synovium highlighted the potential role of AP-1 in synovial hyperplasia and abnormal immune responses (Asahara *et al.,* 1997).

RASFs are internally driven by upregulated AP-1 and release the IL-1β cytokine. Both AP-1 and IL-1β influence each other's gene expression and activity, resulting in a coordinated cross-talk that is fundamental to arthritic joint destruction, such as the development of the characteristic synovial overgrowth or pannus (Shiozawa and Tsumiyama, 2009). AP-1 is upregulated during the G_1 phase of the cell cycle and directly transactivates Wee1 kinase, which in turn, inhibits mitotic cell division by phosphorylating cdc2 and therefore, AP-1 promotes the development of synovial hyperplasia (Kawasaki *et al.,* 2003). This is consistent with earlier transfection studies which showed that the overexpression of c-Fos in cultured

synovial cells promoted the growth of the synovial cells *in vitro* (Kuroki *et al.,* 1993). The hyperplastic synovial overgrowth contributes to cartilage degradation through the action of MMPs, many of which are transcriptionally regulated by AP-1. AP-1 is also involved in physiological bone development, however in RA, AP-1 is implicated in bone resorption. RANKL is a known inducer of JNK, which phosphorylates and activates the Jun proteins of AP-1. Subsequently, AP-1 stimulates NFATc1, which is a key transcription factor involved in osteoclast differentiation (Wagner and Matsuo, 2003). Other than osteoclasts, MMPs contribute to the bone destruction associated with RA, which supports the role of AP-1 in regulating the expression of important genes involved in the joint destruction in RA and highlights the potential of therapeutically targeting AP-1 for the treatment of RA.

The effect of AP-1 inhibition has been studied to confirm the pathological roles of AP-1 in RA. For example, Shiozawa *et al.,* (1997) delivered short double-stranded DNA oligonucleotides, containing the AP-1 consensus sequence, into CIA mice via intraperitoneal injection. These 'decoy' oligonucleotides recognise and bind to AP-1, rendering the transcription factor incapable of binding to the promoters of their target genes. The authors reported the inhibition of arthritic joint destruction in a sequence-specific and dose-dependent manner by the 'decoy' AP-1 oligonucleotides, which inhibited gene expression at the transcriptional level (Shiozawa *et al.,* 1997). Using a different approach, Aikawa *et al.,* (2008) designed and synthesised a specific small inhibitor of c-Fos/AP-1 using 3D pharmocophore modelling based on the crystal structure of the AP-1-DNA complex. The daily oral administration of the drug, named T-5224, in CIA mice resulted in the inhibition of most MMPs, including MMP-3, -9, -13 as well as IL-1β and also inhibited activities of osteoclasts and synovial cells. These studies confirmed the important role of AP-1 in the pathogenesis of RA by selectively inhibiting AP-1, which resolved joint destruction in a preclinical model of RA (Aikawa *et al.,* 2008).

1.3.2. CCAAT/enhancer-binding protein (C/EBP)

The CCAAT/enhancer-binding proteins (C/EBP) are a family of six structurally similar yet functionally distinct basic leucine zipper (bZIP) transcription factors: C/EBPα, -β, -δ, -γ, -ε, and –ζ. These C/EBP family members are defined by their conserved carboxy-terminal domains consisting of a leucine-zipper dimerisation domain and basic DNA-binding domains and an amino terminal activation domain (Lekstrom-Himes and Xanthopoulos, 1998). Dimerisation to form homo- or heterodimers with related proteins is a prerequisite for the DNA-binding activity of C/EBP family members which is mediated by the basic region. With small variations, the C/EBP transcription factors recognise and bind to the consensus sequence 5'-RTTGCGYAAY-3', (where R is A/G and Y is C/T) in the promoter regions of target genes to regulate a range of physiological events including tissue differentiation, immune responses and energy metabolism (Osada *et al.,* 1996; Ramji and Foka, 2002).

The tissue/cell specific expression profile of the various C/EBP transcription factors has been observed at the mRNA and protein level however, the discrepancies between different reports may be due to species-specific differences and the use of different techniques and experimental conditions. In general, C/EBPα is highly expressed in the liver, intestine, lung, adipose tissue, adrenal gland and placenta whereas constitutive expression of C/EBPβ is highest in the liver, intestine, adipose tissue, lung, spleen, kidney and myelomonocytic cells. C/EBPδ is expressed in adipose tissue, lung and intestine, C/EBPγ, and C/EBPζ are ubiquitously expressed whilst expression of C/EBPε mRNA and protein are restricted to myeloid and lymphoid cells. Also, the expression profile of the various C/EBP isoforms is regulated by physiological and pathological signals by a range of extracellular mediators (Ramji and Foka, 2002).

The C/EBPβ mRNA produces at least three isoforms, 38kDa (LAP*; liver-enriched transcriptional activating protein), 35kDa (LAP) and 20kDa (LIP; liver-enriched transcriptional inhibitory protein), where LAP and LIP are the major polypeptides produced in cells. The LAP protein shares sequence homology (71%) in its DNA-binding and leucine zipper domains with C/EBPα, and the LAP gene was shown to encode the liver-enriched transcriptional activator

protein that was identified as C/EBPβ (Descombes *et al.,* 1990; Ossipow *et al.,* 1993). The IL-6 cytokine is known to stimulate the production of acute-phase proteins from the liver during inflammation and C/EBPβ was also identified as the nuclear factor for IL-6 expression (NF-IL6) which is induced by LPS, IL-1 and IL-6 and binds to the IL-1 response elements within the IL-6 promoter. Interestingly, C/EBPβ (NF-IL-6) also binds to the promoter of several other cytokine genes such as TNFα, IL-8 and GM-CSF implying that C/EBPβ regulates the expression of genes involved in acute phase reactions, inflammation and haemopoiesis (Akira *et al.,* 1990). Importantly, C/EBPβ is significantly upregulated in the synovial lining and sublining cells of synovial tissue from RA patients (Pope *et al.,* 1999) and the correlated activation of C/EBPβ with serum C-reactive protein (CRP) and synovial IL-6 mRNA levels suggesting that C/EBPβ may play a regulatory role in the inflammatory process in RA joints (Nishioka *et al.,* 2000).

The transition from proliferation to hypertrophic differentiation of chondrocytes is an important process in the destructive process in RA. C/EBPβ promotes the hypertrophic phenotype of chondrocytes by directly transactivating the p57^{Kip2} cell cycle factor that plays a crucial role in chondrocyte proliferation and differentiation (Hirata *et al.,* 2009). Hypertrophic chondrocytes produce cartilage degrading proteases which contribute to cartilage destruction in RA and C/EBPβ has been shown to directly bind to the MMP-13 promoter resulting in the enhanced expression of MMP-13 in chondrocytes in inflammatory arthritis (Hayashida *et al.,* 2009).

1.3.3. Egr-1

Early growth response-1 (Egr-1) gene was independently discovered by a number of laboratories searching for factors regulating cell growth and proliferation. Egr-1 was first discovered following the stimulation of PCl2-cells with neuronal growth factor (NGF) in a model for neuronal differentiation and was initially referred to as nerve growth factor-induced clone A, NGFIA (Milbrandt, 1987) and later became also known as KROX24 (Chavrier *et al.,* 1989),

TIS8 (Kujubu *et al.,* 1987) and ZIF268 (Christy and Nathans, 1989). The name Egr-1 was eventually established by Sukhatme *et al.,* (1988) who demonstrated the inducible expression of Egr-1 in growth-quiescent fibroblasts exposed to serum (Sukhatme *et al.,* 1988).

The ~3.4Kb Egr-1 transcript has a short half-life (10 to 20 minutes) and encodes a nuclear phosphoprotein predicted to contain 533 amino acids. The protein is highly rich in proline and serine residues and the carboxyl terminal end has three tandem repeats of Cys_2-His_2 zinc finger domains specific for the consensus sequence 5'-CGCCCCCGC-3' and acts as a transcription factor (Sukhatme, 1990). Other members of the Egr family include Egr-2, Egr-3 and Egr-4 which have related DNA-binding zinc finger domains and share the same consensus sequence but the differences in their flanking regions suggest specific functions of the individual proteins (Decker *et al.,* 2003).

As an 'immediately early' gene product, the Egr-1 transcription factor is rapidly activated in a range of cells by multiple extracellular agonists such as growth factors and cytokines, and environmental stresses including hypoxia, fluid shear stresses and vascular injury. Also, the induction of Egr-1 by pharmacological activators, such as TPA, suggests that protein kinase C–dependent pathways are involved (Sukhatme, 1990). Following activation, Egr-1 is transiently expressed and binds to Egr-1 binding sites within the promoters of genes expressed in a range of cell types including thymocytes, B-cells, monocytes, fibroblasts, kidney cells and neurons and is associated in the differentiation of B-cells, T-cells, PC19 embryonic cells and monocytes (Aicher *et al.,* 1999a; Decker *et al.,* 2003).

Pathogenesis-relevant Egr-1 target genes include growth factors, cytokines, receptors and proteases. For example, the studies of Shin *et al.,* (2010) showed that MMP-9 is directly targeted by the Egr-1 transcription factor, which together with NFκB, synergistically activates both the basal and TNFα-induced MMP-9 promoter activity (Shin *et al.,* 2010). The TNFα promoter also contains Egr-1 binding sites, however, Egr-1 is thought to play an auxiliary role in TNFα gene regulation (Krämer *et al.,* 1994). Using *in situ* hybridisation, Aicher *et al.,* (1994)

detected elevated Egr-1 mRNA expression in regions of the synovial membrane predominantly populated by synovial macrophages and fibroblasts, compared to non-RA tissue (Aicher *et al.,* 1994), which was also confirmed using RT-qPCR detection by the studies of Grimbacher *et al.,* (1997). Egr-1 expression in RASFs is inducible by TNFα which may exacerbate the inflammatory process in the joint by autocrine stimulation and therefore the elevated Egr-1 expression may be used as a diagnostic marker to characterise synovial fibroblasts in RA patients (Grimbacher *et al.,* 1998). In fibroblasts, the most proximal serum response element in the Egr-1 promoter is the major positive regulator whereas binding to a cAMP responsive element may serve as a negative regulatory signal for Egr-1 transcription in fibroblasts. Therefore, the overexpression of Egr-1 in RASFs may result from either the activation of the RASFs by PKC signals and serum response factor activity or the failure of Egr-1 repressing signals and also, the enhanced expression of Egr-1 suggests an important role in maintaining the activated phenotype of RASFs (Aicher *et al.,*1999b).

1.3.4. Ets-1

The Ets (E26 transformation-specific) proteins belong to a superfamily of transcription factors defined by their conserved winged helix-turn–helix ETS domain. This unique DNA binding domain (ETS domain) is composed of 85 amino acids and comprises three α-helices and four β-strands arranged in the order H1-S1-S2-H2-H3-S3-S4 (Dittmer, 2003). To date, 27 mammalian *ETS* genes have been identified and are classified into 12 distinct subgroups based on their ETS domain sequence homology (Hollenhorst *et al.,* 2004).

The Ets transcription factors recognise and bind to the core 5'-GGA(A/T)-3' Ets-binding site through their ETS domain, in the promoter regions of target genes (Nye *et al.,* 1992). These transcription factors are known to act as positive or negative regulators of the expression of genes involved in a range of biological processes (e.g. signalling cascades, cellular proliferation, development, differentiation, haematopoiesis, adhesion, immune responses,

apoptosis, invasion, tissue remodelling and angiogenesis) in various cell types including B-cells, endothelial cell, fibroblasts and neoplastic cells (Shaikhibrahim and Wernert, 2012).

The highly conserved *ETS-1* gene is considered to be the founding member of the *ETS* gene family of transcription factors and was originally discovered as the cellular progenitor of the retroviral *v-ETS* oncogene (v-ets) that is associated with *v-MYB* in the genome of the E26 avian leukaemia retrovirus (Nunn *et al.,* 1983). The human *ETS-1* gene, located on chromosome 11, encodes a single mRNA of 6.8Kb (Watson *et al.,* 1985) and Ets-1 has been established as an important transcription factor in the regulation of angiogenesis under both physiological and pathological conditions. By *in situ* hybridisation, Wernert *et al.,* (1992) reported the upregulated expression of Ets-1 mRNA during angiogenesis e.g. in embryonic development, granulation tissue formation and especially during tumour vascularisation whereas mature vessels without angiogenic activity did not express Ets-1 mRNA. The *in vitro* studies in human umbilical vein endothelial cells (HUVECs), demonstrated transient and inducible expression of Ets-1 and urokinase-type plasminogen activator (u-Pa) mRNA by TNFα stimulation which suggested that Ets-1 may regulate the transcription of the matrix degrading proteases e.g. u-Pa during angiogenesis and tumour invasion (Wernert *et al.,* 1992).

During angiogenesis, the endothelial cells produce matrix-degrading proteases which degrade the basement membrane that surrounds pre-existing capillaries or venules and are therefore important in the invasive properties of endothelial cells and the formation of new blood vessels. The studies of Iwasaka *et al.,* (1996) not only demonstrated the inducible expression of Ets-1 mRNA by the angiogenic growth factors, aFGF, bFGF, VEGF, EGF, in HUVECs, immortalised HUVECs and human omental microvascular endothelial cells (HOMECs) but also confirmed the role of Ets-1 in regulating the expression of proteinases during angiogenesis. Antisense oligodeoxynucleotides (ODNs) against Ets-1 mRNA efficiently blocked the synthesis of the Ets-1 transcription factor by human endothelial cells in response to angiogenic factors which also inhibited the expression of the proteases, u-Pa and MMP-1, as well as the migration of

endothelial cells (Iwasaka *et al.,* 1996). Using a similar approach, Wernert *et al.,* (1999) reported the *in vivo* reduction of angiogenesis in chick embryos following the delivery of Ets-1 antisense ODNs (Wernert *et al.,* 1999). Similarly, Hashiya *et al.,* (2004) demonstrated the ability of Ets-1 to stimulate angiogenesis in a rat hindlimb ischemia model following Ets-1 gene delivery using hemagglutinating virus of Japan (HVJ)-liposome mediated transfection of plasmid DNA encoding the Ets-1 gene. The promoter regions of the angiogenic factors vascular endothelial growth factor (VEGF) and hepatocyte growth factor (HGF) contain Ets-1 binding sites and overexpression of Ets-1 upregulated both VEGF and HGF in the rat hindlimb (Hashiya *et al.,* 2004). Collectively, these results demonstrate that Ets-1 regulates angiogenesis through the upregulation of matrix degrading proteases e.g. u-Pa and MMP-1 and the upregulation of angiogenic growth factors e.g. VEGF and HGF which bind to their specific receptors on endothelial cells and initiate a signal transduction cascade to activate the required transcription factors, thus placing Ets-1 upstream of the angiogenic cascade.

Evidence also implicates the Ets-1 transcription factor in RA; using *in situ* hybridisation, immunohistochemistry and RT-qPCR, Redlich and colleagues (2001) reported the overexpression of Ets-1 mRNA in the synovial tissues from all RA patients evaluated, particularly in the synovial lining layer and the sublining areas. Both IL-1β and TNFα induced the upregulation and nuclear localisation of Ets-1 in cultured synovial fibroblasts from RA patients, suggesting the important role of Ets-1 in cytokine-mediated inflammatory cascade in RA (Redlich *et al.,* 2001). Using similar experimental techniques, Wernert *et al.,* (2002) reported the elevated expression of Ets-1 mRNA and Ets-1 protein within endothelial cells of newly developing small blood vessel (during angiogenesis) in the synovial membranes from patients with RA compared to weak Ets-1 expression in those from OA patients, suggesting the importance of Ets-1 in the regulation of inflammatory angiogenesis in RA (Wernert *et al.,* 2002). Taken together, Ets-1 is an important transcription factor in the pathogenesis of RA due to (1) the regulation of target matrix degrading enzymes and potential contribution to the destructive processes in RA, (2) the involvement of Ets-1 in angiogenesis, which may

contribute to pannus expansion and (3) due to the inducible expression of Ets-1 by TNFα and IL-1β and (4) the potential role of Ets-1 as a major mediator downstream of these pro-inflammatory cytokines.

1.3.5. Hypoxia Inducible Factor-1α (HIF-1α)

The healthy synovium has a rich blood supply to provide the avascular articular cartilage with oxygen and nutrients (Walsh and Pearson, 2001). However in RA, synovial hyperplasia and the migrating, invasive pannus increases the distance between the proliferating cells and the nearest blood vessels which exceeds the oxygen diffusion limit (>100-200µm) resulting in local hypoxia and hypoperfusion (Gaber *et al.,* 2005). Consequently, there is an increased demand of oxygen and nutrients which promotes the cellular transition from oxidative phosphorylation to anaerobic glycolytic metabolism and stimulates the local formation of new capillary blood vessels, 'angiogenesis', from pre-existing vasculature to increase the blood supply to the synovium and improve the influx of oxygen and nutrients to the cells (Paleolog, 2002). Synovial angiogenesis also ensures a constant supply of cytokines and growth factors which perpetuates inflammation, synovial hyperplasia and infiltration of immune cells and therefore, angiogenesis is recognised as a key event in the pathophysiology and perpetuation of RA (Koch, 2003).

Angiogenesis plays a central role in normal physiological processes such as embryonic development, wound and fracture healing and in the female reproductive cycle, and the process is closely correlated with the hypoxic environment (Paleolog, 2009). The existence of hypoxia in RA was demonstrated over forty years ago by Lund-Olesen (1970) who recorded the oxygen tension (pO_2) measurements in the synovial fluid of patients with RA using a Clarke-type electrode and reported the mean synovial fluid pO_2 in RA knee joints to be 27mmHg which was significantly lower compared to 43mmHg in patients with OA and 63mmHg in healthy controls (Lund-Olesen, 1970). Subsequent studies also reported decreased oxygen tension (hypoxia) and decreased glucose levels as well as increased

carbon dioxide, lactate and acetate levels in RA joints, which are consistent with anaerobic metabolism (Treuhaft and McCarty, 1971; Ahlqvist, 1984). Further confirmation was provided by the studies of Naughton *et al.,* (1993) who used magnetic resonance spectroscopy to display a profile of low molecular weight metabolites consistent with hypoxia (Naughton *et al.,* 1993) and Peters *et al.,* (2004) who reported the induction of hypoxia during inflammatory responses in AIA rat joints by using hypoxyprobe-1, which is a compound that binds to cells with low oxygen tension to detected hypoxia in structures such as the synovium, pannus, bone marrow and articular cartilage (Peters *et al*., 2004).

Although there is much evidence reporting the increased blood supply to the hypoxic synovium, angiogenesis is unlikely to be adequate to compensate for the increased demand of oxygen and nutrients. In addition to synovial hyperplasia and increased metabolic demand, other contributory factors to hypoxia in RA joints include the accumulation of synovial fluid and joint movement, as evidenced by the studies of Jawed *et al.,* (1997). This group demonstrated the significant increase in intra-articular pressure after isometric quadricep contractions in patients with RA synovitis, which temporally impedes synovial blood flow, causes oxidative injury and exacerbates the pre-existing hypoxic environment in the joint (Jawed *et al.,* 1997).

Hypoxia induces profound changes in the cellular gene expression profile with the ultimate aim of restoring oxygen homeostasis, fundamental for cell survival. The hypoxia-inducible factor (HIF) is a 'master regulator' of this cellular response to hypoxia and molecular cloning of the HIF-1 transcription factor by Wang *et al.,* (1995) revealed that HIF-1 binds to the 3'-enhancer region of the erythropoietin (EPO) gene and is necessary for transcriptional activation of EPO in hypoxia. The HIF-1 transcription factor is composed of a 120kDa α-subunit complexed with a 91-94 kDa β-subunit; HIF-1α is closely related to Sim and HIF-1β is also known as ARNT (aryl hydrocarbon receptor nuclear translocator) and both subunits are basic helix-loop-helix (bHLH) proteins which share highly similar PAS (Per/ARNT/Sim) domains (Wang *et al.,* 1995).

Importantly, HIF-1β is constitutively expressed and oxygen independent whereas HIF-1α contains an ODD (oxygen-dependent degradation domain), which is essential for proteolytic regulation of this subunit under normoxia. In normal physiological conditions, HIF-1α undergoes rapid hydroxylation of the proline residues, Pro402 and Pro564 within the ODD by the prolyl hydroxylase enzymes (PHD) which also require molecular oxygen, iron and 2-oxoglurate as fundamental co-factors in the reaction. The hydroxylated HIF-1α subunit is recognised by the tumour suppressor Von Hippel-Lindau protein (pVHL), which is a component of a multiprotein ubiquitin ligase complex (pVHL-E3-ligase complex) responsible for targeting HIF-1α for polyubiquitination. Following polyubiquitination, HIF-1α subunits are degraded by the 26S proteasome. The impaired function of pVHL or PHD or the inhibition of the hydroxylation reaction results in a constitutively stabilised HIF-1α and activated HIF transactivation, as demonstrated during hypoxia (Gaber *et al.,* 2005; Kenneth and Rocha, 2008).

Cellular HIF-1α protein levels increase exponentially to a maximum level at approximately 0.5% which corresponds to pO_2 values of 10-15mmHg (Jiang *et al.,* 1996). During these hypoxic conditions, hydroxylation of proline residues is suppressed due to the lack of oxygen and therefore, HIF-1α is not targeted by pVHL mediated proteasomal degradation resulting in the accumulation of high levels of HIF-1α, which subsequently translocates to the nucleus and binds to HIF-1β to form a stabilised, transcriptionally active HIF heterodimer, leading to HIF target gene expression (Kenneth and Rocha, 2008). The heterodimer contains an N-terminus bHLH domain which mediates the binding of the HIF transcription factor complex to hypoxia-responsive elements (HREs) with the consensus sequence 5'-RCGTGG-3' within the promoter region of HRE-containing target genes (Wang and Semenza, 1993). More than 60 HIF targets have been identified and include genes involved in angiogenesis, erythropoiesis, glycolytic metabolism, iron metabolism, pH control, apoptosis, cell proliferation, cell migration and matrix metabolism (Gaber *et al.,* 2005).

VEGF is the best characterised proangiogenic stimulus and VEGF gene transcription is directly activated by HIF-1α during hypoxia (Forsythe *et al.,* 1996). VEGF levels are significantly higher in the serum and synovial fluid of RA patients compared to patients with OA or normal controls (Harada *et al.,* 1998; Lee *et al.,* 2001). Hypoxia is a potent stimulus for VEGF induction in RA synovial membrane cells and VEGF production is further increased in hypoxic RASFs in the presence of inflammatory cytokines such as IL-1β and TGFβ (Berse *et al.,* 1999). VEGF, HIF-1α and the HIF-2α isoform are highly expressed in RA synovial macrophages and fibroblasts (Hollander *et al.,* 2001; Giatromanolaki *et al.,* 2003) and in an experimental AIA rat model (Peters *et al.,* 2004), suggesting that pro-inflammatory cytokines during inflammation and synovial hypoxia promotes the upregulation of HIF-1α and transcriptional activation of VEGF in the joint leading to the induction of synovial angiogenesis which perpetuates synovitis. Targeting various points within the HIF-1α pathway, whether inhibiting the transcription factor itself or downstream products such as VEGF, have demonstrated anti-angiogenic effects in RA (Semenza, 2000).

1.3.6. Nuclear Factor κB (NFκB)

NFκB is a collective name for dimeric transcription factors which belong to the Rel family of proteins and include RelA (p65), c-Rel, RelB, NFκB1 (p50 and its precursor p105) and NFκB2 (p52 and its precursor p100). These NFκB subunits form a variety of homodimers and heterodimers to produce the NFκB transcription factors, each of which positively or negatively regulate the expression of their own set of genes and also share a 300-amino acid N-terminal Rel homology domain (RHD), which mediates their DNA binding, dimerisation and nuclear translocation (Tak and Firestein, 2001). NFκB transcription factors can bind to different variations of the NFκB consensus sequence 5'-GGGRNWYYCC-3' to activate, or sometimes, repress target gene transcription in a cell-type and stimulus-specifc manner. Only p65, c-Rel and RelB possess C-terminal transactivation domains (TADs), enabling them to initiate

transcription. However, despite the absence of TADs in p52 and p50, these subunits can positively regulate transcription through heterodimerisation with TAD-containing NFκB subunits or by interacting with non-Rel proteins capable of transactivation. Conversely, p50 and p52 homodimers can compete with TAD-containing dimers for binding to κB sites on the DNA, to negatively regulate transcription. The most prevalent activated form of NFκB is a heterodimer consisting of a p50 or p52 subunit and p65, often referred to as the 'classic' NFκB and is involved in the 'classical' or 'canonical' NFκB pathway (Hayden and Ghosh, 2012).

In unstimulated cells, canonical NFκB dimers are found sequestered in the cytosol and are rendered inactive by the Inhibitor of κB (IκB). All canonical IκB proteins (IκBα, IκBβ, IκBε) contain C-terminal ankyrin repeat domains (ARDs) which form a large interaction surface around the NFκB dimer and inhibits the nuclear localisation signals of NFκB to prevents its entry into the nucleus. A wide range of stimuli including pro-inflammatory cytokines (i.e. TNFα and IL-1β), chemokines, growth factors, bacterial products and viral proteins, UV radiation, ischemia/reperfusion and oxidative stress can activate the ubiquitous NFκB p65/p50 dimer through the canonical NFκB signalling pathway. During inflammatory signalling, the engagement of TNFR, IL-1R and TLRs with their cognate ligands results in the phosphorylation-dependent activation of the IκB kinases (IKK), composed of the two catalytic subunits IKKα and IKKβ, and the scaffolding protein IKKγ/NFκB essential modulator (NEMO). The p65/p50 NFκB complex is predominantly bound by IκBα and IKK phosphorylates IκBα and other IκB proteins, thereby targeting them for ubiquitination and degradation by the 26S proteasome. The NFκB nuclear localisation signal is exposed and enables the active NFκB to translocate into the nucleus and bind to the NFκB-binding site in the promoter regions of target genes. Many NFκB-target genes are involved in a broad range of physiological and pathological processes including; maintenance of the immune system, skeletal system, control of cell survival, differentiation, proliferation, inflammation, apoptosis and also the activation of genes involved in cancer, autoimmune diseases, neurodegenerative diseases, cardiovascular disease, diabetes, RA and many others (Hayden and Ghosh, 2012). The activation of NFκB

is often transient and the termination of NFκB activity is primarily achieved by the ability of NFκB upregulating its own inhibitors of the IκB family. The newly synthesised IκB enters the nucleus and removes NFκB from the DNA to relocate and sequester NFκB in the cytosol. However, during pathological conditions such as chronic inflammation, the continuous presence of NFκB-activating stimuli appears to outperform the inhibitory feedback loop, resulting in an elevated and constitutive activity of NFκB (Hoesel and Schmid, 2013).

The non-canonical NFκB pathway is typically induced by the binding of CD40 ligand, BAFF, RANKL and lymphotoxin-β, to their cognate receptors, resulting in the activation of the NFκB inducing kinase (NIK) which phosphorylates and activates predominantly IKKα. IKKα phosphorylates the precursor protein p100 resulting in its ubiquitination and partial degradation to give rise to the mature p52 subunit and subsequently, the p52/RelB dimer, which can translocate to the nucleus to activate gene transcription. In contrast to the canonical pathway, the non-canonical pathway is dependent on IKKα and independent of IKKβ and NEMO (Simmonds and Foxwell, 2008).

NFκB (particularly canonical p65/p50) is fundamental to the pathogenesis of RA by positively regulating a wide range of genes encoding molecules involved the disease, including cytokines (e.g. TNFα, IL-1β, IL-2, IL-6, IL-8, IL-12, IFNγ, GM-CSF), adhesion molecules (e.g. E-selectin, ICAM-1, VCAM-1), chemokines (e.g. MIP-1α, MCP-1, RANTES) and many other genes involved in cell proliferation, apoptosis and cell cycle progression. NFκB is also inducible by many of these pro-inflammatory cytokines resulting in a positive feedback loop, which promotes the constitutive and high activity of NFκB in RA (Roman-Blas and Jimenez, 2006). Activated NFκB has been detected in human RA synovial tissue (e.g. Handel *et al.,* 1995) and particularly in RASFs (Miagkov *et al.,* 1998) as well as in different animal models of RA including AIA rats, CIA mice, streptococcal cell wall (SCW)-induced and pristane-induced arthritis in rats (Makarov, 2001). Substantial evidence has demonstrated the important role of NFκB in the various processes in RA, especially, during synovial inflammation which underlies the pathology of RA. It is widely recognised that NFκB regulates the

expression of the vast majority of pro-inflammatory cytokines involved in RA, and studies including those which have inhibited the activity of NFκB (e.g. via the delivery of genes encoding dominant negative IKKβ mutant), have provided confirmation of the importance of NFκB for the induction of pro-inflammatory cytokines in human fibroblast-like synoviocytes (Aupperle *et al.,* 2001).

The characteristic synovial overgrowth is largely attributed to enhanced cell proliferation and decreased apoptosis. Many NFκB-target genes are pro-apoptotic stimuli e.g. TNFα and c-Myc, and in the hyperplastic synovium, NFκB activation delivers anti-apoptotic signals which enhance cell survival. Therefore, NFκB contributes to the proliferation of synovial cells by inducing the expression of c-Myc and cyclin D1 proteins required for cell cycle progression, as well as inhibiting the cytotoxic effects of c-Myc and TNFα and inducing the expression of anti-apoptotic molecules such as B-cell CLL/lymphoma 2 (Bcl-2) family members (Barkett and Gilmore, 1999).

Cartilage degradation and bone resorption by RASFs is also mediated by NFκB, which transcriptionally activates several MMPs e.g. MMP-1 which has NFκB-binding sites in its promoters, and NFκB is also important for the induction of MMP-3, -9 and -13. NFκB activation may also have an important role in the differentiation and maturation of osteoclasts in RA, as implied by the development of osteopetrosis in *nfκb1$^{-/-}$nfκb2$^{-/-}$* double-knockout mice due to the accumulation of immature osteoclasts (Iotsova *et al.,* 1997). Furthermore, Jimi *et al.,* (2004) showed that the delivery of a cell-permeable peptide inhibitor of the IKK complex, which is a selective inhibitor of the NFκB pathway, blocked osteoclastogenesis and prevented inflammatory bone destruction in CIA mice (Jimi *et al.,* 2004). Overall, the inhibition of NFκB activity using knockout mice, decoy oligonucleotides, IκB mutants, or selective inhibitors of IKK complexes appear to be effective therapeutic strategies for the treatment of RA, however, the systemic blockade of NFκB activity can result in detrimental effects due to its involvement in numerous physiological processes. Nevertheless, these studies have confirmed the fundamental role of NFκB in the pathogenesis of RA.

1.3.7. Co-operative activity of the candidate transcription factors

The candidate transcription factors HIF-1α, AP-1, Egr-1, Ets-1, C/EBPβ and NFκB orchestrate many of the inflammatory and destructive processes in RA and their co-operative/synergistic activity has also been described in RA. Interestingly, the regulation of many promoters by ETS transcription factors often depends upon their physical interaction with unrelated transcription factors on composite DNA elements (Goetze et al., 2001). This is particularly true for Ets-1 which physically interacts with transcription factors bound to their cognate binding sites located close to the Ets-1-binding site. For example, Ets-1 family members and AP-1 (Fos/Jun) transcription factors functionally cooperate in a large number of promoters which is vital for the regulated expression of numerous genes including cytokines (Gottschalk *et al.,* 1993), MMPs (Logan *et al.,* 1996) and the TNFα promoter which contains Ets-1 binding sites in direct juxtaposition to the AP-1 binding sites (Krämer *et al.,* 1995). Also, Ets-1, NFκB and AP-1 synergistically transactivate the human granulocyte-macrophage colony stimulating factor (GM-CSF) promoter in an interdependent manner in T cells (Thomas *et al.,* 1997).

In addition to homo- and heterodimersation of C/EBP transcription factors, these proteins can form protein-protein interactions with related and unrelated nuclear transcription factors. For example, C/EBPβ has been shown to synergistically interact with the NFκB to activate the transcription of IL-6 and IL-8, both of which contain adjacent C/EBPβ and NFκB binding sites in their promoters (Matsusaka *et al.,* 1993). C/EBPβ is also known to cooperate with many other transcription factors including AP-1 in the regulation of genes including collagenase and cytokines such as IL-8 and TNFα (Ramji and Foka, 2002).

Despite the ubiquitous synthesis of Egr-1 transcription factors, Egr-1-regulated genes are expressed in a tissue-specific manner and such specificity may be due to interactions with other transcription factors that bind to adjacent sites in a given promoter. For example, Egr-1

synergistically interacts with nuclear factor of activated T cells (NFAT) to form heterodimers which bind in close contact in the IL-2 and TNFα promoters to regulate the gene expression of these cytokines (Decker *et al.,* 1998; Decker *et al.,* 2003). Synergistic function of Egr-1 with other nuclear factors has also been reported e.g. with the homeobox protein Ptx1 and the steroidogenic factor Sf1 in the regulation of the luteinizing hormone LH-b expression (Tremblay and Drouin, 1999), with the p65 protein in NFκB p50 gene induction (Cogswell *et al.,* 1997), with the CBP/p300 in the transcription of the lipoxygenase gene (Silverman *et al.,* 1998), and with the transcriptional inhibitors NGFI-A binding protein (NAB)-1 and NAB-2 (Russo *et al.,* 1995; Svaren *et al.,* 1996), among others which contribute to tissue-specific expression of ubiquitously expressed nuclear factors such as Egr-1.

Ets-1 and HIF-1α are colocalised in the hypoxic synovium of inflamed joints of rats with adjuvant-induced arthritis, suggesting that both hypoxia and HIF-1α may be involved in the upregulation of Ets-1 during joint inflammation (Peters *et al.,* 2004). In addition to the HIF family, hypoxia activates a number of other important transcription factors such as NFκB, AP-1, and p53 among others, where there is often crosstalk and cooperative activity between the transcription factors (Kenneth and Rocha, 2008).

1.4. MicroRNAs

1.4.1. MicroRNA discovery

MicroRNAs (miRNAs) are endogenously produced, small (~22 nucleotides) non-coding RNAs involved in the post-transcriptional regulation of target messenger RNAs (mRNAs). MiRNAs were discovered by Lee *et al.,* (1993) through screening for genes that control the developmental timing in the nematode *Caenorhabditis elegans (C.elegans)*. It was revealed that the lineage-4 (lin-4) gene did not encode a protein but instead encoded two small transcripts, 22 and 61 nucleotides (nt) long; the longer 61nt transcript was predicted to be the stem-loop precursor of the functional 22nt transcript. Sequence comparisons of these lin-4 RNA transcripts with the 3' untranslated region (3'UTR) of the lin-14 mRNA revealed multiple antisense complementary sites to the lin-4 sequence within the lin-14 3'UTR (Lee *et al*., 1993; Wightman *et al.,* 1993) which was in support of the earlier prediction that the lin-14 3'UTR was an important region involved in mediating lin-14 repression by the lin-4 transcripts (Wightman *et al.,* 1991). It was subsequently demonstrated that lin-4 downregulated the LIN-14 protein without impairing the lin-14 mRNA levels by interacting with the lin-14 3'UTR to translationally repress lin-14 mRNA, which in turn, regulated the timing of stage transitions in the early development of C.elegans larvae (Lee *et al*., 1993; Wightman *et al.,* 1993).

Seven years later in 2000, Reinhart and colleagues identified the let-7 gene that transcribes a temporally regulated 21nt RNA involved in the late larval developmental stage in *C.elegans.* The let-7 RNA was demonstrated to regulate the expression of lin-14 and lin-28 in a similar mechanism to lin-14 regulation by lin-4 RNA, as described by Lee *et al*., (1993) and consequently, lin-4 and let-7 RNA were termed short temporal RNAs (stRNAs) due to their small size and temporal expression during *C.elegans* development (Reinhart *et al.,* 2000).

Let-7 homologs were identified in other bilateral animals including mammals (Pasquinelli *et al.,* 2000) and soon after, over one hundred additional genes transcribing small RNA species

were identified in *Drosophila,* humans and in worms (Lagos-Quintana *et al.,* 2001; Lau *et al.,* 2001; Lee and Ambros, 2001). Analogous to the lin-4 and let-7 stRNAs, these endogenously expressed RNA species were also ~22nt in length and potentially processed from larger stem loop precursors. These RNA products were generally conserved in evolution and few RNAs exhibited temporal expression whilst the majority displayed cell-specific expression profiles. Therefore, stRNAs and other tiny RNAs with similar features and unknown functions were referred to as microRNAs (Lagos-Quintana *et al.,* 2001; Lau *et al.,* 2001; Lee and Ambros, 2001).

MiRNAs have since been found in plants, animals, green algae and viruses (Griffiths-Jones *et al.,* 2008) and are documented in the online miRBase repository for miRNA sequence data and annotation (miRBase version 20; www.mirbase.org), which currently contains 24,521 entries from rodents, primates, birds, fish, worms, flies, plants and viruses. Interestingly, other types of small RNAs such as endogenous small interfering RNAs (siRNAs) (Reinhart and Bartel, 2002; Ambros *et al.* 2003) and Piwi-interacting RNAs (piRNAs) (Aravin *et al.,* 2007) have been identified in animals, plants and fungi, which have similar RNA silencing functions to endogenous miRNAs but differ in their biogenesis.

1.4.2. MiRNA biogenesis

The production of the functional, mature ~22nt miRNA involves multiple, complex enzymatic steps, summarised in Figure 1.6.
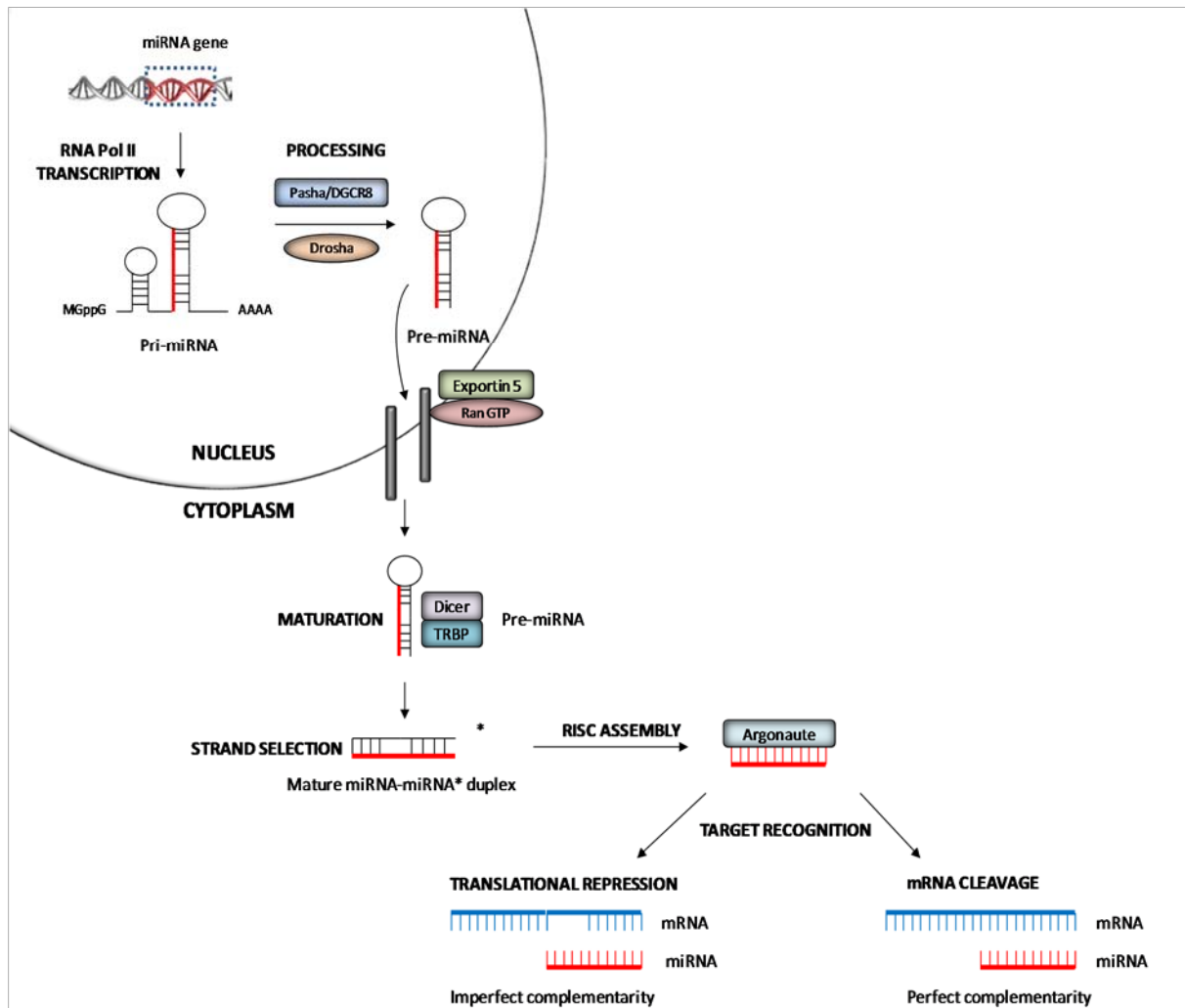
Figure 1.6. A schematic representation of the biogenesis and function of microRNAs. The primary microRNA (pri-miRNA) is transcribed from the miRNA gene by RNA polymerase II and then processed by Drosha and Pasha/DGCR8 in the nucleus to form the hairpin precursor miRNA (pre-miRNA). Pre-miRNA is actively transported into the cytoplasm via the Exportin 5 pathway and then further processed by Dicer and TRBP to form the ~22nt mature miRNA duplex. The passenger strand (*) is degraded and the selected guide strand associates with Argonaute to translationally repress or degraded target mRNAs containing imperfect and perfect complementary miRNA target sites in their 3'UTR, respectively. Figure 1.6 was redrawn, with modifications, from Kai and Pasquinelli, (2010).

MiRNA genes are often expressed from their own promoters and are located in intergenic regions or on the antisense strand of annotated genes (Zhou *et al,* 2007). The majority of miRNA genes are commonly found in the introns of protein-coding genes or non-coding host genes while few miRNA genes are located in the exons of non-coding genes (Rodriguez *et al.,* 2004). The transcription of multiple miRNAs as a single transcriptional unit, known as a 'cluster' is a common phenomenon (Lau *et al.,* 2001), which often results in these miRNAs being related and considered as family members. Generally, a family of miRNAs share nucleotides 2-8, called the 'seed region', located in the 5'-end of the miRNA, which is a vital sequence involved in target mRNA recognition (Bartel, 2004).

MiRNA biogenesis proceeds in the nucleus with the transcription of primary miRNA   (pri-miRNA) by RNA polymerase II (Lee *et al.,* 2002), although evidence of RNA polymerase III-mediated transcription has been reported (Borchert *et al.,* 2006). The transcription of a miRNA cluster or an individual miRNA gene results in either a group of pri-miRNA stem-loop structures or a single hairpin pri-miRNA. The resulting pri-miRNAs (>1Kb) contain partial-self complementarity of ~33 nt, a terminal loop and internal runs of uridine residues, to form the stem-loop structure, flanked by a 5'-m^{7}g cap and a 3'-polyadenylated tail which are characteristic of RNA polymerase II transcripts (Bartel, 2004).

The initial processing of pri-miRNA occurs in the nucleus and involves the enzymatic activity of the RNase III endonuclease, Drosha and its dsRNA binding protein partners; DiGeorge Syndrome Critical Region Gene 8 (DGCR8) and Partner of Drosha (Pasha). These proteins assemble to form the Microprocessor complex which cleaves near the base of the stem loop of pri-miRNA to release the ~60-70nt precursor miRNA (pre-miRNA) hairpin-loop intermediate (Lee *et al.,* 2003).

Following nuclear processing, the pre-miRNA is actively exported from the nucleus to the cytoplasm. The specific length and the 3' overhang of the pre-miRNA is recognised by the export receptor Exportin-5 and Ran-GTP and is transported through the nuclear pore (Yi *et*

al., 2003). The pre-miRNA is further processed by a second RNase III endonuclease Dicer and the HIV-transactivating response RNA-binding protein (TRBP). TRBP increases the cleavage efficiency of Dicer while Dicer cleaves the hairpin-loop of the pre-miRNA to release an unstable 19-25nt double stranded miRNA: miRNA* duplex. The mature miRNA serves as the 'guide' strand and is incorporated into the RNA-induced silencing complex (RISC), whilst the miRNA* 'passenger' strand is typically degraded. It has been proposed that miRNA guide strand selection is largely determined by the thermodynamic stability of the miRNA duplex and the strand that is less stably paired at its 5'-terminus is more likely to be selected as the guide strand (Khvorova *et al.,* 2003; Finnegan and Pasquinelli, 2013).

The guide miRNA is subsequently loaded onto a member of the Argonaute (Ago) protein family and the resulting miRNA-Ago complex forms the core subunit of the RISC complex. The Argonaute protein is directed by the mature miRNA to silence the target mRNAs by translational repression or mRNA cleavage; a fate that is dependent on the degree of complementarity between the microRNA seed region and the 3'UTR of the target mRNA. Perfect complementarity between the miRNA seed region and the target mRNA results in Ago-2-catalysed mRNA cleavage, which is commonly observed in plant miRNAs. Similarly, double–stranded RNA molecules of an endogenous or exogenous source, are also processed by Dicer to generate short duplex small interfering RNA, siRNA, which typically silence gene expression by mRNA cleavage due to perfect base pairing to target mRNAs. In contrast, most animal miRNAs bind to their targets with imperfect complementarity due to mismatches and bulges between the miRNA seed region and the target mRNA 3'UTR, resulting in translational repression or deadenylation of the target mRNA leading to destabilisation and mRNA degradation (Bartel, 2004).

Interestingly, the vast majority of animal miRNA target sites are found in the 3'UTR of the target mRNA however, there are examples of the miRNA target sites located in the 5'UTR (Lytle *et al.,* 2007) and coding regions (Forman *et al.,* 2008). Nevertheless, the 6-8nt miRNA seed region is the most important region for mRNA post-transcriptional regulation. For

example, many miRNA family members share very similar or identical seed sequences, while the 3'-end of the miRNA can differ significantly and such similarities allow various miRNAs to regulate overlapping targets and one miRNA can potentially target multiple mRNAs. Therefore, the conservation of the miRNA seed region among species has proven to be an efficient criterion in predicting miRNA targets (Ghosh *et al.,* 2007).

1.4.3. MicroRNA target prediction and validation

The accurate prediction and validation of miRNA targets is the foundation to associating each miRNA to a specific function which can provide insight into the role of miRNAs in biological processes. Computational predictions of miRNA targets are invaluable tools in narrowing down potential miRNA target sites for experimental validation and current computational prediction tools such as miRanda (Enright *et al.,* 2003) and TargetScan (Lewis *et al.,* 2003) apply methods that are both diverse in approach and performance. Nevertheless, the vast majority of these computational algorithms use four common characteristics of miRNAs and their target genes; (i) the seed region complementarity between miRNA and target genes; (ii) the conservation of the miRNA target sites among different species, (iii) the thermal stability of the miRNA/mRNA duplex (iv) the ability of the miRNA 5'-end to bind to the target gene stronger than that of the 3'-end (Peterson *et al.,* 2014). Based on one or more of these principles, computational algorithms predict miRNA target genes which can be experimentally verified. Various experimental approaches have been employed for validation and differential expression analysis of predicted miRNA targets in biological samples.

Earlier miRNA expression studies were commonly performed using northern blot analysis combined with polyacrylamide gels which simultaneously allowed quantification of the expression levels and size determination of the mature and precursor miRNAs (Lagos-Quintana *et al.,* 2001; Lau *et al.,* 2001; Lee and Ambros, 2001). However, the laborious, poor sensitivity and limited sample throughout of northern blot analysis led to the development of

highly-sensitive and high-throughput miRNA expression profiling techniques such as Real-time qPCR, miRNA-microarray analysis, and improved *in situ* hybridisation and northern blotting analysis using locked nucleic acid (LNA)-modified oligonucleotide probes.

MiRNA *in situ* hybridisation is a variation of immunohistochemical staining which involves the hybridisation of a labelled RNA with a complementary strand within suitably prepared tissue. Although this method is not highly quantitative, *in situ* hybridisation enables the detection of miRNAs localised to specific parts of the tissue or specific cells. Similarly, miRNA northern blot analysis also employs labelled RNA to hybridise with complementary sequences in RNA samples after which the RNA bands are separated by gel electrophoresis and then probed. The practical disadvantage of both techniques is the relatively low sensitivity that has since been improved by using LNA-modified oligonucleotide probes. The exceptional hybridisation affinity of LNA-modified probes towards RNA has been exploited in northern blot analysis to demonstrate 10 times more sensitivity than traditional oligonucleotides (Valoczi *et al*, 2004) and high efficiency for miRNA profiling using *in situ* hybridisation (Kloosterman *et al.,* 2006) and high throughput miRNA detection using miRNA-microarrays (Castoldi *et al.,* 2006).

Microarrays are powerful tools for miRNA expression analysis and also employ the principle of probe hybridisation with a complementary strand within the sample. In contrast to northern blot analysis in which each miRNA is detected separately, miRNA-microarray uses a single microarray chip containing thousands of probes e.g. LNA-modified oligonucleotides, to enable a highly efficient screening process to identify the expression of several hundreds of miRNAs in the same sample (Yin *et al.,* 2008). Despite the practical advantages of microarray technology, this method is limited by the requirement of high concentrations of sample for efficient hybridisation and signal generation and the poor sensitivity for rare targets, therefore microarray validation is often coupled with a more sensitive technique such as Real-time qPCR (Schmittgen *et al.,* 2004). Real-time qPCR is arguably the most commonly used miRNA expression profiling technique due to its high sensitivity, efficiency and its highly quantitative and high-throughput screening of miRNA expression in small amounts of sample. The

combination of bioinformatic target prediction and experimental validation studies have enabled the identification of miRNAs which exhibit specific expression profiles in a range of cell types. MiRNAs are no longer regarded as biological anomalies and there is substantial evidence implicating miRNAs in the regulation of an array of mRNAs involved in various physiological and pathological conditions.

1.4.4. MicroRNAs in RA

MiRNAs are important regulators of gene expression in many biological processes and therefore, changes in miRNA expression patterns can promote the development and/or exacerbate pathologies. Numerous miRNAs exhibit deregulated expression in RA, many of which play important roles in RASFs and therefore, the pathogenesis of RA.

1.4.4.1. Overexpressed miRNAs in RA

Taganov *et al.,* (2006) reported the induced expression of miR-146a in response to a variety of microbial components and pro-inflammatory cytokines e.g. LPS, CpG, IFNα and TNFα in a human monocyte cell line (THP-1 cells) in an NFκB-dependent manner. It was proposed that miR-146a might negatively regulate cytokine signalling by downregulating its mRNA targets encoding TNF receptor-associated factor 6 (TRAF6) and IL-1 receptor-associated kinase 1 (IRAK1), which are key molecules downstream of TNFα and IL-1β signalling (Taganov *et al.,* 2006). Interestingly, a specific polymorphism in the 3'UTR of IRAK1 is associated with increased susceptibility to developing RA, which opens the possibility to investigate how miRNAs and polymorphisms in their target mRNAs can lead to RA in specific ethnic groups and in association with environmental factors (Chatzikyriakidou *et al.,* 2010).

The upregulated expression of miR-146a is not confined to immune cells and has also been observed in the synovial fibroblasts and synovial tissue isolated from patients with RA

compared to OA patients and miR-146a was induced by TNFα and IL-1β, suggesting that miR-146a mRNA is expressed in synovial fibroblasts in response to these pro-inflammatory cytokines. It was also revealed that miR-146a-expressing cells in the synovial tissues were predominantly CD68[+] macrophages, CD3[+] T cells and CD79[+] B cells (Nakasa *et al.,* 2008).

Similarly, Stanczyk *et al.,* (2008) reported the constitutive expression of miR-146a and also miR-155 in RASFs and RA synovial tissue compared to OA synovial fibroblasts and tissues and miR-155 expression in *ex vivo* RASFs was further induced by TNFα, IL-1β, LPS, bacterial lipoprotein and poly (I-C) stimulation. Interestingly, the overexpression of miR-155 in RASFs was associated with downregulated expression of MMP-3 and reduced induction of MMP-3 and MMP-1 by TLR ligands and cytokines, suggesting the role of miR-155 in modulating the destructive properties of RASFs (Stanczyk *et al.,* 2008).

Evidently, miR-155 and miR-146a have proven to play important roles in various inflammatory responses but appear to have opposing impacts in RA. These observations were supported by *in vivo* studies. For example Nakasa *et al.,* (2011) showed that the administration of the inhibitory double-stranded miR-146a prevented joint destruction in CIA mice but did not completely suppress inflammation (Nakasa *et al.,* 2011). Conversely, Bluml *et al.,* (2011) used CIA and K/BxN serum-transfer arthritis in wild type (WT) and miR-155-null mice to examine the role of miR-155 in the pathogenesis of autoimmune arthritis. Their data showed that the miR-155-null mice did not develop CIA or generate pathogenic autoreactive B- and T-cells, compared to WT mice. The miR-155-null mice also had significantly decreased levels of IL-17 and IL-22, which impaired Th17 polarisation of miR-155-null mouse T cells. Furthermore, in the K/BxN serum-transfer arthritis model, which is dependent on innate effector mechanisms, the miR-155-null mice had significantly reduced local bone destruction due to reduced generation of osteoclasts. These studies demonstrated that miR-155 is essential in adaptive and innate immune reactions leading to autoimmune arthritis and may therefore represent a potential therapeutic target for the treatment of RA (Bluml *et al.,* 2011).

Interestingly, Pauley *et al.,* (2008) reported that the expression of miR-16, miR-132, miR-155 and miR-146a was significantly higher in the circulating peripheral blood mononuclear cells (PBMCs) from RA patients compared to healthy control individuals and the high expression of miR-16 and miR-146a correlated with active disease. Importantly, the detection of miRNAs in the easily obtainable peripheral blood cells opens the possibility to identify biomarkers which can be monitored during the disease course without the need for invasive surgical procedures to isolate joint tissues and cells for miRNA analysis and the results of this study suggest that miR-16 and miR-146a might be useful biomarkers of disease activity (Pauley *et al.,* 2008). This observation was confirmed by the studies of Murata *et al.,* (2010) who reported the significantly elevated expression of miR-16, miR-146a, miR-155 and miR-223 in the synovial fluid from RA patients compared to OA control samples and showed a significant correlation of plasma miR-16 and miR-146a with tender joint counts and DAS28 score (Murata *et al.,* 2010). Furthermore, Fulci *et al.,* (2010) showed that miR-223 is overexpressed in peripheral T-lymphocytes from RA patients compared to healthy individuals, particularly naive CD4[+] T-lymphocytes, which are one of the major infiltrating cells in the RA synovial tissues, suggesting that miR-223 could target T-cell responses and therefore contribute to the onset and pathogenesis of RA (Fulci *et al.,* 2010).

The IL-17 pro-inflammatory cytokine is recognised as a crucial factor in inflammation and bone destruction in RA due to its ability to induce the expression of pro-inflammatory cytokines and RANKL. Niimoto *et al.,* (2010) reported the significant upregulation of let-7a, miR-26, miR-146a/b, miR-150 and miR-155 in IL-17-producing CD4[+] T lymphocytes from RA patients and reported the strong expression of miR-146a and IL-17 in PBMCs in patients with high disease activity. Also, miR-146a was intensely expressed in the hyperplastic RA synovium with high expression of IL-17 from RA patients with high disease activity and IL-17-producing cells expressing miR-146a, indicating that miR-146a is associated with IL-17 expression in the PBMC and synovium in RA patients (Niimoto *et al.,* 2010).

Furthermore, Alsaleh *et al.,* (2009) reported the upregulated expression of miR-346 in RASFs and showed that miR-346 was able to indirectly regulate the release of the pro-inflammatory cytokine IL-18 through indirect inhibition of the Bruton's tyrosine kinase in LPS-activated RASFs, which is involved in the miR-346-regulated expression of IL-18 (Alsaleh *et al.,* 2009).

1.4.4.2. Suppressed miRNAs in RA

Downregulated miRNAs have also been reported in RA. For example miR-124, which is involved in cell proliferation, exhibited significantly decreased expression in RASFs as compared with OASFs. This study by Nakamachi et al., (2009) showed that the transfection of miR-124 precursor (pre-miR-124) and therefore the overexpression of miR-124, led to suppressed proliferation of RASFs and arrested cell cycle at the G_1 phase. Furthermore, the two targets of miR-124, CDK2 (cyclin-dependent kinase 2) and MCP-1 (monocyte chemotactic protein-1), which are involved in the inflammatory process of RA, were significantly downregulated by the induction of pre-miR-124a in RASFs, suggesting that miR-124 plays an important role in the regulatory mechanisms of RASFs (Nakamachi et al., 2009).

The interaction of pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs) with pattern recognition receptors such as TLRs, has been shown to activate RASFs. TLR2 expression is strongly upregulated in RASFs in response to TLR2 ligands and TLR2 mRNA is a target of miR-19. Interestingly, miR-19a and miR-19b belong to the same cluster and were found to be significantly downregulated in RASFs stimulated with LPS and BLP (bacterial lipopeptide Pam_3CSK_4). Overexpression of miR-19a/b using synthetic mimics decreased TLR2 protein expression and also IL-6 and MMP-3 secretion was significantly downregulated in activated RASFs transfected with either mimic. These studies demonstrated the important role for miR19a/b in the regulation of IL-6 and MMP-3 release by controlling TLR2 expression, suggesting that miR-19a/b can act as negative regulators of inflammation (Philippe *et al.,* 2012).

The infiltration and accumulation of T cells in the synovium of RA joints is a hallmark pathological feature and Li *et al.,* (2010) reported the elevated expression of miR-146a in CD4$^+$ T cells isolated from RA synovial fluid, which was positively correlated with the levels of TNFα and also the downregulated expression of miR-363 and miR-498 in the same cells (Li *et al.,* 2010).

Recently, Zhu *et al.,* (2012) reported the significantly downregulated expression of miR-23b in inflammatory lesions of patients with RA or SLE as well as in the tissue samples from mouse models of lupus (lupus-prone MRL-*lpr* mice), RA (CIA mice) and multiple sclerosis (experimental autoimmune encephalomyelitis (EAE) mice). Interestingly, IL-17 significantly downregulated miR-23b expression in human primary fibroblast-like synoviocytes (FLS), mouse primary kidney cells and astrocytes. The authors showed that miR-23b suppressed IL-17-, IL-1β-, and TNFα-mediated signalling by directly targeting the 3'UTR of TAK1-binding protein 2 (TAB2), TAB3 and IKKα mRNA, which are known to promote NFκB activation in response to TNFα and IL-1β. Therefore, miR-23b plays an important role in suppressing the pathology in multiple autoimmune diseases by limiting the pro-inflammatory signalling pathways in tissue-resident cells (Zhu, *et al.,* 2012).

1.4.5. Manipulating miRNA activity

MiRNAs are integrally involved in health and disease and there is much evidence suggesting that the replacement or overexpression of miRNAs may be therapeutically beneficial in diseases, including cancer and RA. In general, there are two approaches of miRNA-based therapeutics; miRNA mimicry or miRNA inhibition.

Endogenous miRNAs can be suppressed using decoy targets which inhibit endogenous miRNAs that show gain-of-function in disease. This method involves the cellular introduction of synthetic antisense oligonucleotides, termed antagomiRs or AMO (anti-miRNA oligonucleotides) which are modified miRNA targets designed with full or partial

complementarity to a specific miRNA and act by 'inhibiting the inhibitor'. The antagomiRs bind with high affinity to the active miRNA strand to actively compete with endogenous mRNA targets for binding to the specific miRNA. Since binding of the antagomiR with the miRNA is irreversible, the miRNA:antagomiR duplex is unable to be processed by RISC and/or degraded resulting in silencing of the endogenous miRNA and consequently, increased expression of the target mRNA (Krutzfeldt *et al.,* 2005). Synthetic antagomiRs can be modified to enhance stability against nucleases and resistance to degradation, improve base pairing thermodynamics and improve transfection efficient e.g. by including 2'-*O*-methyl (2'-OMe) modified ribose sugars, terminal phosphorothioates and also, a cholesterol group at the 3'-end and LNA-modifications. The significant disadvantage of synthetic antagomiRs is their transient miRNA knockdown (days to weeks) and as demonstrated by Scherr *et al.,* (2007) long-term inhibition of specific miRNA function can be achieved following the delivery of lentivirally encoded recombinant antagomiRs, highlighting the advantages of miRNA-gene therapy (Scherr *et al.,* 2007). To date, antagomiRs have been successfully used *in vivo* and have also reached the stage of clinical trials in human patients (Hydbring and Badalian-Very, 2013), with particular success in the phase II clinical trials of LNA-modified oligonucleotides designed to inhibit the liver specific miR-122 for the treatment of Hepatitis C virus (Miravirsen, Santaris Pharma A/S).

The success of antagomiR oligonucleotides is generally dependent on delivering a dose sufficient to saturate the cellular pool of the target miRNA, which often requires repeated administration in large doses. However, an alternative gene therapy method developed by Ebert *et al.,* (2007) showed that miRNA target sites expressed at a high level from a vector could also function as competitive inhibitors of the cognate miRNA. This method of inhibiting the action of endogenous miRNAs was achieved through the use of decoy targets called 'miRNA sponges'. MiRNA sponges are artificial transcripts expressed from strong promoters, containing multiple, tandem binding sites to the candidate miRNA, engineered into their 3'-UTR. Following the delivery of the vector into the cells, the 'sponge' transcript containing the

miRNA target sites accumulate in the cell and deplete the endogenous miRNAs by acting as competitive inhibitors of the corresponding miRNAs thereby preventing miRNA interaction with their natural mRNA targets (Ebert et al., 2007). The level of miRNA suppression is dependent on the number of target sites in the transcript 3'-UTR and sponges with more copies of miRNA target sites have a greater affinity for their cognate miRNA which increases the overall target miRNA suppression. This effect can be enhanced by designing the miRNA binding sites with imperfect complementarity to the miRNA to enable saturation of the candidate miRNA at lower concentrations compared to perfectly complementary sites, which can be attributed to the different mechanisms of miRNA regulation (Brown and Naldini, 2009). The interaction of miRNAs with their target mRNAs is largely dependent on complementary base pairing between the miRNA seed region and mRNA 3'UTR, therefore, recombinant sponges can be engineered with binding sites complementary to the seed sequences of multiple miRNAs, allowing a single sponge to repress a whole family of related miRNAs, unlike antagomiRs which appear to be specific for one miRNA (Ebert et al., 2007). In addition to the number of miRNA target sites in the sponge, the efficacy of miRNA sponges is also dependent on the concentration of endogenous miRNA relative to the sponge miRNA. Maximal sponge expression can be achieved through the use of a strong promoter and lentiviral vectors to enable stable chromosomal insertions of the sponge and therefore, prolonged expression without the need for repeated administration (Ebert and Sharp, 2010). The use of miRNA sponges and antagomiRs in the induction of miRNA loss-of-function are attractive alternatives to genetic knockouts and provide a versatile approach to investigating miRNA biology in a broad range of cells as well as harbouring therapeutic potential (Gentner and Naldini, 2012).

Conversely, some diseases may be caused by the loss or reduced expression of a specific miRNA, therefore the replacement of a specific miRNA may be therapeutically relevant. MiRNAs which exhibit loss-of-function in disease can be replaced by introducing enzymatically generated or chemically synthesised miRNAs into the cells, thus increasing the concentration of the candidate miRNA in the cell to induce post-transcriptional silencing of the upregulated

mRNA (Amarzguioui *et al.,* 2005). The introduced double-stranded RNA (dsRNA) mimetics are similar in structure to siRNAs and essentially 'mimic' the endogenous activity of the specific miRNA to enhance regulation of the target mRNA. Exogenous miRNA mimics are often administered as individual dsRNA oligonucleotides however, as with antagomiRs for miRNA inhibition, the oligonucleotides require modifications such as the introduction of 2'-OMe and LNA, for improved stability and activity (Czauderna *et al.,* 2003). However, transient expression and the requirement for multiple administration of the oligonucleotide to maintain the therapeutic effects are drawbacks associated with the delivery of miRNA mimic oligonucleotides and can be improved by expressing the short hairpin RNA (shRNA) from a strong promoter in a vector-based system, although limitations are also imposed by the choice of vector and promoter. Exogenous miRNAs expressed from promoters within vectors as a shRNA can mimic pre-miRNAs and are transported to the cytoplasm and processed by Dicer to yield the functional miRNA that target specific mRNAs (Amarzguioui *et al.,* 2005). The miRNA mimics are practically indistinguishable from endogenous miRNAs and therefore, unlikely to cause toxicity.

1.4.6. Exploiting endogenous miRNAs for regulated gene therapy

Combining miRNA regulation with gene therapy allows regulation of gene expression at the post-transcriptional level. Substantial progress has been made in exploiting the endogenous activity and distinct expression profiles of miRNAs to achieve restricted gene expression in a desired cell type, tissue, developmental stage or in response to specific stimuli, while minimising gene expression elsewhere (Gentner and Naldini, 2012). This approach employs a similar vector design to miRNA sponges but produces the opposite outcome; the perfectly complementary candidate miRNA target sites are incorporated into the 3'UTR of a therapeutic transgene, thereby subjecting the transgene to post-transcriptional degradation by the specific miRNA and preventing gene expression in cells and/or environments that express high levels

of the corresponding miRNA. Therefore, the transgene will be expressed in a particular cell type and/or environment in which the miRNA is not expressed (Broderick and Zamore, 2011). The pioneering research of Brown *et al.,* (2006) provided proof-of-principle of miRNA-mediated regulation by gene transfer. This concept was developed on the observation that immune responses were elicited against the transgene product and vector, despite the use of a hepatocyte-specific promoter which was designed to restrict transgene expression to hepatocytes and prevent expression within other cells, especially in antigen-presenting cells, to induce immune tolerance. Therefore, the authors engineered a lentiviral vector with four perfectly complementary target sites of the haematopoietic cell-specific miR-142-3p into the 3'UTR of the GFP reporter gene and showed a 100-fold reduction in GFP expression in haematopoietic cells, particularly monocytes and dendritic cells (i.e. antigen presenting cells (APCs), whilst retaining high GFP expression in non-haematopoietic cells (Brown *et al.,* 2006). As expected, the i.v. injection into haemophilia B mice (factor IX knockout) with the lentiviral vector expressing the human coagulation factor IX gene under the control of a hepatocyte-specific promoter, triggered an anti-factor IX immune response and the clearance of transduced hepatocytes (Brown *et al.,* 2007c). However, following the systemic delivery of the hepatocyte-specific promoter combined with four miR-142-3p target sites in the 3'UTR of the human factor IX gene within the lentiviral vector, the gene expression was confined to the hepatocytes resulting in the stable correction of haemophilia B mice without the induction of immune-mediated vector clearance. Furthermore, this strategy conferred human factor IX immune tolerance and the expression of miR-142-3p target did not cause saturation of miR-142-3p or loss of endogenous miR-142-3p function, which are vital concerns about the safety of this approach (Brown *et al.,* 2007b).

MiRNAs have also been exploited for stem cell gene transfer, for example, Papapetrou *et al.,* (2009) generated lentiviral-encoded antigen receptors with miR-181 target sites, which is highly expressed in thymocytes. The vectors were introduce into mouse haematopoietic stem cells (HSCs) and the expression of the antigen receptors was suppressed in late thymocytes,

thereby enabling the cells to avoid clonal deletion in the thymus. Consequently, receptor expression was fully restored in post-thymic T cells to confer protection against a subsequent challenge with antigen-expressing hCD19[+] tumours, thereby demonstrating its potential application in cancer immunotherapy (Papapetrou *et al.,* 2009).

MiRNA-mediated regulation of gene expression has also demonstrated promising results in cancer gene therapy where many groups have exploited the differential miRNA expression between normal and cancer cells to improve tumour targeting. Similar to the strategy described above, the vector containing target sites complementary to a highly expressed miRNA in normal tissue but low or absent in tumour cells, can induce high gene expression in tumour cells with minimal expression in normal cells due to the abundance of the corresponding miRNA. For example, the studies of Suzuki *et al.* (2008) were conducted on the observation that intratumoural injection of adenoviral vectors often disseminate into the systemic circulation and transduce the liver resulting in severe hepatotoxicity. Therefore to overcome this problem, the authors incorporated four perfectly complementary target sites to the liver-specific miR-122 into the 3'UTR of the suicide gene. Following the intratumoural delivery of the vector, the suicide gene was highly expressed in the tumour cells to achieve significant anti-tumour effects and dramatically reduced HSV-thymidine kinase-ganciclovir-induced hepatotoxicity, thereby demonstrating a safe and efficient suicide gene therapy strategy (Suzuki *et al.,* 2008).

MiRNA-mediated concepts have been successfully applied to other platforms including the redirected oncolytic virus tropism (Kelly *et al.,* 2008) and reducing the toxicity of viral vaccines (Barnes *et al.,* 2008). Taken together, the exploitation of endogenous miRNAs holds great promise for efficient gene regulation to achieve targeted and stable gene expression. However, the success of this approach relies on the incorporation of multiple, tandem and perfectly complementary target sites, differential miRNA expression in the cell/tissue/environment under study and importantly, a high threshold of miRNA concentration (Brown *et al.,* 2007b). MiRNA targeting could potentially be harnessed to regulate therapeutic

gene expression in response to inflammation as a novel approach to achieve inflammation-regulated gene expression by exploiting inflammation-repressed miRNAs in the diseased tissue, i.e. miR-23b in RASFs.

1.5. Thesis Aims and Objectives

The underlying aim of this study is to create a novel gene therapy lentiviral vector which exploits the endogenous activity of differentially expressed transcription factors and miR-23b to transcriptionally and post-transcriptionally regulate gene expression, respectively, during inflammation and/or hypoxia. The objectives are:

1.5.1. Transcriptional regulation of gene expression

1. To construct and screen plasmid DNA with inflammation-inducible composite promoters

The endogenous activities of the upregulated candidate transcription factors (AP-1, C/EBPβ, Egr-1, Ets-1, HIF-1α and NFκB) will be exploited by constructing novel composite synthetic promoters, comprised of their cognate binding sites positioned upstream of a minimal CMV promoter and the firefly (*Photinus pyralis*) luciferase gene, using random ligation and Assembly PCR cloning techniques.

I hypothesise that the inflammation-inducible composite promoters will be multi-responsive to respective stimuli and induce synergistic luciferase gene expression during combined inflammatory and hypoxic stimulation, as a result of simultaneous binding of the respective inducible transcription factors. High-throughput transfection of the plasmid DNA constructs into 293T cells will enable the identification of the best synthetic promoters for further analysis, as determined by luciferase quantification.

2. To construct and screen lentiviral inflammation-inducible composite promoters and assess their activity *in vivo*

Lentiviral vectors possess attractive features for experimental gene therapy i.e. the ability to integrate into the host genome with low immunogenicity and allow long-term and stable gene expression. Therefore, the selected composite promoters will be cloned into lentiviral vectors and the resulting lentiviral particles will firstly be used to generate stable cell lines: assessment of composite promoter kinetics will further refine the selection of suitable promoters for analysis *in vivo*. The lentiviral particles of the selected composite promoters will also be delivered into the paws of mice with carrageenan-induced paw inflammation. Bioluminescence imaging will be used to monitor the changes in luciferase gene expression during inflammation and the best composite promoter will be used to transcriptionally-regulate therapeutic gene expression *in vivo*.

3. Assessment of the therapeutic efficacy of mTNFR-Fc and IL-1Ra genes expressed from lentiviral inflammation-inducible composite promoters

The luciferase gene in the selected composite promoter will be replaced with mTNFR-Fc and IL-1Ra therapeutic genes, in individual constructs. Therapeutic protein expression from lentivirally-transduced cells will be quantified using ELISA. The therapeutic efficacy of transcriptionally-regulated mTNFR-Fc and IL-1Ra expression will be assessed following the local delivery of the candidate lentiviral particles into the paws of mice with carrageenan-induced paw inflammation.

1.5.2. Post-transcriptional regulation of gene expression

1. To construct a novel inflammation-responsive miR-23b-regulated system

The inflammation-repressed expression profile of miR-23b in NIH3T3 mouse embryonic fibroblasts will be confirmed using Real-time qPCR. Luciferase gene expression will be post-transcriptionally regulated by cloning the binding sites of miR-23b into the 3'UTR of the luciferase gene, positioned downstream of constitutive CMV and SFFV promoters in plasmid DNA and lentiviral constructs, respectively.

I hypothesise that luciferase gene expression will be highest during the inflamed state (when miR-23b activity is repressed) and significantly downregulated during the uninflamed state (when miR-23b activity is highest), in transiently transfected and lentivirally-transduced NIH3T3 mouse embryonic fibroblasts.

1.5.3. Dual regulation of gene expression

1. To construct a novel transcriptionally and post-transcriptionally regulated system

The binding sites of the inflammation-repressed miR-23b will be cloned into the 3'UTR of the luciferase gene, positioned downstream of an inflammation-inducible NFκB-responsive synthetic promoter, in a plasmid DNA construct.

I hypothesise that luciferase gene expression will be highest during the inflamed state (when NFκB is activated and miR-23b activity is repressed) and significantly downregulated during the uninflamed state (when NFκB is inactivated and miR-23b activity is highest) in NIH3T3 mouse embryonic fibroblasts.

CHAPTER 2:

Materials and Methods

2.1. Tissue culture

2.1.1. Culture of adherent cell lines

Human embryonic kidney 293T cells and mouse embryonic fibroblast NIH3T3 cells were grown in complete Dulbecco's Modified Eagle's Medium (DMEM) (Lonza Group Ltd., Switzerland) containing 10% heat-inactivated foetal bovine serum (Hi-FBS; Gibco, Life Technologies Corp, California, USA), 4.5 g/L glucose (Sigma-Aldrich Corp, St Louis, MO, USA), 100 IU/ml penicillin (Sigma-Aldrich), 100 µg/ml streptomycin (Sigma-Aldrich) and 2 mM L-glutamine (Sigma-Aldrich). Cells were maintained at 37°C in a humidified 10% CO_2 incubator. Growth medium was changed every three days and cells were trypsinised using Trypsin-Versene® (Trypsin-EDTA; Lonza Ltd) when 80-90% confluent.

2.1.2. Long term storage of cell lines

Following trypsinisation of a 70-80% confluent cell monolayer, the cells were pelleted by centrifugation at 1200 rpm for 4 minutes at room temperature. The cell pellet was resuspended in 3 ml freezing medium, comprised of 90% Hi-FBS and 10% dimethylsulfoxide (DMSO; Sigma-Aldrich) and stored in cryovials in 1 ml aliquots. The cells were frozen slowly by initialling storing the cells at -70°C overnight and then transferring the cells to liquid nitrogen for long-term storage. Frozen cells were re-cultured by rapidly thawing the cells in a water bath at 37°C, with gentle agitation. The DMSO was removed from the cells by adding 9 ml complete DMEM medium to the thawed cells and centrifuging the tubes at 1200 rpm for 4 minutes at room temperature. The cell pellet was resuspending in 2 ml complete DMEM medium and transferred to a 75 cm^2 flask containing 10 ml complete DMEM medium.

2.1.3. Routine mammalian cell transfections using FuGENE® 6 Transfection Reagent

To minimise experimental variability, mammalian cells were routinely co-transfected in 96-well tissue culture plates with a total concentration of 200 ng/well DNA (180 ng firefly expressing recombinant plasmid DNA and 20 ng renilla expressing plasmid pRL-CMV (Promega Corp., Madison, WI) using FuGENE® 6 Transfection Reagent (Promega), unless stated otherwise. The day before transfection, 293T or NIH3T3 cells were trypsinised and the live cells were counted using 0.4% Trypan-Blue (Sigma-Aldrich) exclusion on a haemocytometer and $2x10^4$ cells/well were seeded in 96-well tissue culture plates in 100 μl of complete DMEM medium. Briefly, 0.6 μl FuGENE® was diluted in serum-free DMEM medium, vortexed for one second and incubated at room temperature for 5 minutes. A total concentration of 200 ng/well DNA was added to the diluted FuGENE® and incubated at room temperature for 15 minutes to allow the formation of the FuGENE®: DNA complex. Following incubation, 6 μl of the FuGENE®: DNA complex was added to the cells, in a 'drop-wise manner'. Triplicate transfections were performed in each experiment, unless stated otherwise. Twenty-four hours post-transfection, the cells were stimulated with various stimuli.

2.1.4. Mammalian cell stimulation

The choices of stimuli were collated from published studies (Chapter 1). Following 24 hours post-transfection, the cell medium was removed from the transfected cells and replaced with 100 μl of 0.5% FBS DMEM alone for the unstimulated control cells. For the activation of HIF-1α, 100 μl of 0.5% FBS DMEM alone was added to the cells and the plate was incubated in hypoxic conditions at 0.1% O_2, 10% CO_2 and 89.9% N_2 in a CO_2 incubator (New Brunswick, an Eppendorf company, Hamburg, Germany) for 18 hours. For the activation of inflammation-responsive transcription factors and/or all transcription factors, the cells were treated with 100 μl of 0.5% FBS DMEM containing 10 ng/ml human TNFα (30001A; Peprotech, NJ, USA), 10 ng/ml PMA (P1585; Sigma-Aldrich) or a combination of 10 ng/ml TNFα and 10 ng/ml PMA in 0.1% hypoxia for 18 hours.

For the downregulation of miR-23b expression, the cells were treated with 100 µl of 0.5% FBS DMEM containing 10 ng/ml TNFα or 50 ng/ml human or mouse IL-17A (hIL-17A #8928SC, mIL-17A #5227SC; Cell Signalling Technology Inc, MA, USA) or a combination of 10 ng/ml TNFα and 50 ng/ml hIL-17A/mIL-17A for 18 hours.

2.2. Quantification of protein expression

2.2.1. Firefly luciferase assay

Post-stimulation (18 hours), the cell medium was removed and the cells were lysed in 50 µl 1x Passive Lysis Buffer (Promega) and the firefly luciferase expression was quantified using the Luciferase Assay System (Promega) following the manufacturer's instructions, with minor modifications. The cell debris was pelleted by centrifugation at 3,000 rpm for 5 minutes and 10 µl of the cell lysate was transferred into a 96-well opaque luminometer plate (Costar, Corning Inc, Corning, New York, USA). The intracellular firefly luciferase expression in the cell lysates was quantified by adding 50 µl of Luciferase Assay Reagent (Promega), which was preloaded into the luminometer (MLX Microtiter® Plate Luminometer, Dynex Technologies Inc, Chantilly, VA, USA), to the lysates. The Revelation Software (Dynex Technologies Inc) was programmed to perform a 2 second measurement delay and then a 10 second measurement read for luciferase activity, which was expressed as relative light units (RLU). After use, the luminometer tubing was thoroughly washed with water and 70% ethanol.

2.2.2. Renilla luciferase assay

The intracellular renilla luciferase expression was quantified using the Renilla Luciferase Assay System (Promega) following the manufacturer's instructions, with minor modifications. Briefly, 10 µl of the cell lysate was transferred into a 96-well opaque luminometer plate and 20 µl of Renilla Luciferase Assay Reagent (Promega) was dispensed into each well from the

preloaded luminometer. Renilla luciferase expression was quantified using the same method as firefly luciferase expression quantification (section 2.2.1). After use, the luminometer tubing was thoroughly washed with water and 70% ethanol. Firefly luciferase values were normalised to renilla luciferase values in the same sample, to minimise experimental variability, unless otherwise stated.

2.2.3. BCA protein assay

The total protein concentration in the lysates of transfected and transduced stable cells was quantified using the Pierce® BCA Protein Assay Kit (Thermo Scientific, Rockford, IL, USA) following the manufacturer's protocol. The protein concentration was measured from the bovine serum albumin (BSA) standard curve which ranged from 2 mg/ml - 0.03 mg/ml BSA in sterile distilled H_2O (dH_2O). In duplicate, 10 µl of each BSA serial dilution or 10 µl of cell lysate was transferred into each well of a 96-well microtiter plate. A volume of 200 µl of Working Reagent (50 parts BCA Reagent A and 1 part BCA Reagent B) was added to each well and incubated at 37°C for 30 minutes. The measurement of absorbance at 562 nm was performed using a Tecan Genios microplate reader (Tecan Group Ltd, Mannedorf, Switzerland) with Magellan 4 software. The BSA standard curve was plotted as the BSA standards concentration (mg/ml) versus absorbance at OD_{562nm}, using GraphPad® Prism 5 (GraphPad® Software Inc, California, USA). The linear equation obtained from the BSA standard curve was used to calculate the protein concentration of the lysate samples. Firefly luciferase values were normalised to the total protein concentration in the same samples and expressed as relative light units per milligram of protein (RLU/mg), unless stated otherwise.

2.3. Bacterial manipulation

2.3.1. Preparation of chemically-competent DH5α and GT115 *E.Coli* cells

Escherichia coli (E.coli) DH5α and GT115 strains were treated with calcium chloride to produce chemically-competent bacterial cells. Miller Luria Bertani (LB) media (10 g Tryptone, 10 g NaCl and 5 g Yeast Extract in 1L sterile dH_2O, pH 7.5) and low salt Lennox LB media (10 g Tryptone, 5 g NaCl and 5 g Yeast Extract in 1L sterile dH_2O, pH 7.5) was used to grow DH5α and GT115 cells respectively.

Frozen glycerol stocks (stored at -80°C) of DH5α or GT115 bacteria were streaked on antibiotic-free LB-agar plates (10 g tryptone, 10 g NaCl, 5 g yeast extract and 15 g agar in 1L sterile dH_2O, pH 7.5) and low-salt LB-agar plates (10 g tryptone, 5 g NaCl, 5 g yeast extract and 15 g agar in 1L sterile dH_2O, pH 7.5) respectively and the plates were incubated at 37°C overnight. A single colony of DH5α or GT115 bacteria was inoculated directly into   5 ml antibiotic-free LB media or low-salt LB media, respectively and incubated overnight at 37°C with shaking at 225 rpm. The following day, the 5 ml DH5a and GT115 inoculums were transferred into 500 ml of LB media and low-salt LB media, respectively and incubated at 37°C with agitation at 225 rpm until the OD_{600} reached 0.4-0.6. The culture was transferred into two 250 ml centrifuge tubes and centrifuged at 4°C for 10 minutes at 5,000 rpm. The LB supernatant was discarded and the bacterial pellet was resuspended in 50 ml of cold 0.1M $MgCl_2$ and centrifuged at 4°C for 10 minutes at 2,000 rpm. The supernatant was discarded and the bacterial pellet was resuspended in 50 ml of cold 0.1M $CaCl_2$ and incubated on ice for 20 minutes and then centrifuged at 4°C for 10 minutes at 2,000 rpm. The supernatant was discarded and the bacterial pellet was resuspended in 12.5 ml of cold 0.1M $CaCl_2$ containing 14% glycerol (v/v) and stored immediately at -80°C in 100 µl aliquots.

2.3.2. Ligation reactions and transformation of chemically-competent bacterial cells

Ligation of DNA fragments was performed using a vector to insert ratio of 1:3 (molar) or an excess of x3 insert per molar of vector DNA. The reaction volume of 20µl contained the plasmid DNA vector, DNA insert, 2 µl 10x T4 DNA Ligase Buffer (New England Biolabs, NEB Inc., Ipswich, UK) and 1 µl T4 DNA Ligase (400,000 U/ml, NEB). Ligation reactions were incubated overnight at 4°C and transformed into chemically-competent DH5α or GT115 cells using the heat-shock method. Recombinant pGL3 plasmids containing the f1 origin of replication (ori) were transformed into DH5α cells. In contrast, recombinant pCpG plasmids, containing the R6Kγ ori were transformed in GT115 cells, which are a *pir* mutant *E.coli* strain also deficient in *Dcm* methylation.

An aliquot of 100 µl of chemically-competent *E.coli* bacterial cells was thawed on ice and incubated with the 20 µl ligation reaction on ice for 30 minutes. The bacterial cells were heat-shocked at 42°C for 90 seconds and immediately incubated on ice for 2 minutes. The bacterial cells were added to 5 ml of antibiotic-free LB media and low-salt LB media, for the DH5α and GT115 bacterial transformations, respectively. The tubes were incubated at 37°C with shaking at 225 rpm for 90 minutes and 400 µl of the DH5α and GT115 bacterial cultures were streaked on LB-agar plates containing 100 µg/ml Carbenicillin (Carbenicillin Direct, UK) and low-salt LB-agar plates containing 25 µg/ml Zeocin (Invitrogen Corp., Paisley, UK) respectively and incubated overnight at 37°C.

2.3.3. Amplification and purification of miniprep plasmid DNA

Single DH5α or GT115 transformed bacterial colonies were picked from agar plates using sterile 10 µl pipette tips and selectively grown in 5 ml of LB media containing 100 µg/ml Carbenicillin or 5 ml of low-salt LB media containing 25 µg/ml Zeocin, respectively and incubated overnight at 37°C with shaking at 225 rpm. Miniprep plasmid DNA was purified from bacteria using the PureLink® Quick Plasmid Miniprep Kit (Invitrogen) following the

manufacturer's instructions. DH5α and GT115 bacterial cells were pelleted in 1.6 ml Eppendorf tubes by centrifuging the tubes at room temperature at 12,000 g for 4 minutes and resuspending the pellet in 250 µl Resuspension Buffer (containing RNase A). The resuspended bacterial cells were lysed with 250 µl Lysis Buffer and mixed by inverting the tube 6 times and incubating at room temperature for 5 minutes. The cell debris and proteins were precipitated and removed from the lysates by adding 350 µl Precipitation Buffer to the lysates and centrifuging the samples at 12,000 g for 10 minutes at room temperature. The supernatant was loaded onto the Spin Columns, placed inside a 2 ml Wash Tube and centrifuged at 12,000 g for 1 minute at room temperature. The flow-through was discarded and the column was washed with 500 µl Wash Buffer W10 and centrifuged at 12,000 g for 1 minute at room temperature. The flow-through was discarded and the column was further washed with 700 µl Wash Buffer W9 and centrifuged at 12,000 g for 1 minute at room temperature. Residual wash buffer was removed from the column by centrifugation at 12,000 g for 2 minutes at room temperature. Plasmid DNA was eluted from the column by adding 75 µl of TE Buffer (10 mM Tris HCl and EDTA 1 mM, pH 8.0), preheated to 65°C, to the column, incubating at room temperature for 1 minute and centrifuging at 12,000 g for 2 minutes at room temperature. The correct construction of recombinant plasmid DNA and also the integrity of the DNA was analysed by restriction enzyme digestion and/or DNA sequencing. Purified miniprep DNA was stored at -20°C.

2.3.4. Amplification and purification of maxiprep plasmid DNA

The overnight DH5α or GT115 miniprep bacterial cultures (50 µl) were amplified in 250 ml of LB media containing 100 µg/ml Carbenicillin or 250 ml of low-salt LB media containing 25 µg/ml Zeocin, respectively. The cultures were incubated overnight at 37°C with shaking at 225 rpm and the maxiprep DNA was purified from the bacteria using the PureLink® HiPure Plasmid DNA Purification Kit (Invitrogen) according to the manufacturer's guidelines. After 18-20 hours, the bacterial cells were harvested by centrifugation at 4,000 g for 10 minutes at 4°C in a

Beckman XL-90 Ultracentrifuge (Beckman Coulter, CA, USA) and then resuspended in 10 ml Resuspension Buffer (containing RNase A). The bacterial cells were lysed with 10 ml Lysis Buffer, mixed and then incubated at room temperature for 5 minutes. A precipitate was then formed by adding 10 ml Precipitation Buffer to the lysate and immediately mixed to obtain a homogenous solution. The cell debris and proteins were removed from the lysate by centrifugation at 6,000 g for 45 minutes at 4°C (Beckman XL-90 Ultracentrifuge) and the supernatant was loaded onto the HiPure Maxi Column (which had been pre-equilibrated with 30ml Equilibration Buffer) and the solution was allowed to drain from the column by gravity flow. The column was washed with 60 ml Wash Buffer which flowed by gravity flow and the flow-through fraction was discarded. The purified DNA was eluted from the column by adding 15 ml Elution Buffer to the column, in a sterile 50 ml Falcon Tube (Corning Inc.). The DNA was precipitated by adding 10.5 ml Isopropanol to the eluate, thoroughly mixed, stored at -20°C for 60 minutes and then centrifuged at 4,000 g for 60 minutes at 4°C in an Eppendorf 5810R Centrifuge. The supernatant was discarded and the DNA pellet was washed with 5 ml 70% Ethanol at room-temperature and centrifuged at 4,000 g for 30 minutes at 4°C (Eppendorf 5810R Centrifuge). The supernatant was removed and the DNA pellet was air-dried for 5 minutes and resuspended in 500 µl TE Buffer. Purified maxiprep DNA was stored at -20°C and the integrity of the DNA was analysed by restriction enzyme digestion and/or DNA sequencing.

2.3.5. Quantification of nucleic acids

The quantity of nucleic acids (plasmid DNA, PCR products and/or RNA) was calculated using the NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA) by measuring the absorbance of light with a wavelength of 260 nm (A_{260}). An absorption of 1 at A_{260} equals a concentration of 50 µg/ml double-stranded DNA.

2.4. DNA analysis and cloning

2.4.1. Analytical restriction enzyme digestion

Approximately 1 µg plasmid DNA was routinely digested with restriction enzymes (NEB) to verify and/or identify correctly constructed recombinant plasmid DNA. The reaction comprised 1 µg plasmid DNA, restriction enzyme(s), compatible 10x NEB Buffer and sterile dH_2O in a final volume of 20 µl. The volume of enzyme used in the reactions varied depending on the total volume of the reaction but never exceeded 10% (v/v) to prevent star activity of the restriction enzyme. Reactions were incubated for 1-2 hours at the optimal temperature of the restriction enzyme in the compatible 10x NEB buffer, as recommended by NEB. DNA digestion was verified by agarose gel electrophoresis.

2.4.2. Standard agarose gel electrophoresis

The integrity of RNA and/or DNA was verified by horizontal gel electrophoresis on a 0.8-2% (w/v) agarose gel in 0.5 x TAE buffer (National Diagnostics, GA, USA), depending on the size of the fragment. The gels were prepared by dissolving agarose in 0.5 x TAE buffer (0.02 M Tris-acetate, 0.5 mM EDTA, pH 8.3) by boiling in a microwave oven. Once cooled, ethidium bromide (10 mg/ml in H_2O; BDH Ltd, UK) or SYBR® Safe (10,000X concentrate in DMSO, Invitrogen) was added to obtain a final concentration of 1 µg/ml ethidium bromide or 1x SYBR® Safe, for visualisation of DNA bands. The gel was cast into a mould and combs were inserted horizontally into the gel, which was set at room temperature. A final concentration of 1x Loading dye (8 ml glycerol, 500 µl 10x TBE buffer (890 mM Tris-borate, 890 mM boric acid, 20 mM EDTA, pH 8.3), 0.025 g bromophenol blue, 0.025 g xylene cyanol and sterile dH_2O to 10 ml total volume) was mixed with the RNA or DNA samples before loading onto the agarose gels. A 1Kb+ or 50 bp DNA ladder (Invitrogen) was simultaneously loaded alongside the DNA samples to provide size markers. Gels were electrophoresed using the voltage and duration required for the separation of the required DNA/RNA bands, which was typically at 90V for 45-

90 minutes, and then visualised by exposure to ultra-violet (UV) light at 570-640 nm using a CCD camera on a gel documentation system (UVP BioDoc-It® Imaging System; UVP LLC, CA, USA). The image was captured using the adjoining Sony Digital Graphic Printer (UP-D897; Sony Corp).

2.4.3. Gel extraction of DNA fragments

DNA fragments excised from agarose gels were purified from agarose using the PureLink® Quick Gel Extraction kit (Invitrogen) according to the manufacturer's instructions. The band of interest was excised from the agarose gel using a sterile scalpel blade under UV light at 570-640 nm using a UV transilluminator (Syngene, Cambridge, UK) and the gel slice transferred into a sterile 1.6 ml Eppendorf tube and weighed. For routinely used ≤2% agarose gels, 1 volume of gel was dissolved in three volumes of Gel Solubilisation Buffer (L3) and incubated at 50°C for a total of 15 minutes, with mixing by regular inversions at 3 minute time-intervals. One gel volume of isopropanol was added to the dissolved gel slice and the mixture was subsequently applied onto the Quick Gel Extraction Column, placed inside a wash tube, and centrifuged at 12,000 g for 1 minute (Eppendorf 5415 microcentrifuge). The flow-through was discarded and 500 µl Wash Buffer (W1) was added to the column and centrifuged at 12,000 g for 1 minute. The flow-through was discarded and the residual wash buffer and ethanol was removed from the column by centrifugation at maximum speed for 2 minutes. The column was placed in a sterile 1.6 ml Eppendorf tube and the DNA was eluted from the column by adding 50 µl Elution Buffer (E5) and incubating the DNA for 1 minute at room temperature following by centrifugation at 12,000 g for 1 minute.

2.4.4. Purification of plasmid DNA and PCR products

Plasmid DNA and PCR products were purified using the PureLink® PCR purification Kit (Invitrogen) according to the manufacturer's instructions. The Binding Buffer B2 was routinely used for DNA and/or PCR product purification, unless the use of Binding Buffer HC was stated. Approximately 4 volumes of the appropriate Binding Buffers (B2 or HC) were mixed with 1 volume of the PCR product or plasmid DNA (50-100 μl) and the sample was applied to the PureLink® Spin Column. The column was placed in a collection tube and centrifuged at room temperature at 10,000 g for 1 minute (Eppendorf 5415 microcentrifuge), after which the flow-through was discarded. The column was washed by adding 650 μl Wash Buffer and centrifugation at room temperature at 10,000 g for 1 minute. The flow-through was discarded and the residual Wash Buffer was removed by centrifugation at 12,000 g at room temperature for 3 minutes. The column was transferred into a sterile 1.6 ml Eppendorf tube and the DNA was eluted from the column by adding 50 μl Elution Buffer and incubating the column at room temperature for 1 minute followed by centrifugation at the 12,000 g for 2 minutes. The recovered DNA was used for subsequent cloning, unless stated otherwise.

2.5. Random Ligation Cloning Method: construction of pGL3-4bp-composite synthetic promoters

2.5.1. Expression construct: pGL3mCMV

The promoter-less pGL3-Basic plasmid (Promega) encoding the luciferase reporter gene (25 μg) was digested with 2.5 μl XhoI (NEB), 2.5 μl HindIII (NEB), 10 μl 10x NEB Buffer 2 and sterile dH$_2$O in a 100 μl reaction mixture. Following an overnight incubation at 37°C, the 4797 bp DNA fragment was isolated by gel extraction using the PureLink® Quick Gel Extraction Kit (Invitrogen) to generate the pGL3-Basic cloning vector. Oligonucleotides containing the minimal region of the human cytomegalovirus promoter (mCMV promoter) from -52 bp to -14 bp of the wild-type promoter (Boshart et al., 1985; GenBank Accession #K03104) were synthesised with 5'-XhoI and 3'-HindIII overhangs by Sigma-Aldrich. The oligonucleotides

(Table 2.1) were resuspended in sterile TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) at a final concentration of 2 mg/ml. The forward and reverse mCMV oligonucleotides were annealed using the boiling method; 2 µl of each mCMV-oligonucleotide was added to 46 µl of TE buffer in a reaction volume of 50 µl and boiled for 5 minutes in a water bath and then allowed to cool to room temperature overnight. The annealed mCMV-oligonucleotides were cloned into the equivalent XhoI/HindIII site of pGL3-Basic vector in a ligation reaction comprising 10 µl annealed mCMV oligonucleotides, 1 µl pGL3-Basic vector, 2 µl 10x T4 DNA Ligase Buffer (NEB), 1 µl T4 DNA Ligase (NEB) and   6 µl H_2O and incubated at 4°C overnight. The resulting construct was referred to as pGL3mCMV (Fig 3.2 A, section 3.2.1). The ligation reactions were used to transform chemically-competent DH5α *E.coli* cells and the plasmid DNA was selectively grown in LB media containing 100 µg/ml Carbenicillin. Large scale plasmid DNA was isolated using the PureLink® HiPure Plasmid DNA Purification Kit (Invitrogen) and the positive recombinant pGL3mCMV constructs were verified by restriction analysis and DNA Sanger sequencing (Genome Centre, Queen Mary, London, UK) using the reverse sequencing primer GL2, which anneals to the 5' end of the luciferase gene and sequences towards the promoter (Appendix 2).

2.5.2. Cloning vector: pGL3mCMV

pGL3mCMV (25 µg) vector was linearised in a reaction comprising 3 µl NheI, 10 µl 10x NEB Buffer 2 and sterile distilled water in a final reaction volume of 100 µl. Following an overnight incubation at 37°C, the linear DNA was separated from uncut and/or partially cut DNA by gel electrophoresis. The linear DNA was gel extracted using the PureLink® Quick Gel Extraction Kit (Invitrogen) and subsequently dephosphorylated using Calf Intestinal Phosphatase (CIP) (NEB) to prevent religation of the linear DNA. Dephosphorylation of the 5'- and 3'- ends of pGL3mCMV required incubating 50 µl pGL3mCMV, 3 µl CIP (NEB), 10 µl 10x NEB Buffer 2 and sterile dH_2O in a final volume of 100 µl for 1 hour at 37°C. A final concentration of   50 mM EDTA was added to the reaction to inhibit the CIP enzyme and the DNA was subsequently

purified using the PureLink® PCR Purification Kit (Invitrogen) to yield 50 µl of the linear pGL3mCMV cloning vector (Fig 3.2 B, section 3.2.1).

2.5.3. Oligonucleotide design for pGL3-4bp-composite synthetic promoters

The oligonucleotides containing the TFBSs for AP-1, C/EBP, Egr-1, Ets-1, HIF-1α and NFκB were designed with phosphorylated 5'-CTAG overhangs complementary to the NheI overhangs. The oligonucleotides were synthesised at Sigma-Aldrich as HPLC purified forward and reverse oligonucleotides (Table 2.1) and each oligo was resuspended at 2 mg/ml in sterile TE buffer and annealed using the boiling method, as previously described in section 2.5.1.

	Oligonucleotide Sequences (5'-3')
Forward XhoI mCMV Oligo	5'- <u>TCGAG</u>GCCTGTAGGCGTGTACGGTGGGAGGCTTATATAAGCAGAGCTCA -3'
Reverse HindIII mCMV Oligo	5'- <u>AGCTT</u>GAGCTCTGCTTATATAAGCCTCCCACCGTACACGCCTACAGGCC -3'
AP-1 Forward Oligo	5'- <u>CTAG</u>**TGAGTCA** -3'
AP-1 Reverse Oligo	5'- <u>CTAG</u>**TGACTCA** -3'
C/EBPβ Forward Oligo	5'- <u>CTAG</u>**ATTGCGCAAT** -3'
C/EBPβ Reverse Oligo	5'- <u>CTAG</u>**ATTGCGCAAT** -3'
Egr-1 Forward Oligo	5'- <u>CTAG</u>**TTGCGTGGGCGT** -3'
Egr-1 Reverse Oligo	5'- <u>CTAG</u>**ACGCCCACGCAA** -3'
Ets-1 Forward Oligo	5'- <u>CTAG</u>**CCGGAAGTTCC** -3'
Ets-1 Reverse Oligo	5'- <u>CTAG</u>**GGAACTTCCGG** -3'
HRE Forward Oligo	5'- <u>CTAG</u>**ACGTGG** -3'
HRE Reverse Oligo	5'- <u>CTAG</u>**CCACGT** -3'
NFκB Forward Oligo	5'- <u>CTAG</u>**GGAATTTC** -3'
NFκB Reverse Oligo	5'- <u>CTAG</u>**GAAATTCC** -3'

Table 2.1. List of TFBS-oligonucleotides for cloning the pGL3-4bp-composite synthetic promoter constructs. Overhangs are underlined

2.5.4. Construction of pGL3-4bp-composite synthetic promoters

The annealed TFBS-oligonucleotides, with phosphorylated 5'-CTAG overhangs, were pooled together and cloned into the equivalent site of the pGL3mCMV vector in a reaction comprising 1 µl of each annealed TFBS-oligonucleotide (n=6), 2 µl pGL3mCMV vector, 2 µl 10x T4 DNA Ligase Buffer, 1 µl T4 DNA ligase and sterile dH$_2$O to a final volume of 20 µl. The ligation was incubated at 4°C overnight and then transformed into chemically-competent DH5α *E.coli* cells and purified using the PureLink® Quick Plasmid Miniprep Kit (Invitrogen). Positive recombinant plasmid DNA constructs from this ligation were referred to as pGL3-4bp-composite synthetic promoters (Fig 3.2 C, section 3.2.1).

2.6. Random Ligation Cloning Method: construction of pCpG-4bp-composite synthetic promoters

2.6.1. Expression construct: pCpGmCMV-Luc[+]

The pCpGmCMV-Luc[+] expression vector was generated by the ligation of two fragments from two different expression vectors; pCpG-mSEAP (InvivoGen, San Diego, CA, USA) and pGL3mCMV (section 2.5.1). The mCMV promoter and luciferase reporter gene was isolated from the pGL3mCMV expression vector by digesting 25 µg pGL3mCMV with 2.5 µl NheI, 2.5 µl AfeI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a final reaction volume of 100 µl. Following an overnight incubation at 37°C, the DNA fragments were separated by agarose gel electrophoresis and the 2143 bp DNA fragment was isolated by gel extraction using PureLink® Quick Gel Extraction Kit (Invitrogen) (Fig 3.8 A, section 3.3.3)

The pCpG-mSEAP plasmid contains two nuclear matrix attachment regions from the 5'-region of the human IFN-β gene and the β-globin gene. This fragment was isolated by digesting 25 µg pCpG-mSEAP with 3 µl SbfI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a 100 µl reaction and incubated overnight at 37°C. The reaction was purified using the PureLink® PCR Purification Kit (Invitrogen) and the DNA was treated with Klenow (DNA polymerase I, Large Klenow fragment, NEB) to generate blunt ended DNA which can be ligated to the AfeI

blunt overhang of the pGL3mCMV DNA. Briefly, 50 µl of purified linear pCpG-mSEAP DNA was incubated with 3 µl Klenow (NEB), 10 µl 10x NEB Buffer 2, 10 µl 2 mM dNTPs and sterile dH$_2$O in a final reaction volume of 100 µl at 25°C for 15 minutes. The reaction was stopped following the addition of a final concentration of 10mM EDTA and the Klenow enzyme was heat-inactivated by incubating the reaction at 75°C for 20 minutes. The DNA was purified using the PureLink® PCR Purification Kit (Invitrogen) and the resulting 50 µl DNA was digested with 2.5 µl NheI, 2.5 µl XbaI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a final reaction volume of 100 µl and incubated at 37°C overnight. The DNA fragments were separated by agarose gel electrophoresis and the 2194 bp fragment was isolated by gel extraction using PureLink® Quick Gel Extraction Kit (Invitrogen) to produce 50 µl of purified pCpG-mSEAP vector (Fig 3.8 B, section 3.3.3).

The pCpGmCMV-Luc$^+$ construct was generated by ligating 6 µl pGL3mCMV fragment,   2 µl pCpG-mSEAP vector, 2 µl 10x T4 DNA Ligase Buffer (NEB), 1 µl T4 DNA ligase (NEB) and sterile dH$_2$O in a ligation reaction of 20 µl. The ligation reaction was incubated at 4°C overnight and then transformed into chemically-competent GT115 *E.coli* bacterial cells. The resulting large scale recombinant pCpGmCMV-Luc$^+$ DNA was purified using the PureLink® HiPure Plasmid DNA Purification Kit (Invitrogen) (Fig 3.8 C, section 3.3.3).

2.6.2. Cloning vector: pCpGmCMV

The pCpGmCMV-Luc$^+$ construct (25 µg) was digested with 2.5 µl NheI, 2.5 µl XhoI, 10 µl 10x NEB Buffer 2 and sterile dH$_2$O in a 100 µl reaction. Following an overnight incubation at 37°C, the DNA fragments were separated by gel electrophoresis and the DNA was purified from the agarose using the PureLink® Quick Gel Extraction Kit (Invitrogen) to yield 50 µl of purified pCpGmCMV vector (Fig 3.8 C, section 3.3.3).

2.6.3. Cloning vector: pCpGmCMV-66bp spacer

Forward and reverse spacer-oligonucleotides with 5'-XbaI, 3'-XhoI overhangs and an internal NheI restriction enzyme site were synthesised by Sigma-Aldrich (Table 2.2) to achieve a 66 bp space between the proximal TFBS and the TATA box within the pCpGmCMV vector. The oligonucleotides were resuspended to 2 mg/ml in sterile TE Buffer and annealed using the boiling method. The spacer oligonucleotides were ligated into the compatible NheI/XhoI site within the pCpGmCMV vector (section 2.6.2) in a ligation which was subsequently transformed into chemically-competent GT115 cells. The resulting maxiprep pCpGmCMV-66bp DNA (25 µg) was linearised with 3 µl NheI (NEB), 10 µl 10x NEB Buffer 2 and sterile dH$_2$O in a 100 µl reaction and incubated overnight at 37°C. The linear DNA was separated by gel electrophoresis and the DNA was purified from the excised agarose using the PureLink® Quick Gel Extraction Kit (Invitrogen). The resulting linear pCpGmCMV-66bp vector was dephosphorylated by adding 50 µl purified pCpGmCMV-66 bp vector to 3 µl CIP (NEB), 10 µl 10x NEB Buffer 2 and sterile dH$_2$O in a final volume of 100 µl and incubating the reaction at 37°C for 1 hour. A final concentration of 50 mM EDTA was added to the reaction to inhibit the CIP enzyme which was then purified using the PureLink® PCR Purification Kit (Invitrogen) to yield 50 µl of the linear pCpGmCMV-66bp vector (Fig 3.18 A and B, section 3.4).

2.6.4. Oligonucleotide design for pCpG-4bp-composite synthetic promoters

The oligonucleotides containing the TFBSs for AP-1, HIF-1α and NFκB were designed with phosphorylated 5'- CTAG overhangs which are complementary to the NheI overhang of the pCpGmCMV-66bp cloning vector. The HPLC purified forward and reverse oligonucleotides were synthesised at Sigma-Aldrich (Table 2.2) and were resuspended at 2 mg/ml in TE buffer and annealed using the boiling method, as previously described in section 2.5.1.

	Oligonucleotide Sequences (5'-3')
AP-1 Forward Oligo	5'- <u>CTAG</u>**TGAGTCA** -3'
AP-1 Reverse Oligo	5'- <u>CTAG</u>**TGACTCA** -3'
HRE Forward Oligo	5'- <u>CTAG</u>**ACGTGC** -3'
HRE Reverse Oligo	5'- <u>CTAG</u>**GCACGT** -3'
NF-κB Forward Oligo	5'- <u>CTAG</u>**GGGACTTTCC** -3'
NF-κB Reverse Oligo	5'- <u>CTAG</u>**GGAAAGTCCC** -3'
Forward Spacer-66bp Oligo	5'- <u>CTAG</u>ACGCGT<u>GCTAGC</u>TCGCGATCTTATGATCTGGATCCATGC-3'
Reverse Spacer-66bp Oligo	5'-<u>TCGAG</u>CATGGATCCAGATCATAAGATCGCGA<u>GCTAGC</u>ACGCGT-3'

Table 2.2. List of TFBS-oligonucleotides used to construct the pCpG-4bp-composite synthetic promoter constructs. Overhangs are underlined and TFBSs are highlighted in bold. The internal NheI restriction enzyme site within the spacer-66bp-oligonucleotides is also underlined.

2.6.5. Construction of pCpG-4bp-composite synthetic promoters

The annealed oligonucleotides, with phosphorylated 5'-CTAG overhangs, were cloned into the compatible site within the NheI digested pCpGmCMV-66bp vector in a ligation reaction comprising equimolar 3 μl AP-1, 3 μl HRE and 3 μl NFκB annealed oligonucleotides, 2 μl pCpGmCMV-66bp vector, 2 μl 10x T4 DNA Ligase Buffer (NEB), 1 μl T4 DNA ligase (NEB), and sterile dH$_2$O to a final volume of 20 μl. The ligation reaction was incubated at 4°C overnight and the resulting library of miniprep recombinant plasmid DNA, which were purified from transformed GT115 positive colonies, were referred to as pCpG-4bp-composite synthetic promoter constructs (Fig 3.18 C, section 3.4).

2.7. Assembly PCR Cloning Method

2.7.1. Design of oligonucleotides for Assembly PCR

The oligonucleotides required for the Assembly PCR reaction were ordered from Sigma-Aldrich and contained an NFκB (variable or fixed), HRE, AP-1 motif or a spacer sequence,

flanked by annealing sequences of 10 bp, 15 bp, 20 bp or 25 bp. The 5'-Stop-NheI and 3'-Stop-XhoI oligonucleotides contained NheI and XhoI restriction enzyme sites, respectively, and annealing sequences of 10 bp, 15 bp or 20 bp. All oligonucleotide sequences were screened using the TRANSFAC® database (Wingender *et al.,* 2000) to verify the absence of other mammalian TFBSs. The desalted oligonucleotides were resuspended at 100 µM in sterile TE buffer and all oligonucleotides used in the Assembly PCR reactions described in this thesis are listed in Table 2.3.

	DNA Sequence (5'-3')
Variable NFκB- 30 bp Forward primer	5'-ATCTCTGCGATGAACCTCACCATGT**GGGRNNYYCC**ACAAGGTGCCTCTTATGATCTGGAT-3'
Variable NFκB- 30 bp Reverse primer	5'- GTGAGGTTCATCGCAGAGATCTTGT**GGRRNNYCCC**ACATGATCCAGATCATAAGAGGCAC-3'
Fixed NFκB- 30 bp Forward primer	5'-ATCTCTGCGATGAACCTCACCATGT**GGGACTTTCC**ACAAGGTGCCTCTTATGATCTGGAT-3'
Fixed NFκB- 30 bp Reverse primer	5'- GTGAGGTTCATCGCAGAGATCTTGT**GGAAAGTCCC**ACATGATCCAGATCATAAGAGGCAC-3'
Fixed NFκB- 20 bp Forward primer	5'-ATCTCTGCGATGAACCTCAC**GGGACTTTCC**GTGCCTCTTATGATCTGGAT-3'
Fixed NFκB- 20 bp Reverse primer	5'- GTGAGGTTCATCGCAGAGAT**GGAAAGTCCC**ATCCAGATCATAAGAGGCAC-3'
Fixed NFκB- 15 bp Forward primer	5'-TGCGATGAACCTCAC**GGGACTTTCC**GTGCCTCTTATGATC-3'
Fixed NFκB- 15 bp Reverse primer	5'- GTGAGGTTCATCGCA**GGAAAGTCCC**GATCATAAGAGGCAC-3'
Fixed AP-1- 30 bp Forward primer	5'-ATCTCTGCGATGAACCTCACCATGT**TGAGTCA**ACAAGGTGCCTCTTATGATCTGGAT-3'
Fixed AP-1- 30 bp Reverse primer	5'- GTGAGGTTCATCGCAGAGATCTTGT**TGACTCA**ACATGATCCAGATCATAAGAGGCAC-3'
Fixed AP-1- 20 bp Forward primer	5'-ATCTCTGCGATGAACCTCAC**TGAGTCA**GTGCCTCTTATGATCTGGAT-3'
Fixed AP-1- 20 bp Reverse primer	5'- GTGAGGTTCATCGCAGAGAT**TGACTCA**ATCCAGATCATAAGAGGCAC-3'
Fixed AP-1- 15 bp Forward primer	5'-TGCGATGAACCTCAC**TGAGTCA**GTGCCTCTTATGATC-3'
Fixed AP-1- 15 bp Reverse primer	5'- GTGAGGTTCATCGCA**TGACTCA**GATCATAAGAGGCAC-3'
Fixed HRE- 20 bp Forward primer	5'-ATCTCTGCGATGAACCTCAC**ACGTGC**GTGCCTCTTATGATCTGGAT-3'
Fixed HRE- 20 bp Reverse primer	5'-GTGAGGTTCATCGCAGAGAT**GCACGT**ATCCAGATCATAAGAGGCAC-3'
Fixed HRE- 15 bp Forward primer	5'-TGCGATGAACCTCAC**ACGTGC**GTGCCTCTTATGATC-3'

Fixed HRE- 15 bp Reverse primer	5'-GTGAGGTTCATCGCA**GCACGT**GATCATAAGAGGCAC-3'
Fixed HRE- 10 bp Forward primer	5'-TGAACCTCAC**ACGTGG**GTGCCTCTTA-3'
Fixed HRE- 10 bp Reverse primer	5'- GTGAGGTTCA**GCACGT**TAAGAGGCAC-3'
5nt Spacer -20 bp Reverse primer	5'- GTGAGGTTCATCGCAGAGAT**CATGG**ATCCAGATCATAAGAGGCAC-3'
5nt Spacer -15 bp Reverse primer	5'- GTGAGGTTCATCGCA**CATGG**GATCATAAGAGGCAC-3'
10nt Spacer -20 bp Reverse primer	5'- GTGAGGTTCATCGCAGAGAT**ACAGACATGG**ATCCAGATCATAAGAGGCAC-3'
10nt Spacer -15 bp Reverse primer	5'- GTGAGGTTCATCGCA**ACAGACATGG**GATCATAAGAGGCAC-3'
5'- Stop-NheI 20 bp primer	5'-CAGTT<u>GCTAGC</u>GTGCCTCTTATGATCTGGAT-3'
5'- Stop-NheI 15 bp primer	5'-CAGTT<u>GCTAGC</u>GTGCCTCTTATGATC-3'
5'- Stop-NheI 10 bp primer	5'-CAGTT<u>GCTAGC</u>GTGCCTCTTA-3'
3'- Stop-XhoI 20 bp primer	5'-GGATT<u>CTCGAG</u>ATCCAGATCATAAGAGGCAC-3'
3'- Stop-XhoI 15 bp primer	5'-GGATT<u>CTCGAG</u>GATCATAAGAGGCAC-3'
3'- Stop-XhoI 10 bp primer	5'-GGATT<u>CTCGAG</u>TAAGAGGCAC-3'
3'- Stop-XhoI- 0bp-SalI 20 bp primer	5'-GGATT<u>GTCGACTCGAG</u>ATCCAGATCATAAGAGGCAC-3'
3'- Stop-XhoI- 5bp-SalI 20 bp primer	5'-GGATT<u>GTCGAC</u>CATGG<u>CTCGAG</u>ATCCAGATCATAAGAGGCAC-3'
3'- Stop-XhoI- 9bp-SalI 20 bp primer	5'-GGATT<u>GTCGAC</u>CAGACATGG<u>CTCGAG</u>ATCCAGATCATAAGAGGCAC-3'
3'- Stop-XhoI- 14bp-SalI 20 bp primer	5'-GGATT<u>GTCGAC</u>GGATACAGACATGG<u>CTCGAG</u>ATCCAGATCATAAGAGGCAC-3'

Table 2.3. List of oligonucleotides used for the Assembly PCR method. TFBSs are highlighted in bold. The NheI, XhoI or SalI restriction sites within the 'Stop' oligonucleotides are underlined.

2.7.2. Assembly reaction (x10 cycles)

Modifications were made to an online published Assembly PCR protocol by Team Heidelberg (http://2009.igem.org/Team:Heidelberg/Project_Synthetic_promoters). Briefly, an initial oligonucleotide mix containing 5 µl of 100 µM forward TFBS-oligonucleotide, 5 µl of 100 µM reverse TFBS-oligonucleotide, 12.8 µl of 100 µM 5'-Stop-NheI and 12.8 µl of 100 µM 3'-Stop-XhoI oligonucleotides in a final reaction volume of 35.6 µl. The oligonucleotide mix was diluted 1:100 in sterile dH_2O and a volume equal to the final reaction volume before dilution was added to the assembly PCR reaction (x10 cycles). For example, 35.6 µl diluted oligonucleotide mix was added to 12 µl 5x Phusion HF Buffer (NEB), 6 µl 2 mM dNTPs (Invitrogen), 0.5 µl Phusion DNA Polymerase (NEB) and sterile dH_2O to a final volume of   60 µl. The oligonucleotides were assembled in a ten cycle PCR reaction to create a double stranded DNA template using the thermocycler program: heated lid at 110°C, initial denaturation at 95°C for 5 minutes, 10 cycles of 95°C for 30 seconds, 60°C for 45 seconds, 72°C for 45 seconds and an additional extension at 72°C for 2 minutes in the G-Storm Thermocycler (model GS482, Gene Technologies Limited, Essex, UK). The reaction was purified from excess dNTPs and unassembled oligonucleotides using the PureLink® PCR Purification Kit (Invitrogen) to yield 50 µl of purified double stranded assembled PCR products.

2.7.3. Amplification reaction (x25 cycles)

Purified assembled PCR products were used as the template for the amplification PCR reaction (x25 cycles). Three identical reactions comprising 15 µl of assembled PCR products, 10 µl 5x Phusion HF Buffer (NEB), 5 µl 2mM dNTPs (Invitrogen), 5 µl 10 µM   5'-Stop-NheI, 5 µl 10 µM 3'-Stop-XhoI oligonucleotides and 0.5 µl Phusion DNA Polymerase (NEB) and sterile dH_2O to a final volume of 50 µl was PCR amplified using the thermocycler program: heated lid at 110°C, initial denaturation of 95°C for 5 minutes, 25 cycles of 95°C for 30 seconds, 60°C for 45 seconds, 72°C for 45 seconds and an additional extension at 72°C for

5 minutes (G-Storm Thermocycler GS482). After PCR, the three identical reactions were combined and purified following two successive high cut-off PCR purification steps using the Binding Buffer HC provided in the PureLink® PCR Purification Kit (Invitrogen) which served as an alternative to gel extraction. The high cut-off property of Buffer HC eliminated PCR products <300 bp which included the amplification primers and failed PCR products.

2.7.4. Routine digestion of Assembly PCR products

The amplified PCR products (50 µl) were digested with 2.5 µl NheI (NEB), 2.5 µl XhoI (NEB), 10 µl 10x NEB Buffer 2 and sterile dH$_2$O to a final volume of 100 µl. The reaction was incubated at 37°C overnight, followed by a high cut-off purification using the PureLink® PCR Purification Kit (Invitrogen). The digested PCR products were cloned upstream of the mCMV promoter within the equivalent NheI/XhoI site of the pCpGmCMV cloning vector (section 2.6.2) or within the pCpGmCMV-Xbp cloning vectors, where X bp is the spacer length which was introduced to increase the spacing between the proximal TFBS and the TATA box (section 2.9.1).

2.8. Assembly PCR Cloning Method: Construction of NFκB and AP-1-responsive promoters with increased spacing between the TFBSs

PCR products with NFκB or AP-1 motifs, separated by varied degrees of spacing between the TFBSs, were generated by assembling TFBS-oligonucleotides and spacer-oligonucleotides with different length annealing sequences. Nine promoter libraries with a   15 bp – 60 bp space between the NFκB or AP-1 motifs were constructed by combining the oligonucleotides specified in Table 2.4 using the Assembly PCR method described in the preceding section 2.7 (schematically presented in Fig 3.11, section 3.3.5).

	15bp between TFBS	20bp between TFBS	30bp between TFBS	35bp between TFBS	40bp between TFBS	45bp between TFBS	50bp between TFBS	55bp between TFBS	60bp between TFBS
Forward TFBS oligo **30 bp annealing sequence**			●					●	●
Reverse TFBS oligo **30 bp annealing sequence**			●						
Forward TFBS oligo **20 bp annealing sequence**		●				●	●		
Reverse TFBS oligo **20 bp annealing sequence**		●							
Forward TFBS oligo **15 bp annealing sequence**	●			●	●				
Reverse TFBS oligo **15 bp annealing sequence**	●								
5nt spacer reverse oligo **20 bp annealing sequence**						●		●	
10nt spacer reverse oligo **20 bp annealing sequence**							●		●
5nt spacer reverse oligo **15 bp annealing sequence**				●					
10nt spacer reverse oligo **15 bp annealing sequence**					●				
5'- Stop-NheI **20 bp annealing sequence**		●	●			●	●	●	●
3'- Stop-XhoI **20 bp annealing sequence**		●	●			●	●	●	●
5'- Stop-NheI **15 bp annealing sequence**	●			●	●				
3'- Stop-XhoI **15 bp annealing sequence**	●			●	●				

Table 2.4. List of oligonucleotides incorporated into the x10 cycle PCR reaction to generate promoters with nine variations of spacing between the TFBSs. The four oligonucleotides indicated by a dot were assembled in the x10 cycle PCR reaction to generate the specified spacing between the TFBSs.

2.9. Assembly PCR Cloning Method: construction of NFκB and AP-1-responsive promoters with increased spacing between the proximal TFBS and TATA box

2.9.1. Cloning vectors: pCpGmCMV-Xbp

The cloning vectors pCpGmCMV-Xbp, where X bp is the spacing between the proximal TFBS and the TATA box, were generated by initially combining 5 µl of 100 µM forward HRE-oligonucleotide, 5 µl of 100 µM reverse HRE-oligonucleotide, 12.8 µl of 100 µM 5'-Stop-NheI oligonucleotide and 12.8 µl of the various 3'-Stop XhoI oligonucleotides (100 µM) listed in Table 2.5, in four individual reactions of 35.6 µl. All oligonucleotides contained 20 bp annealing sequences to generate the specific spacing between the proximal TFBS and the TATA box.

Spacing between proximal TFBS and TATA box	60bp	66bp	70bp	74bp
Forward TFBS oligo 20 bp annealing sequence	●	●	●	●
Reverse TFBS oligo 20 bp annealing sequence	●	●	●	●
5'- Stop-NheI 20 bp annealing sequence	●	●	●	●
3'- Stop-XhoI- 0bp- SalI 20 bp annealing sequence	●			
3'- Stop-XhoI- 5bp- SalI 20 bp annealing sequence		●		
3'- Stop-XhoI- 9bp- SalI 20 bp annealing sequence			●	
3'- Stop-XhoI- 14bp- SalI 20 bp annealing sequence				●

Table 2.5. List of 3'-Stop-XhoI-oligonucleotides incorporated into the x10 cycle PCR reaction to generate promoters with spacing between the proximal TFBS and the TATA box. The four oligonucleotides indicated by a dot were assembled in the x10 cycle PCR reaction. The PCR products were amplified using the standard 5'-Stop-NheI and 3'-Stop-XhoI primers in the x25 cycle PCR reaction.

The four individual oligonucleotide mixtures were diluted 1:100 in sterile dH$_2$O and assembled and amplified following the Assembly PCR protocol described in section 2.7. The resulting

purified PCR products contained a 5'-NheI, an internal XhoI and a 3'-SalI restriction enzyme site. Therefore, 50 µl PCR product was digested with 2.5 µl NheI (NEB), 2.5 µl SalI-HF (high-fidelity, NEB), 10 µl 10x NEB Buffer 4 and sterile dH$_2$O to a final volume of 100 µl. The reaction was incubated at 37°C overnight, followed by a high cut-off purification using the PureLink® PCR Purification Kit (Invitrogen). The NheI/SalI digested PCR products were cloned into the compatible NheI/XhoI site of the pCpGmCMV cloning vector (section 2.6.2) in four standard ligation reactions which were then transformed into chemically-competent GT115 cells from which the plasmid DNA constructs were isolated.

The pCpG-Xbp cloning vectors were generated by digesting the four individual plasmid DNA constructs (25 µg) with 2.5 µl NheI (NEB), 2.5 µl XhoI (NEB), 10 µl 10x NEB Buffer 2 and sterile dH$_2$O in a reaction volume of 100 µl. The reaction was incubated at 37°C overnight and the DNA fragments were separated by gel electrophoresis. The DNA was purified from agarose using the PureLink® Gel Extraction Kit (Invitrogen) to yield 50 µl of purified pCpG-Xbp cloning vectors for the Assembly PCR products (Fig 3.15, section 3.3.6).

Table 2.6 lists the pCpG-Xbp cloning vectors used to generate the specified spacing between the proximal TFBS and the TATA box.

3'-Stop- XhoI primers 20 bp annealing sequence	pCpGmCMV-Xbp- cloning vector (digested with NheI/XhoI)	Distance between proximal TFBS and TATA box within the promoter
3'-Stop- XhoI (current)	pCpGmCMV	55 bp
3'-Stop- XhoI- 0bp- SalI	pCpGmCMV-0bp	60 bp
3'-Stop- XhoI- 5bp- SalI	pCpGmCMV-5bp	66 bp
3'-Stop- XhoI- 9bp- SalI	pCpGmCMV-9bp	70bp
3'-Stop- XhoI- 14bp- SalI	pCpGmCMV-14bp	74 bp

Table 2.6. List of pCpGmCMV-Xbp-cloning vectors used to generate the required spacing between the proximal TFBS and the TATA box.

2.9.2. Construction of synthetic promoters with increased spacing between proximal TFBS and TATA box

PCR products with NFκB or AP-1 motifs, each separated by 20 bp space, were generated using the Assembly PCR protocol described in section 2.7. The resulting NheI/XhoI digested PCR products were cloned into the equivalent site within the pCpGmCMV (section 2.6.2) or the pCpGmCMV-Xbp cloning vectors (section 2.9.1) to generate NFκB- and AP-1-responsive promoters with various degrees of spacing between the proximal TFBS and the TATA box (Table 2.6).

2.10. Construction of pCpG-20bp-composite synthetic promoters

The oligonucleotide mix used to construct the composite promoters included 1.6 µl of each 100 µM forward NFκB-, AP-1- and HRE-oligonucleotides with 1.6 µl of each 100 µM reverse NFκB-, AP-1- and HRE-oligonucleotides, 12.8 µl 5'-Stop-NheI and 12.8 µl 3'-Stop-XhoI oligonucleotides in a final volume of 35.2 µl. All oligonucleotides contained 20 bp annealing sequences. The oligonucleotide mix was diluted 1:100 in sterile dH_2O and assembled and amplified using the Assembly PCR protocol (section 2.7). The resulting NheI/XhoI digested PCR products were cloned into the NheI/XhoI site within the pCpGmCMV-5bp cloning vector (section 2.9.1) to generate constructs with a 20 bp space between the TFBSs and a 66 bp between the proximal TFBS and the TATA box (Fig 3.20, section 3.5).

2.11. Construction of pCpG-cluster composite promoters

2.11.1. Cloning vector: pCpG-proximal TFBS cluster

The pCpG-6NFκB-Luc[+], pCpG-8AP-1-Luc[+], pCpG-6HRE-Luc[+] which contain 6NFκB, 8AP-1 and 6HRE clusters and 20 bp between the TFBSs, respectively, were selected from the various promoters libraries in sections 2.8 and 2.9. Each construct (25 µg) was digested with 3 µl NheI (NEB), 10 µl 10x NEB Buffer 2 and sterile dH_2O to a final volume of 100 µl and

incubated at 37°C overnight. The linear DNA was separated by gel electrophoresis and purified from the agarose using the PureLink® Gel Extraction Kit (Invitrogen). The DNA was dephosphorylated following incubation with CIP enzyme (NEB) and subsequently purified using the PureLink® PCR Purification Kit (Invitrogen) to yield 50 µl of the linear cloning vectors; pCpG-proximal 6NFκB, pCpG-proximal 8AP-1, pCpG-proximal 6HRE (Fig 3. 23 A, section 3.6).

2.11.2. Construction of pCpG-clustered composite promoters

The 6NFκB, 8AP-1 and 6HRE clusters were isolated from their respective plasmid DNA constructs by digesting 25 µg of each construct with 2.5 µl NheI (NEB), 2.5 µl XhoI (NEB), 10 µl 10x NEB Buffer 2 and sterile dH$_2$O in a final volume of 100 µl and incubated at 37°C overnight. The DNA fragments were separated by gel electrophoresis and the 6NFκB, 8AP-1 and 6HRE clusters were purified from the agarose using the PureLink® Gel Extraction Kit (Invitrogen). The concentration of the purified clusters was quantified using the Nanodrop and equal concentrations of NheI/XhoI digested 8AP-1 and 6HRE clusters were cloned into the NheI linear pCpG-proximal-6NFκB vector. Similarly, equal concentrations of 6NFκB and 6HRE clusters were cloned into the pCpG-proximal-8AP-1 vector and 6NFκB and 8AP-1 clusters were cloned into the pCpG-proximal-6HRE vector in standard ligation reactions of 20 µl. Due to the cloning strategy, the pCpG-clustered synthetic promoters were comprised of a minimum of 3 clustered TFBSs with the second cluster being cloned in the reverse orientation (Fig 3. 23 B and C, section 3.6).

2.12. MiRNA-mediated regulation of gene expression

2.12.1. Real-time qPCR: miRNA expression profiling

2.12.1.1. Isolation of small RNA from NIH3T3 mouse embryonic fibroblasts

Prior to RNA purification, all surfaces and equipment were cleaned with RNase AWAY® (Fisher Scientific Ltd, Leicestershire, UK) and gloves were changed regularly to prevent RNase contamination and/or RNA degradation. NIH3T3 mouse embryonic fibroblasts were maintained in complete DMEM medium and harvested at ~70% confluency. The cells were seeded at 1 x 10^6 cells into 10 cm^2 tissue culture dishes in 10 ml complete DMEM medium. After 24 hours, the medium was aspirated and replaced with 10 ml of 0.5% FBS DMEM medium only or 0.5% FBS DMEM medium containing 50 ng/ml mouse IL-17A or 10 ng/ml TNFα or a combination of 50 ng/ml mIL-17A with 10 ng/ml TNFα for 18 hours. Post-incubation, the cells were washed with 1 ml sterile PBS and the RNA fraction enriched with small RNA species was isolated using the mirVANA™ miRNA Isolation Kit (Ambion/Applied Biosystems, Texas, USA), following the manufacturer's instructions.

The PBS was carefully removed from the cells and the dish was then placed on ice. The cells were lysed with 600 μl Lysis/Binding Solution and the lysates were collected using a sterile rubber policeman and transferred into a 1.6 ml Eppendorf tube. The tubes were briefly vortexed to completely lyse the cells and 1/10th of the lysate volume of miRNA Homogenate Additive was added to the cell lysate, vortexed and incubated on ice for 10 minutes. Acid Phenol: Chloroform was added in a volume equal to that of the original lysate volume, and vortexed for 1 minute. Following centrifugation at 10,000 g for 5 minutes at room-temperature (Eppendorf microcentrifuge 5415D), the aqueous phase was transferred into a new sterile 1.6 ml Eppendorf tube. The resulting semi-pure RNA sample was then purified using the glass-fiber filter method, where a third of the volume of 100% ethanol was added to the aqueous phase, briefly vortexed and then transferred onto the filter cartridge. The lysate/ethanol sample was then passed through the filter cartridge by centrifugation at 10,000 g for 15 seconds and

the filtrate was collected and measured and two-thirds volume of 100% ethanol was added to the filtrate, briefly vortexed and passed through a second filter cartridge. The sample was centrifuged at 10,000 g for 15 seconds and the flow-through was discarded. The step was repeated until the entire sample had been filtered. Following successive washes with Wash Solutions 1 and 2/3, respectively, the filter was transferred into a fresh collection tube. The RNA was eluted with 100 µl of pre-heated (95°C) nuclease-free water and centrifuged at maximum speed for 30 seconds to yield a 100 µl small RNA-containing eluate. The RNA concentration (ng/µl) was quantified using the Nanodrop spectrophotometer.

2.12.1.2. cDNA Synthesis

Small RNA was used as a template to synthesise cDNA using the QuantiMir RT Kit Small RNA Quantitation System (Systems Biosciences, California, USA). The polyA tail was added to the RNA template in a reaction comprising 10 ng small RNA, 2 µl 5x PolyA Buffer, 1 µl   25 mM $MnCl_2$, 1.5 µl 5 mM ATP, 0.5 µl PolyA Polymerase and RNase-free H_2O in a final volume of 10 µl and incubated at 37°C for 30 minutes. After incubation, 0.5 µl Oligo dT Adaptor was added to the reaction and incubated for 5 minutes at 60°C and then cooled at room temperature for 2 minutes. The final cDNA synthesis step involved the addition of 4 µl 5x RT Buffer, 2 µl dNTP mix, 1.5 µl 0.1M DTT, 1.5 µl RNase-free H_2O and 1 µl Reverse Transcriptase and incubation at 42°C for 60 minutes and 95°C for 10 minutes. The cDNA samples were stored at -20°C.

2.12.1.3. Preparation of miRNA standards using end-point PCR

The sequences of mature miR-23b and the control miR-17-5p, miR-103-3p and miR-191-5p sequences were obtained from the miRBase database (www.mirbase.org) and synthesised as desalted oligonucleotides from Sigma-Aldrich (Table 2.7) to serve as forward PCR primers. The control miR-16-5p and U6 forward PCR primers were supplied with the QuantiMir Kit.

	PCR primer sequences (5'-3')
Forward hsa-miR-23b-3p primer	5'- ATCACATTGCCAGGGATTACC -3'
Forward hsa-miR-17-5p control primer	5'- CAAAGTGCTTACAGTGCAGGTAG -3'
Forward hsa-miR-103-3p control primer	5'- AGCAGCATTGTACAGGGCTATGA -3'
Forward hsa-miR-191-5p control primer	5'- CAACGGAATCCCAAAAGCAGCTG -3'
Forward hsa-miR-16-5p control primer	5'- TAGCAGCACGTAAATATTGGCG -3'
Forward hsa-U6 control primer	5'- CGCAAGGATGACACGCAAATTC -3'

Table 2.7. Forward PCR primers of miRNAs. Forward PCR primers anneal to the mature candidate miRNA in cDNA samples.

Each PCR primer was resuspended at 10 µM in sterile TE buffer and the cDNA samples were diluted 1:50 in nuclease-free water. The miRNAs listed in Table 2.7 were amplified from the unstimulated cDNA sample in an end-point PCR reaction to generate the miRNA standards for the subsequent Real-time qPCR reaction. Briefly, the diluted cDNA (1 µl) was added to 1 µl 10 µM forward miRNA primer, 1 µl 10 µM Universal Reverse Primer (QuantiMir kit), 2 µl 2mM dNTPs, 4 µl 5x Phusion HF Buffer, 0.5 µl Phusion DNA polymerase and nuclease-free water in a final volume of 20 µl, of six individual reactions for each miRNA. The PCR reaction proceeded with a heated lid at 110°C, initial denaturation at 98°C for 5 minutes followed by 30

cycles of 98°C for 10 seconds, 60°C for 60 seconds, 72°C for 30 seconds and a final extension at 72°C for 10 minutes (G-Storm Thermocycler GS482). The PCR reactions (20 µl) were separated on a 2.0% low melting agarose gel (Promega) and the PCR products (~68 bp) were excised from the gel using the PureLink® Quick Gel Extraction Kit (Invitrogen).

The NanoDrop spectrophotometer was used to quantify the concentration of the purified PCR products (ng/µl) and the following calculation was used to calculate the copy number of each miRNA in their corresponding PCR products (10^{10} g/µl), where A = concentration of the PCR product:

Copy of each miRNA (10^{10} g/µl) = (A x 10^{-9} x (6.02 x 10^{23})) / (660 x (PCR product size))

The following calculation was used to determine the initial volume of PCR product required to make the first serial dilution of 0.5 x 10^9 miRNA copies/µl standard sample:

Volume of PCR product (µl) = ((0.5 x 10^9) x 100) / (copy number of miRNA)

The first serial dilution of 0.5 x 10^9/µl standard sample was prepared by diluting the calculated volume of PCR product (µl) in 5 µg/ml tRNA in a total volume of 100 µl. Following a brief vortex, 5 µl of the diluted PCR product was further diluted in 45 µl tRNA in a total volume of 50 µl to generate the 0.5 x 10^8 standard. The preceding standard was further diluted by combining 5 µl of 0.5 x 10^8/µl standard sample with 45 µl tRNA in a total volume of 50 µl to generate the 0.5 x 10^7/µl standard sample. This process was continued until serial dilutions to 0.5 x 10^1/µl standard sample was prepared for each miRNA.

2.12.1.4. Absolute quantification of miRNA expression using Real-time qPCR

Absolute Real-time quantitative PCR was used to quantify the copy number of miR-23b and the control miR-17-5p, miR-103-3p, miR-191-5p, miR-16-5p and U6 in unstimulated, mIL-17A, TNFα and mIL-17A + TNFα–stimulated cDNA from NIH3T3 cells. The 384 well-plate setup was performed using the ABI 7900 Sequence Detection System 2.4 (Applied Biosystems, Life Technologies Corp, California, USA) and the reactions were performed using the Applied Biosystems 7900HT Fast Real-time PCR System.

Each reaction comprised 1 µl cDNA (unstimulated and stimulated cDNA diluted 1:50), 0.4 µl of 10 µM Universal Reverse Primer (QuantiMir), 0.4 µl of each 10 µM miRNA forward primer, 5 µl of 2x Brilliant III qPCR Ultra-Fast SYBR® Green qPCR Master mix (600882; Agilent Technologies, CA, USA), 0.15 µl of 20 µM Rox reference dye (Agilent) and RNase-free water to a final volume of 10 µl.

The reactions required to generate the standard curve comprised 1 µl of each standard ($0.5 \times 10^7 - 0.5 \times 10^1$ copies/µl), 0.4 µl of 10 µM Universal Reverse Primer (QuantiMir), 0.4 µl of each miRNA forward primer, 5 µl of 2x Brilliant III qPCR Ultra-Fast SYBR® Green qPCR Master mix (Agilent), 0.15 µl of 20 µM Rox reference dye (Agilent) and RNase-free water to a final volume of 10 µl.

The no-template control (NTC) reaction comprised 1 µl of 5µg/ml tRNA, 0.4 µl of 10 µM Universal Reverse primer (QuantiMir), 0.4 µl of each 10 µM miRNA forward primer, 5 µl 2x Brilliant III qPCR Ultra-Fast SYBR® Green qPCR Master mix (Agilent), 0.15 µl of 20 µM Rox reference dye (Agilent) and RNase-free water to a final volume of 10 µl.

The reactions were performed using the thermocycler program: stage 1 incubation at 60°C for 2 minutes, stage 2 incubation at 95°C for 3 minutes, stage 3 with 40 cycles of 95°C for 5 seconds, 60°C for 15 seconds, and stage 4 incubation at 95°C for 15 seconds, 60°C for 15 seconds and 95°C for 15 seconds.

2.12.1.5. Real-time qPCR data analysis and normalisation

The PCR data which included the standard curve plots, amplification plots, and dissociation curves was analysed using the SDS 2.4 Software (Applied Biosystems) and presented in Appendix 15. The copy numbers of each miRNA in the unstimulated and stimulated cDNA samples were calculated by comparing the cycle threshold (Ct) values to the standard curve for the corresponding miRNA. The copy numbers of the five control miRNAs quantified in each cDNA sample was entered into the Excel Add-In NormFinder Algorithm (Andersen *et al.,* 2004) which identified the optimal normalisation miRNA among the set of candidates by evaluating the overall expression variability of the candidate normalisation miRNAs and also the variation between sample subgroups of the sample set. Due to low expression variability, miR-191 was selected as the 'best normaliser'. The miR-23b copy numbers in each cDNA sample was normalised to the miR-191 copy numbers in the same samples by dividing the miR-23b copy number by the calculated miR-191 normalisation factors.

2.12.2.   Construction of pcLuc⁺-miR-23b-target expression vectors

2.12.2.1. Cloning vector: pcLuc⁺

The pcLuc⁺ expression vector (25 µg) (Gould *et al.,* 2004) was digested with 2.5 µl XbaI (NEB), 2.5 µl ApaI (NEB), 10 µl 10x NEB Buffer 4 and sterile dH$_2$O to a final volume of 100 µl and incubated at 25°C for 16 hours for optimal ApaI activity and then at 37°C for 5 hours for optimal XhoI activity. The DNA fragments were separated by gel electrophoresis and purified from the agarose using the PureLink® Gel Extraction Kit (Invitrogen) to yield 50 µl of pcLuc⁺ cloning vector (Fig 6.2 A, section 6.1.1).

2.12.2.2. pcLuc⁺-miR-23b-target expression vectors

The miR-23b-target oligonucleotides containing two miR-23b target sites were synthesised with 5'-CTAG and 3'-ApaI overhangs and internal XbaI and EcoRI restriction enzyme sites from Sigma-Aldrich (Table 2.8) and resuspended at 2mg/ml in sterile TE buffer.

	Cloning oligonucleotide sequences (5'-3')
Forward hsa-miR-23b-3p cloning oligo	5'-<u>CTAG</u>**GGTAATCCCTGGCAATGTGAT**CGAT**GGTAATCCCTGGCAATGTGAT**TCTAGAATTC<u>GGGCC</u>-3'
Reverse hsa-miR-23b-3p cloning oligo	5'-<u>C</u>GAATTCTAGA**ATCACATTGCCAGGGATTACC**ATCG**ATCACATTGCCAGGGATTACC**-3'

Table 2.8. Sequences of miR-23b target oligonucleotides. Oligonucleotides were HPLC purified. Overhangs are underlined and miR-23b target sites are highlighted in bold.

Forward and reverse miR-23b-oligonucleotides were annealed using the boiling method and cloned into the compatible XbaI/ApaI site within the pcLuc⁺ cloning vector to create pcLuc⁺-miR-23b-2T, in a standard ligation reaction (Fig 6.2 B, section 6.1.1). The pcLuc⁺-miR-23b-4T constructs with four miR-23b target sequences were generated by digesting 25 µg pcLuc⁺-miR-23b-2T with XbaI and ApaI and cloning a second miR-23b-target oligonucleotides into the equivalent site, using the same method described above. Standard ligation reactions of 20 µl were transformed into DH5α *E.coli* cells and the DNA was isolated from positive colonies using the routine plasmid DNA purification protocols. Correct construction of recombinant DNA was verified by restriction enzyme digestion using EcoRI (Fig 6.2 D, section 6.1.1) and DNA sequencing using the End of Luc⁺ forward sequencing primer (Appendix 12.1 and 12.2).

2.12.3. Construction of pCpG-6NFκB-miR-23b-2T expression vector

2.12.3.1. Cloning vector: 6NFκB-Luc⁺

The pCpG-6NFκB-Luc$^+$ construct (25 µg) was digested with 2.5 µl PpumI (NEB) and 2.5 µl FseI (NEB), 10 µl 10x NEB Buffer 4 and sterile dH$_2$O to a final volume of 100 µl and incubated at 37°C overnight. The DNA fragment containing the 6NFκB synthetic promoter and the 5'-portion of the luciferase gene were separated by gel electrophoresis and purified from the agarose using the PureLink® Gel Extraction Kit (Invitrogen) to yield 50µl of the 6NFκB-Luc$^+$ cloning vector (Fig 6.9 A, section 6.6).

2.12.3.2. pCpG-6NFκB-miR-23b-2T expression vector

The 3'-portion of the luciferase gene and two miR-23b target sites were PCR amplified from the pcLuc$^+$-miR-23b-2T construct using the PCR primers listed in Table 2.9.

	PCR primer sequences (5'-3')
Forward FseI PCR primer	5'-AGCGGTTGCCAAGAGGTTCCATCTGCCA-3'
Reverse FseI PCR primer	5'-ATGTCACGTA<u>GGCCGGCC</u>CGAATTCTAG-3'

Table 2.9. PCR primers used to amplify miR-23b-2T from the pcLuc$^+$-miR-23b-2T construct. PCR primers were synthesised by Sigma-Aldrich and resuspended to 10 µM. The FseI restriction enzyme site within the reverse PCR primer is underlined.

The PCR product was digested with 2.5 µl PpumI (NEB) and 2.5 µl FseI (NEB), 10 µl 10x NEB Buffer 4 and sterile dH$_2$O to a final volume of 100 µl and incubated at 37°C overnight and then purified using the PureLink® PCR Purification Kit (Invitrogen) (Fig 6.9 B, section 6.6). The PCR product, containing PpumI and FseI overhangs, was cloned into the equivalent site within the 6NFκB-Luc$^+$ cloning vector to create the 6NFκB-Luc$^+$-miR-23b-2T construct (Fig 6.9 C, section 6.6).

2.12.4. Transfection of miRNA mimics

The regulation of luciferase mRNA expressed from pcLuc$^+$-miR-23b-target constructs by the miR-23b was verified by co-transfecting the cells with the pcLuc$^+$-miR-23b-target constructs and the synthetic miRNA mimics. Briefly, 3.2 x 10^4 293T cells were seeded on a 48-well plate with 200 µl of antibiotic-free 10% FBS DMEM medium. After 24 hours, 0.5 µl Lipofectamine$^®$ 2000 transfection reagent (Invitrogen) was diluted in 25 µl Opti-MEM I reduced serum medium (Invitrogen), briefly vortexed and incubated for 5 minutes at room temperature. Following the incubation, a DNA mixture of 90 ng pcLuc$^+$-miR-target DNA,   10 ng pRL-CMV and 1 µM miRNA mimic was diluted in Opti-MEM I reduced serum medium in a total volume of 25 µl, which was subsequently added to the diluted Lipofectamine$^®$, briefly vortexed and incubated for a further 20 minutes at room temperature. The 50 µl DNA:miRNA mimic:Lipofectamine$^®$ complex was added to the cells in a drop-wise manner and the cells were incubated under standard conditions for 24 hours. The firefly and renilla luciferase expression was quantified using the Luciferase Assay System (Promega) and Renilla Luciferase Assay System (Promega), respectively, and the firefly luciferase expression was normalised to the renilla luciferase expression.

2.13.   Lentiviral expression vector constructs

2.13.1. Cloning vector: pLV.CMV

The SIN lentiviral plasmid pLV.CMVenh.gp91.eGFP.cHS4, containing the eGFP reporter gene driven by the CMV promoter, positioned between cHS4 insulators, was purchased from Addgene (www.addgene.org; plasmid 30471; Barde *et al.,* 2011) as a bacterial stab and renamed pLV.CMV.GFP for simplification. A sterile glass rod was inserted into the bacterial stab and the culture was streaked on a fresh LB-agar plate containing 100 µg/ml Carbenicillin. The plate was incubated at 37°C overnight after which a single bacterial colony was inoculated in 5 ml of LB media containing 100 µg/ml Carbenicillin and incubated overnight at 37°C, with shaking at 225 rpm. A maxiprep preparation was subsequently setup and the lentiviral plasmid

pLV.CMV.GFP was isolated using the PureLink® HiPure Plasmid DNA Purification Kit (Invitrogen).

pLV.CMV.GFP (25 µg) was digested with 2.5 µl BamHI, 2.5 µl SalI, 10 µl 10x NEB Buffer 4 and sterile distilled water in a total volume of 100 µl and incubated at 37°C overnight. The DNA fragments were separated by gel electrophoresis and the 8357 bp fragment was isolated using the PureLink® Quick Gel Extraction Kit (Invitrogen) to yield 50 µl of purified pLV.CMV cloning vector (Fig 4.1.1 A, section 4.1.1).

2.13.2. Oligonucleotides for cloning the lentiviral vector constructs

The PCR primers were synthesised as desalted oligonucleotides with specified 5'- and 3'- restriction enzymes sites, at Sigma-Aldrich (Table 2.10). PCR primers were resuspended at 100 µM in sterile TE buffer. The HPLC purified mCMV cloning oligonucleotides were synthesised with 5'- SnaBI and 3'-BamHI overhangs, at Sigma-Aldrich (Table 2.10). The forward and reverse mCMV oligonucleotides were resuspended at 2 mg/ml in TE buffer and annealed using the boiling method, as previously described in section 2.3.1

	DNA Sequence (5'-3')
Forward Luciferase BamHI primer	5'- GATGAGCAGGATCCCATGGAAGACG -3'
Reverse Luciferase SalI primer	5'- ATGTACGCGTCGACTCTAGAATTAC -3'
Forward pCpGmCMV ClaI primer	5'- GTCGGATTATCGATGCTAGCGTGCC -3'
Forward pGL3mCMV and mCMV ClaI primer	5'- GTCGGATTATCGATGCGTGCTAGC -3'
Reverse all constructs BstBI primer	5'- ACTCGTAGTTCGAAGTACTCAGCGT -3'
Forward hIL-1Ra BamHI primer	5'- GATGAGCAGGATCCATGGAAATCTGCAGAGGC -3'
Reverse hIL-1Ra SalI primer	5'- ATGTACGCGTCGACCTACTCGTCCTCCTGGAA -3'

Forward mTNFRII-Fc BamHI primer	5'- GATGAGCA<u>GGATCC</u>ATGTACAGGATGCAACTC -3'
Reverse mTNFRII-Fc XhoI primer	5'- ATGTACGC<u>CTCGAG</u>TCATTTACCAGGAGAGTG -3'
Forward pCpGmCMV SnabI primer	5'- TCGGATT<u>TACGTA</u>GGCCAGCTAG -3'
Reverse pCpGmCMV BamHI primer	5'- ATGTACGC<u>GGATCC</u>TGAGCTCTGCTTATATAA -3'
Forward mCMV SnabI overhang cloning oligo	5'- <u>GTA</u>GCCTGTAGGCGTGTACGGTGGGAGGCTTATATAAGCAGAGCTCG -3'
Reverse mCMV BamHI overhang cloning oligo	5'- <u>GATCC</u>GAGCTCTGCTTATATAAGCCTCCCACCGTACACGCCTACAGGCTAC -3'
Forward SFFV promoter SnabI primer	5'- GATGAGCA<u>TACGTA</u> AATTCCTGCAGCCCCGAT -3'
Reverse SFFV promoter BamHI primer	5'- ATGTACGC<u>GGATCC</u>GGTGGCTTTACCAACAGT -3'
Forward lenti-miRNA MfeI primer	5'- CGTTAGC<u>CAATTG</u>GTAATTCTAG -3'
Reverse lenti-miRNA KpnI primer	5'- GGTGGAT<u>GGTACC</u>GAATAGGGCCC -3'

Table 2.10. List of oligonucleotides used to construct the lentiviral expression cassettes. PCR primers and oligonucleotides were synthesised by Sigma-Aldrich. Underlined sequences indicate the candidate restriction enzyme site or overhang.

2.13.3. Cloning vector: pLV.Luc⁺

The pGL3mCMV plasmid was diluted to 10 ng/µl and the luciferase gene was PCR amplified using forward and reverse primers with BamHI and SalI restriction enzyme sites, respectively (Table 2.10). The reaction comprised 1 µl of diluted pGL3mCMV, 2.5 µl 10 µM BamHI forward primer, 2.5 µl 10 µM SalI reverse primer, 5 µl 2 mM dNTPs, 10 µl 5x Phusion HF Buffer, 0.5 µl Phusion DNA Polymerase and nuclease-free water in a total reaction volume of 50 µl. The reaction was performed using the thermocycler program: heated lid at 110°C, initial denaturation at 98°C for 5 minutes, 30 cycles at 98°C for 10 seconds, 60°C for 30 seconds, 72°C for 30 seconds and a final extension at 72°C for 5 minutes.

The luciferase PCR product of 1656 bp was verified by gel electrophoresis and purified using the PureLink® PCR Purification Kit (Invitrogen). The purified PCR product (50 µl) was digested with 2.5 µl BamHI, 2.5 µl SalI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a total volume of 100 µl and incubated at 37°C overnight (Fig 4.1.1 B, section 4.1.1). The digested luciferase PCR product was cloned into the pLV.CMV cloning vector in a ligation reaction of 20 µl which was subsequently transformed into chemically-competent DH5α cells and the plasmid DNA was purified from positive bacterial transformants using the routine methodology described in section 2.3. The resulting purified DNA was referred to as pLV.CMV.Luc⁺.

The CMV promoter and the 5'-portion of the luciferase gene was removed from the pLV.CMV.Luc⁺ construct by digesting 25 µg pLV.CMV.Luc⁺ with 2.5 µl PmeI, 2.5 µl BstBI,   10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a total volume of 100 µl, which was incubated at 37°C for 12 hours for optimal PmeI activity and then at 65°C for 4 hours for optimal BstBI activity. The DNA fragments were separated by gel electrophoresis and the 7882 bp fragment was isolated from agarose using the PureLink® Quick Gel Extraction Kit (Invitrogen). The resulting purified pLV.Luc⁺ served as the cloning vector for the lentiviral synthetic promoters (Fig 4.1.1 C, section 4.1.1).

2.13.4. Construction of LV-composite synthetic promoters expressing luciferase

The candidate pGL3-composite promoter and pCpG-composite promoter plasmid DNA were diluted to 10 ng/µl and the synthetic promoter and 5'- portion of the luciferase gene was PCR amplified from each construct using forward and reverse PCR primers with ClaI and BstBI restriction enzyme sites, respectively (Table 2.10). The reaction comprised 1 µl of diluted plasmid DNA, 2.5 µl 10 µM ClaI forward primer, 2.5 µl 10 µM BstBI reverse primer, 5 µl   2 mM dNTPs, 10 µl 5x Phusion HF Buffer, 0.5 µl Phusion DNA Polymerase and nuclease-free water in a total reaction volume of 50 µl. The reaction was performed using the thermocycler program: heated lid at 110°C, initial denaturation at 98°C for 5 minutes, 30 cycles at 98°C for

10 seconds, 60°C for 30 seconds, 72°C for 30 seconds and a final extension at 72°C for 5 minutes.

The PCR products were purified using the PureLink® PCR Purification Kit and 50 µl of purified DNA was digested with 2.5 µl ClaI, 10 µl 10x NEB Buffer 4 and sterile distilled water in a total reaction volume of 100 µl and incubated at 37°C overnight. The DNA was purified using the PureLink® PCR Purification Kit and the ends of the DNA were blunted by incubating 50 µl purified DNA with 3 µl Klenow (NEB), 10 µl 10x NEB Buffer 2, 10 µl 2 mM dNTPs and sterile dH$_2$O in a final reaction volume of 100 µl at 25°C for 15 minutes. The reaction was terminated by adding a final concentration of 10 mM EDTA to the reaction and heat inactivating the Klenow enzyme at 75°C for 20 minutes. The reaction was purified using the PureLink® PCR Purification Kit and 50 µl of purified DNA was digested with 2.5 µl BstBI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a reaction volume of 100 µl and incubated at 65°C overnight and then purified again. The PCR products with 5'-blunt ends and 3'-BstBI overhangs were cloned into the PmeI (blunt) and BstBI site within the pLV.Luc+ cloning vector in a standard ligation reaction, which restored the luciferase gene and incorporated the candidate synthetic promoters into the lentiviral vector. Following the transformation of DH5α cells, the DNA purified from positive colonies were referred to as LV- x -Luc+, where x is the designated number of the synthetic promoter, e.g. LV-2-Luc+ (Fig 4.1.1 D and E, section 4.1.1).

2.13.5. Construction of LV-SFFV-Luc+

The pUCL-Luc+ plasmid DNA was obtained from Professor Adrian Thrasher, Great Ormond Street, UK (Demaison et al., 2002). The constitutive SFFV promoter and the 5'-portion of the luciferase gene was isolated from pUCL-Luc+ by digesting 25 µg DNA with 2.5 µl EcoRI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a total volume of 100 µl. The reaction was incubated at 37°C for 12 hours and then purified using the PureLink® PCR Purification Kit (Invitrogen). The ends of the DNA were blunted using Klenow enzyme and purified again. The purified DNA (50 µl) was digested with 2.5 µl BstBI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a reaction

volume of 100 µl and incubated at 65°C overnight and then purified again. The DNA fragments were separated on an agarose gel and the 730 bp DNA fragment was isolated and purified using the PureLink® Gel Extraction Kit (Invitrogen). The SFFV-Luc+ fragment was cloned into the pLV.Luc+ vector in a standard ligation reaction to restore the luciferase gene and incorporate the SFFV promoter into the lentiviral vector. The DNA purified from positive bacterial transformants was referred to as LV-SFFV-Luc+ (Fig 4.1.2, section 4.1.2).

2.13.6. Cloning vectors: pLV.CMV.hIL-1Ra and pLV.CMV.mTNFRII-Fc

The therapeutic genes hIL-1Ra and mTNFRII-Fc were amplified from pcIL-1Ra and pFuse-mTNFRII-Fc DNA using PCR primers with BamHI and SalI restriction enzyme sites or BamHI and XhoI restriction enzyme sites, respectively (Table 2.10). The two individual reactions comprised 1 µl of the diluted respective insert DNA (10 ng/µl), 2.5 µl 10 µM BamHI forward primer, 2.5 µl 10 µM reverse primer (reverse hIL-1Ra SalI primer or reverse mTNFRII-Fc XhoI primer) 5 µl 2 mM dNTPs, 10 µl 5x Phusion HF Buffer, 0.5 µl Phusion DNA Polymerase and nuclease-free water in a total volume of 50 µl. The reaction was performed using the thermocycler program: heated lid at 110°C, initial denaturation at 98°C for 5 minutes, 30 cycles at 98°C for 10 seconds, 60°C for 30 seconds, 72°C for 30 seconds and a final extension at 72°C for 5 minutes (G-Storm Thermocycler GS482). The correct sizes of PCR products were verified by gel electrophoresis and purified using the PureLink® PCR Purification Kit (Invitrogen). The purified DNA (50 µl) was digested with 2.5 µl BamHI, 2.5 µl SalI for hIL-1Ra (or 2.5 µl XhoI for mTNFRII-Fc), 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a total volume of 100 µl and incubated at 37°C overnight. The digested and purified PCR products were cloned into the BamHI/SalI site within the pLV.CMV cloning vector in a standard ligation reaction and the subsequent purified maxiprep DNA from positive DH5α transformants were named pLV.CMV.hIL-1Ra or pLV.CMV.mTNFRII-Fc. Each DNA construct (25 µg) was subsequently digested with 2.5 µl PmeI, 2.5 µl BamHI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a total volume of 100 µl and incubated at 37°C overnight. The DNA fragments were separated by

agarose gel electrophoresis and the pLV.hIL-1Ra fragment (6931 bp) and the pLV.mTNFRII-Fc fragment (7756 bp) were isolated and purified from agarose using the PureLink® Gel Extraction Kit (Invitrogen) to generate the pLV.hIL-1Ra and pLV.mTNFRII-Fc cloning vectors, respectively (Fig 5.1 B-E, section 5.1.1).

2.13.7. Construction of LV- synthetic promoters expressing hIL-1Ra or mTNFRII-Fc

The candidate synthetic promoters and SFFV promoter were PCR amplified from their corresponding pCpG-constructs and pUCL-Luc$^+$ DNA respectively, using primers with SnaBI and BamHI restriction enzyme sites (Table 2.10). The PCR reaction comprised 1 µl of diluted plasmid DNA (10 ng/µl), 2.5 µl 10 µM SnaBI forward primer, 2.5 µl 10µM BamHI reverse primer, 5 µl 2 mM dNTPs, 10 µl of 5x Phusion HF Buffer, 0.5 µl of Phusion DNA Polymerase and nuclease-free water in a total reaction volume of 50 µl. The reaction was performed using the thermocycler program: heated lid at 110°C, initial denaturation at 98°C for 5 minutes, 30 cycles at 98°C for 10 seconds, 60°C for 30 seconds, 72°C for 30 seconds and a final extension at 72°C for 5 minutes. The purified DNA (50 µl) was digested with 2.5 µl SnaBI, 2.5 µl BamHI, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a total volume of 100 µl and incubated at 37°C overnight. The digested and purified PCR products were cloned into the compatible PmeI/BamHI site within the pLV.hIL-1Ra and pLV.mTNFRII-Fc cloning vectors to generate the pLV.x.hIL-Ra and pLV.x.mTNFRII-Fc constructs respectively, where x is the designated number of the candidate synthetic promoter or the SFFV promoter (Fig 5.1 F-J, section 5.1.1).

2.13.8. Cloning vector: pLV.SFFV.Luc$^+$

The pLV.SFFV.Luc$^+$ DNA (25 µg) was digested with 2.5 µl EcoRI, 2.5 µl KpnI-HF, 10 µl 10x NEB Buffer 4 and sterile dH$_2$O in a total volume of 100 µl and incubated at 37°C overnight. The SFFV fragment was excised and purified from the agarose gel following electrophoresis to serve as the pLV.SFFV.Luc$^+$ cloning vector (Fig 6.6 A, section 6.5.1).

2.13.9. Construction of lentiviral-miR-23b cassettes

The pcLuc$^+$-miR-23b-2T and pcLuc$^+$-miR-23b-4T plasmid DNAs were diluted to 10ng/µl and the miR-23b-target sites were PCR amplified using primers with MfeI and KpnI restriction enzyme sites. The individual reactions comprised 1 µl of each diluted DNA, 2.5 µl 10µM MfeI forward primer, 2.5 µl 10 µM KpnI reverse primer, 5 µl 2 mM dNTPs, 10 µl of 5x Phusion HF Buffer, 0.5 µl of Phusion DNA Polymerase and nuclease-free water in a total reaction volume of 50 µl. The reaction was performed using the thermocycler program: heated lid at 110°C, initial denaturation at 98°C for 5 minutes, 30 cycles at 98°C for 10 seconds, 50°C for 60 seconds, 72°C for 30 seconds and a final extension at 72°C for 10 minutes. The purified DNA (50 µl) was digested with 2.5 µl MfeI, 2.5 µl KpnI, 10 µl 10x NEB Buffer 4 and sterile distilled water in a total volume of 100 µl and incubated at 37°C overnight. The digested and purified PCR products were cloned into the compatible EcoRI/KpnI site within the pLV.SFFV.Luc$^+$ cloning vector to generate LV-SFFV-miR-23b-2T and LV-SFFV-miR-23b-4T (Fig 6.6 B and C, section 6.5.1).

2.14. Production and titration of lentiviral particles

2.14.1. Production of VSV-G pseudotyped lentiviral particles

All lentiviral work was performed following the safety guidelines of Queen Mary University and experiments were conducted in a designated lentivirus room using a ScanLaf Mars Safety Class 2 Hood. The recombinant lentiviral particles (LPs) were produced using a three-plasmid transient co-transfection into 293T cells which included the constructed recombinant lentiviral transfer plasmids, the gag-pol encoding packaging plasmid pCMVΔR8.2 (kindly provided by Inder Verma, Salk Institute, La Jolla, CA; Naldini *et al.,* 1996b) and the vesicular-stomatitis virus glycoprotein (VSV-G) envelope encoding plasmid pMD.G (Naldini *et al.,* 1996a). Prior to the day of transfection, 9×10^6 293T cells were seeded in 30 ml complete DMEM medium in

a 15 cm^2 tissue culture dish and incubated overnight at 37°C in 5% CO_2. After 24 hours, a DNA solution was prepared by combining 18 µg transfer plasmid, 18 µg pCMVΔR8.2, 4 µg pMD.G with OptiMEM medium (Invitrogen) to a final volume of 1 ml. Polyethyleneimine (PEI; Sigma-Aldrich) was diluted by combining 200 µl PEI and 800 µl of OptiMEM medium. The DNA:PEI complex was formed by incubating the DNA solution with the diluted PEI for 10 minutes at room temperature. One hour before transfection, a final concentration of 25 µg/ml chloroquine was added to the cells to inhibit DNA degradation by lysosomes thereby improving the transfection efficiency. Subsequently, the transfection mix (2 ml) was added in a drop-wise manner to the 293T cells and incubated at 37°C in 5% CO_2. After 16 hours, the medium was carefully removed from the cells and discarded in diluted Virkon® disinfectant and 30 ml of complete DMEM medium was added to the cells. Following 72 hours post-transfection, the medium containing the packaged LPs was collected and the cell debris was removed from the medium by filtration through a 0.45 µM low protein binding filter (Fisher Scientific) and transferred into a centrifuge tube. The LPs were concentrated and collected by ultracentrifugation in a Beckman XL-90 Ultracentrifuge at 23,000 rpm for 2 hours at 4°C. The LP pellets were resuspended in 300 µl of DMEM medium (alone), aliquoted in 20 µl volumes in sterile cryovials and then rapidly stored at -80°C until further use.

2.14.2. Lentiviral particle titration: p24 ELISA assay

The Lenti-X™ p24 Rapid Titer Kit (Clontech Laboratories, CA, USA) was used to quantify the p24 antigen in the concentrated lentiviral preparations as per the manufacturer's instructions. Briefly, the LPs were diluted 1:20,000 in DMEM and a series of five p24 standard dilutions were generated from 200 pg/ml -12.5 pg/ml. Lysis buffer (20 µl) was added to each anti-p24 coated well. A volume of 200 µl of diluted lentivirus, each standard curve dilution and DMEM

medium (negative control) was dispensed into each well, and incubated at 37°C for 60 minutes. The contents of the wells were aspirated and manually washed six times with 1x Wash Buffer (provided in the kit as 20x Wash Buffer) and thoroughly dried by firmly inverting the plate on an absorbent paper towel. The wells were incubated with 100 µl anti-p24 biotin conjugated detection antibody at 37°C for 60 minutes after which the contents of the wells were removed. The wells were washed and dried as previously described and 100 µl Streptavidin-HRP conjugate was dispensed into the wells and incubated at room temperature for 30 minutes. Following the subsequent removal of the well contents, washing and drying of the wells, 100 µl of Substrate Solution was immediately dispensed into the wells and incubated at room temperature for 20 minutes. The reaction was stopped by adding 100 µl of Stop Solution to each well and the absorbance values were immediately read at 450 nm using a microtiter plate reader (Tecan, Magellan 4 Software). The p24 values of the LPs were determined using the p24 standard curve and the LPs within the given concentration of p24 were determined using the formulae below:

- 1 LP contains 8×10^{-5} pg of p24 (derived from $(2000) \times (24 \times 10^{3} \text{Da}) / (6 \times 10^{23})$)

- 1 ng p24 is equivalent to $\sim 1.25 \times 10^{7}$ LP

- For a typical lentivirus vector, there is 1 IFU for every 100-1000 LP

- Therefore, a supernatant titre of 10^{7} IFU/ml $\approx 10^{9} - 10^{10}$ LP/ml or 80-800 ng p24/ml

2.14.3. Generation of stable cell lines

Prior to the day of transduction, 5×10^{4} 293T or NIH3T3 cells were seeded in 3 ml of complete DMEM medium in a 6-well plate. After 24 hours, 15µl of lentiviral particles was added to the cells containing 6 µg/ml polybrene (Sigma-Aldrich). After 72 hours, the cell medium was removed and discarded in diluted Virkon® disinfectant and the adherent cells were carefully

washed with PBS. The cells were trypsinised using the standard protocol and transferred into a 75 cm^2 flask containing 10ml of complete DMEM medium.

2.14.4. Stimulation of stable cell lines

Approximately $2x10^4$ stable 293T or NIH3T3 cells were seeded in a 96-well plate in 100 µl of complete DMEM medium. After 24 hours, the cell medium was removed from the stable 293T cells and replaced with 0.5% FBS DMEM alone (unstimulated) or incubated in hypoxia   (0.1% O_2) or 0.5% FBS DMEM containing 10 ng/ml TNFα, 10 ng/ml PMA, or a combination of TNFα, PMA and hypoxia for 18 hours.

At the same time point, the cell medium was removed from the stable NIH3T3 cells and replaced with 0.5% FBS DMEM alone or 0.5% FBS DMEM containing 50 ng/ml mIL-17A,   10 ng/ml TNFα or a combination of mIL-17A with TNFα for 18 hours.

After 18 hours, the cell medium was aspirated and replaced with 50 µl 1x Glo Lysis Buffer (Promega) to lyse the cells. The luciferase gene expression and total protein content of the lysates were quantified as described in section 2.2. The firefly luciferase expression was normalised to the titre of the corresponding LPs (lenti IFU/µl) and protein content in the cell lysate (mg/ml). In contrast, the therapeutic protein expression was quantified by a sandwich enzyme-linked immunosorbent assay (ELISA).

2.15. ELISA quantification of therapeutic protein expression

The therapeutic protein expression of hIL-1Ra and mTNFRII-Fc in the cell supernatant from unstimulated or inflammatory and/or hypoxic stimulated stable 293T cells was quantified using sandwich ELISA.

2.15.1. mTNFRII-Fc ELISA

For mTNFRII-Fc protein quantification, each well of the high affinity protein-binding 96-well ELISA plate (Nunc Maxisorp®, eBioscience, CA, USA) was coated with 100 µl of rat monoclonal anti-TNFRII antibody (ab7369; Abcam Plc, Cambridge, UK) diluted 1:200 in 100 mM bicarbonate/carbonate coating buffer (3.03g Na_2CO_3, 6g $NaHCO_3$ in 1L dH_2O final volume, pH 9.6). The plate was sealed with adhesive plastic to prevent evaporation and incubated overnight at 4°C. The next day, the diluted capture antibody was aspirated from the wells and the wells were washed 3 times with 200 µl 1x PBS-Tween (8g NaCl, 0.2g KCl, 1.44g Na_2HPO_4, 0.24g KH_2PO_4 and 0.05% (v/v) Tween-20 in 1L dH_2O, pH 7.4) dried thoroughly by firmly blotting against clean paper towels. Non-specific binding sites in the coated wells were blocked by adding 200 µl of blocking buffer (4% Marvel dried skimmed milk in 1x PBS) to each well. The plate was sealed and incubated at room temperature for 2 hours. The blocking buffer was removed from the wells, which were subsequently washed 4 times with 200 µl 1x PBS-Tween and dried thoroughly. Serial dilutions of mTNFRII-Fc standards from 374 ng/ml – 3.74 pg/ml were generated in 0.5% FBS DMEM medium and 100 µl of each mTNFRII-Fc sample and the diluted mTNFRII-Fc standards were added to duplicate wells. The plate was sealed and incubated at room temperature for 3 hours. Following incubation, the contents of wells were aspirated and the wells were washed 4 times with 200 µl 1x PBS-Tween and dried thoroughly. The biotinylated goat polyclonal secondary detection antibody to mouse IgG (ab7067; Abcam) was diluted 1:1000 in 1xPBS-Tween and 100 µl was added to each well. The plate was sealed and incubated at room temperature for 1 hour after which the wells were washed 4 times with 200 µl 1x PBS-Tween and thoroughly dried. The streptavidin biotinylated horseradish peroxidase complex (890803; R&D Systems, MN, USA) was diluted 1:200 in 1x PBS-Tween and 100 µl was added to each well, the plate was sealed and incubated at room temperature for 1 hour. The wells were washed 7 times with 200 µl 1x PBS-Tween and dried thoroughly and 100 µl of a 1:1 mixture of Peroxidase Substrate Solution B (0.02% H_2O_2 in a Citric Acid Buffer; #506500; KPL Inc, MO, USA) and TMB Peroxidase Substrate (0.4 g/L

3,3',5,5'- tetramethylbenzidine in an organic base; #507601; KPL Inc) was added to each well. The plate was sealed and incubated at room temperature for 30 minutes in the dark. The reaction was stopped by adding 100 µl 4M sulphuric acid into each well and the absorbance values were immediately read at 450 nm using a microtiter plate reader (Tecan, Magellan 4 Software).

2.15.2. hIL-1Ra ELISA

The hIL-1Ra protein concentration in the hIL-1Ra stable cell supernatants was quantified using the ELISA protocol described in the preceding section 2.15.1, with the modifications required for hIL-1Ra detection:

- Recombinant human IL-1Ra/IL-1F3 monoclonal capture antibody (MAB280; R&D) was diluted 1:100 with 100mM bicarbonate/carbonate coating buffer.
- Serial dilutions of hIL-1Ra standards from 40 ng/ml –0.4 pg/ml were generated in 0.5% FBS DMEM medium and 100 µl of each hIL-1Ra sample and the diluted hIL-1Ra standards were added to duplicate wells.
- Recombinant human IL-1Ra/IL-1F3 biotinylated polyclonal goat secondary detection antibody (BAF280; R&D) was diluted 1:200 in 1xPBS-Tween.

Standard curves of mTNFRII-Fc and hIL-1Ra protein concentration versus absorbance at 450 nm were plotted in GraphPad® Prism 5. The linear equation was used to calculate the mTNFRII-Fc and hIL-1Ra protein concentration in the samples, which was normalised to the lentiviral titre of their respective LPs.

2.16. *In vivo* procedures

2.16.1. Animals

Adult male CD1 mice (6-8 week old, 21-25g) were treated according to the approved UK Home Office and institutional guidelines.

2.16.2. Intraplantar delivery of lentivectors into mouse hind paws

The mice were anaesthetised with AErrane (Isofluorane; Baxter Healthcare Ltd, Thetford, Norfolk) using Boyle's apparatus (British Oxygen Company BOC, London, UK). The hind paws were sprayed with 70% ethanol to disinfect the site and 25 µl of lentiviral particles containing 260,000 lentiviral IFU of luciferase lentivectors (LV-2-Luc+, LV-9-Luc+, LV-12-Luc+, LV-4NFκB-Luc+, LV-mCMV-Luc+, LV-SFFV-Luc+) were delivered into both hind paws.

For the subsequent experiments which delivered therapeutic lentivectors, 25 µl of lentiviral particles containing 830,000 lentiviral IFU of therapeutic lentivectors were delivered into the left hind paws by intraplantar injection and an equivalent volume of sterile saline was injected into the control right hind paw using the same method. The lentivectors were allowed to integrate into the cell genome for 7 days.

2.16.3. Induction of carrageenan-induced paw inflammation

Seven days post-lentiviral injection (luciferase or therapeutic LPs), paw inflammation was induced by an intraplantar injection of 50 µl of 1% λ-carrageenan solution (0.1g λ-carrageenan powder (Sigma-Aldrich) in 10 ml sterile saline) into the left hind paws. The control right hind paws were injected intraplantarly with sterile saline.

2.16.4. Real-time bioluminescence imaging of luciferase gene expression

At 0, 3, 24 and 72 hours post-carrageenan injection, the mice were injected intraperitoneally with 200 µl of luciferin K^+ Salt (30 mg/ml; Promega) and anaesthetised with Isofluorane using Boyle's apparatus. After exactly 15 minutes, the anaesthetised mice were photographed (0.2-second exposure) and imaged for light emission (5 minutes on medium sensitivity) with the IVIS® Lumina II (Caliper Life Sciences Corp, Hopkinton, MA, USA). Bioluminescence images were overlaid on gray-scale photographs which were obtained with a 12-cm field of view, a binning of 8 and a 1/ *f* stop and open filter. The regions of interest (ROI) were defined manually over both hind paws and the background photon flux was defined in control regions of the same size. Light emission was quantified as photons per steradian per square centimetre (photons/second/cm^2/sr) using Living Image® Software (Caliper Life Sciences Corp.).

2.16.5. Paw caliper measurements

The thickness of hind paws was measured at 0, 3, 24, 48, 72 and 96 hours post-carrageenan injection using POCO 2T calipers (Kroeplin Längenmesstechnik, Schlüchtern, Germany). After 96 hours post-carrageenan injection, the mice were terminated by cervical dislocation.

2.17. Statistical analyses

Comparisons between the specified experimental and control groups was calculated using the two-tailed, unpaired Student's t-test (Microsoft Excel 2007, Microsoft Corp., Redmond, WA, USA), unless stated otherwise. Values for $p \leq 0.05$ were considered statistically significant. Graphs were drawn using GraphPad® Prism 5, where the numbers (or titles) on the X-axis correspond to the individual constructs and the data represents the mean ± SD of triplicate values, unless stated otherwise.

CHAPTER 3:

Transcriptional Regulation of Luciferase Gene Expression during Disease Activity:

Design and initial *in vitro* selection of transcriptionally regulated promoters

3.1. Introduction

Rheumatoid arthritis (RA) is a chronic, systemic inflammatory disease which predominantly affects the synovial joints. Pro-inflammatory cytokines, such as TNFα and IL-1β, coordinate many of the inflammatory and destructive processes in RA joints. In addition, RA joints are often hypoxic, where low oxygen concentrations initiate angiogenesis which facilitates the infiltration of immune cells and provides oxygen and nutrients to the hypoxic tissue, promoting further overgrowth of the synovium. The inflammatory and hypoxic environment activates various transcription factors (TFs) such as NFκB and HIF-1α respectively, resulting in the transcription of pro-inflammatory mediators that further exacerbate chronic inflammation within the RA joint.

The endogenous activity of inflammation and/or hypoxia-responsive TFs can be exploited by constructing gene therapy vectors with composite synthetic promoters, containing the binding sites of responsive TFs (TFBSs). In this way, therapeutic gene expression can be controlled at the transcriptional level through interactions of TFs with their TFBSs within the composite synthetic promoter to create inflammation-inducible, multi-responsive and synergistically-inducible expression systems. Localised delivery of inflammation-regulated therapy into the RA joint would allow high expression of the therapeutic proteins during disease flare and minimal therapeutic protein concentrations during periods of remission, as schematically depicted in Figure 3.1.
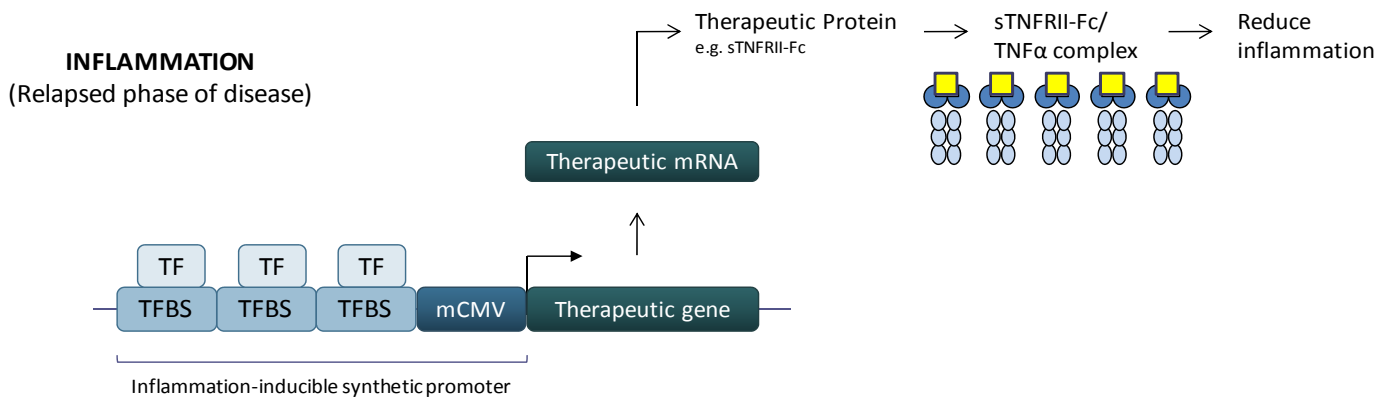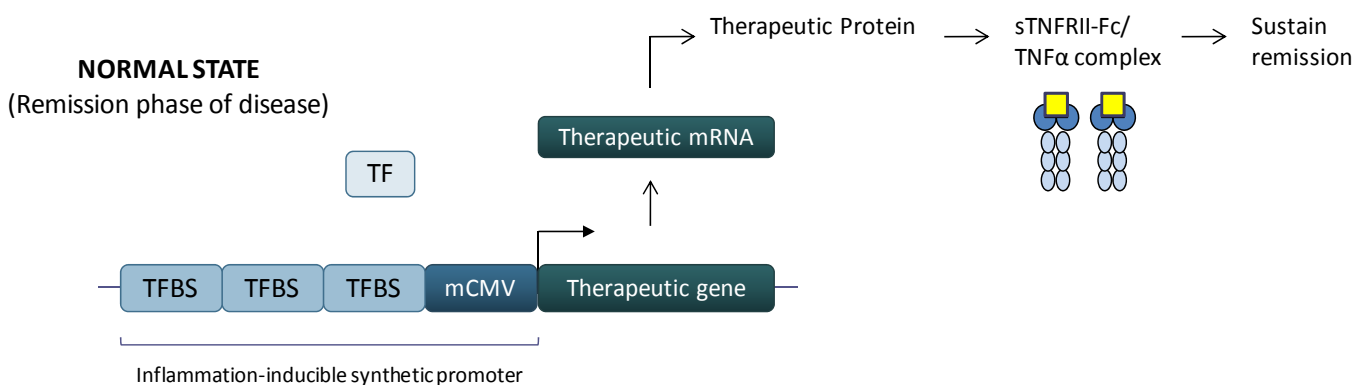
Figure 3.1. Inflammation-inducible synthetic promoters can regulate therapeutic protein expression in response to the level of disease activity in RA joints. During the relapsed phase of RA, the joint is significantly inflamed and often hypoxic, resulting in the activation of responsive TFs. Upon activation, the TFs bind to their cognate binding sites within the synthetic promoter to initiate transcription of the downstream therapeutic gene e.g. soluble TNFRII-Fc (sTNFRII-Fc). The resulting sTNFRII-Fc proteins can bind to TNFα to form sTNFRII-Fc/TNFα complexes which reduce the concentration of free TNFα within the joint to prevent TNFα-mediated signalling and consequently decrease the level of inflammation [A]. In contrast, during the remission phase of RA, the level of inflammation/hypoxia is substantially lower resulting in low level activation of the responsive TFs. Subsequently, the therapeutic protein will be expressed at a low level to sustain remission [B].

The aim of the work described in Chapter 3 is to verify the concept of inflammation-inducible regulation of gene expression by composite synthetic promoters. The level of firefly luciferase gene (Luc$^+$) expression serves as a surrogate marker for the expected expression levels of therapeutic genes. Therefore, the initial *in vitro* functional analysis of promoter activity was correlated to the magnitude of luciferase gene expression, thus serving as a readout for the activity of the synthetic promoter.

Due to the presence of different TFBSs within the composite promoters, I predicted that the composite promoters would be multi-responsive to different individual stimuli and additively or synergistically-inducible in response to combined inflammatory and hypoxic stimulation. This unexplored concept for RA gene therapy is evaluated in Chapter 3.

3.2. Characterisation and functional analysis of inflammation-inducible pGL3-4bp-composite synthetic promoters

3.2.1. Construction of the pGL3-4bp-composite synthetic promoter library using the Random Ligation Cloning Method

The schematic diagram below (Fig 3.2) depicts the cloning procedure used to generate the pGL3mCMV-composite synthetic promoter constructs.

A. pGL3 Basic construct was digested with XhoI and HindIII to serve as the cloning vector. The annealed mCMV oligonucleotides with 5'- XhoI and 3'- HindIII overhangs were cloned between these sites, upstream of the luciferase gene within the pGL3 Basic vector to generate pGL3mCMV (Fig 3.2 A) (DNA sequence in Appendix 2).
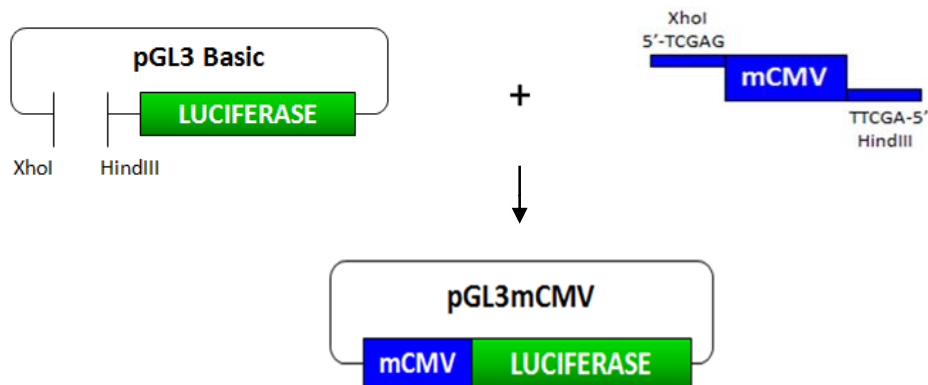
B. pGL3mCMV was digested with NheI. The 5' and 3' ends of pGL3mCMV were dephosphorylated by incubating the linear DNA with Calf Intestinal Phosphatase (CIP) enzyme to prevent recircularisation of the pGL3mCMV cloning vector (Fig 3.2 B).
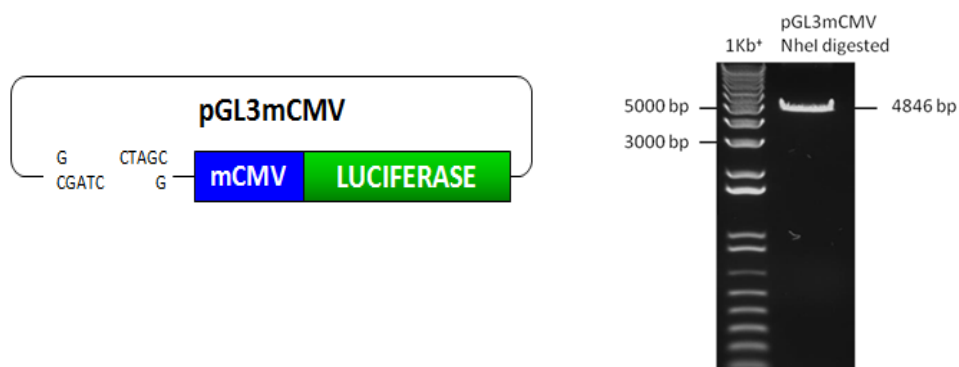
C. Oligonucleotides containing the candidate TFBSs were ordered with phosphorlyated 5'-CTAG overhangs, which were compatible to the NheI overhang. The six annealed TFBS-oligonucleotides were randomly ligated (in tandem) upstream of the mCMV promoter within the pGL3mCMV vector to generate a library of constructs with varying composition and numbers of TFBSs (separated by a 4bp space) within the composite promoters (Fig 3.2 C). The selected pGL3-4bp-composite promoters were sequenced (Appendix 3).
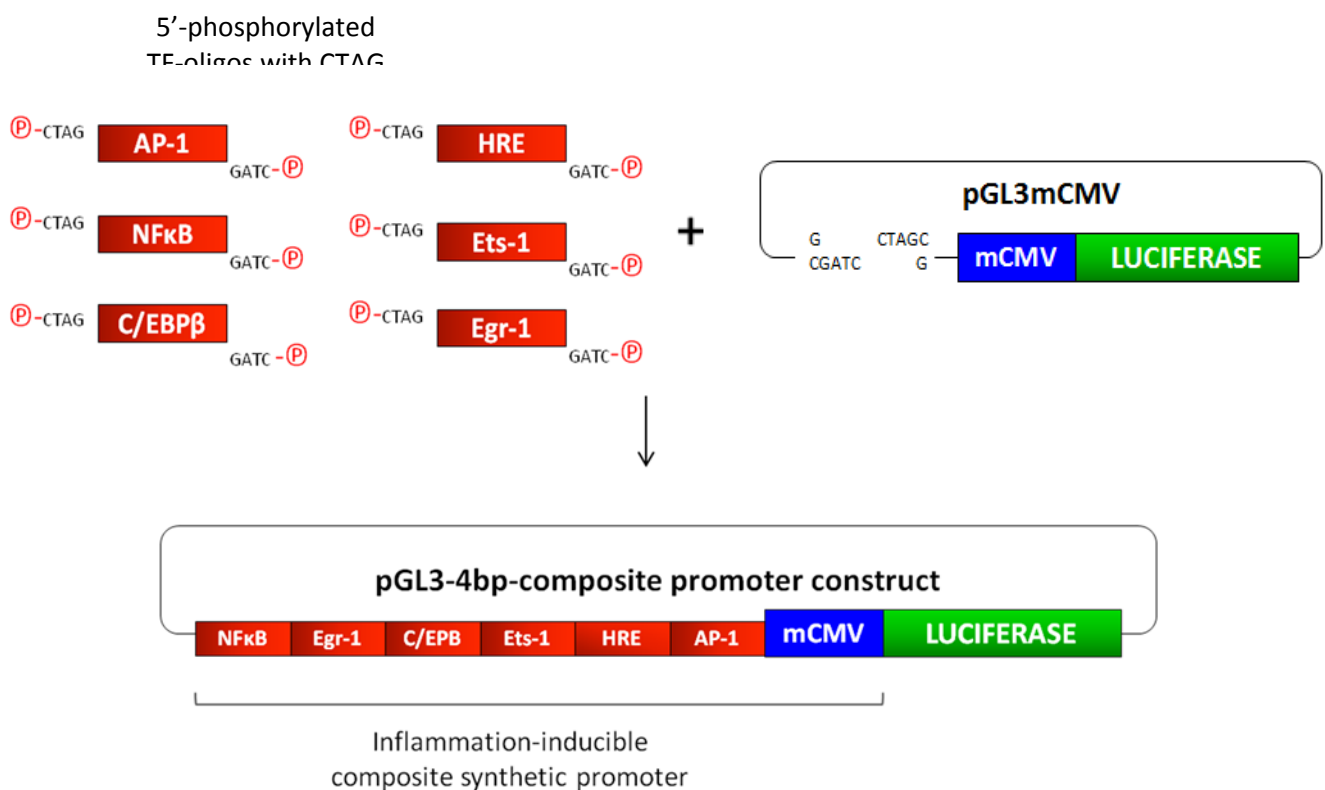
Figure 3.2. Schematic representation of the random ligation cloning method used to generate the pGL3-4bp-composite synthetic promoters. The pGL3 Basic construct was systematically modified to generate a library of constructs containing composite synthetic promoters with diverse TFBSs, separated by a 4bp space.

3.2.2. Characterisation and functional analysis of pGL3-4bp-composite synthetic promoters

To examine the concept of inflammation-regulated, multi-responsive and synergistically-inducible composite synthetic promoters, a library of 250 pGL3-4bp-composite promoters were generated and transfected into 293T cells. The cells were treated with a combined stimulus of hIL-1β (10 ng/ml) and hypoxia (0.1% O_2) to induce all candidate TFs and to mimic the inflammatory and hypoxic microenvironment within the RA joint. This preliminary screening of synthetic promoter activity identified the most inducible promoters which displayed moderate/high fold inductions in response to combined hypoxia and hIL-1β stimulation (n=25).

The functionality of the selected 25 promoters was further characterised by co-transfecting the plasmid DNA constructs with pRL-CMV in 293T cells. Transfected 293T cells were unstimulated, incubated in hypoxia (0.1% O_2) or stimulated with TNFα (10 ng/ml), PMA (10 ng/ml) or their combination (TNFα and PMA, in hypoxia). The normalised firefly luciferase values are presented in Figure 3.3 and fold inductions are presented in Figure 3.4.
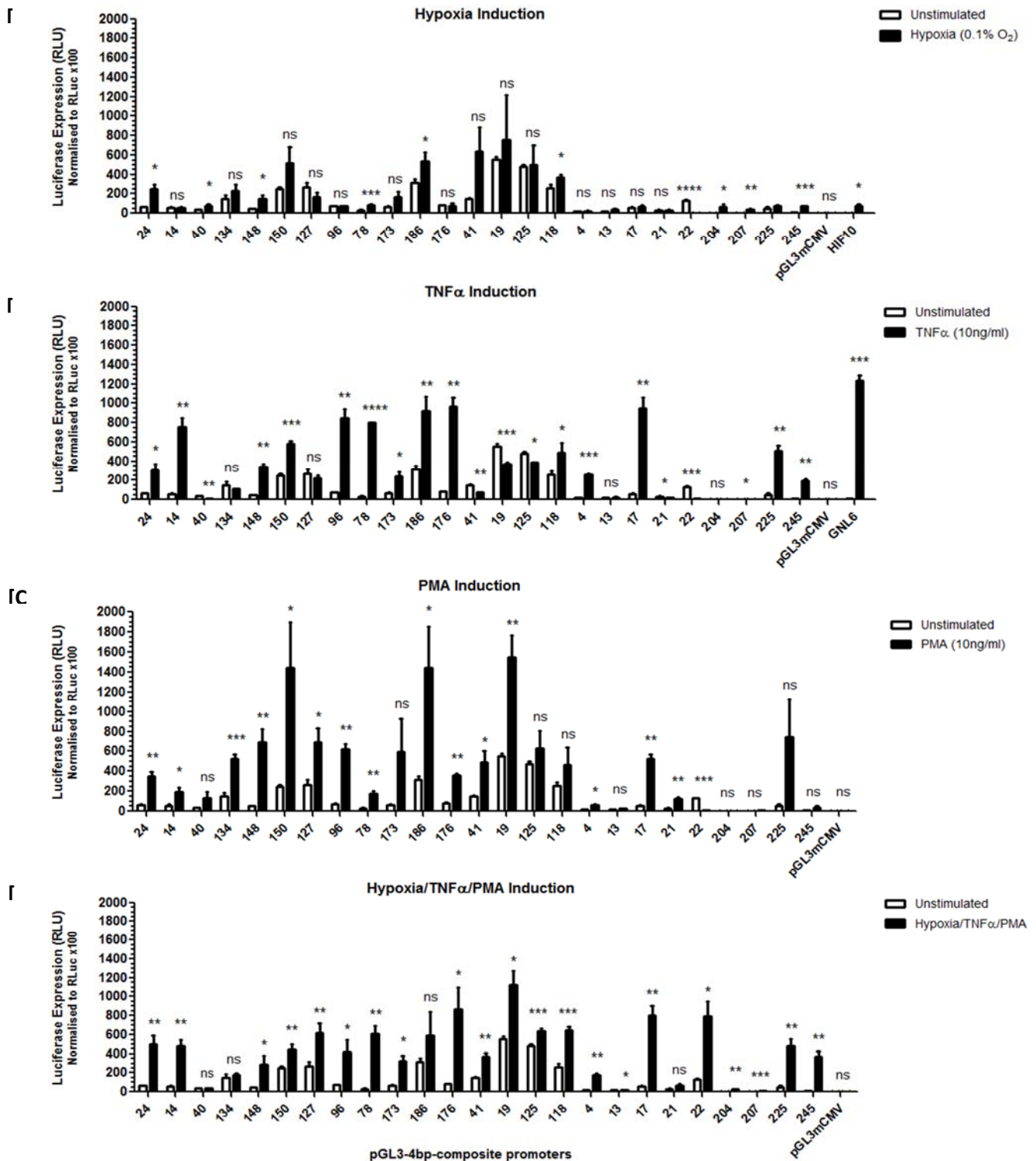
Figure 3.3. Functional characterisation of the pGL3-4bp-composite promoters. 293T cells (20,000) were seeded in a 96 well plate. After 24 hours, the selected pGL3-composite promoter constructs (n=25; 180 ng) were co-transfected with pRL-CMV (20 ng) into 293T cells. The cells were unstimulated, incubated in hypoxia [A], stimulated with TNFα [B], PMA [C] or their combination [D] for 18 hours. Transfections were carried out in triplicate and normalised to renilla luciferase from pRL-CMV. The hypoxia responsive construct pGL3-HIF10 and inflammation-responsive GNL6 were included in
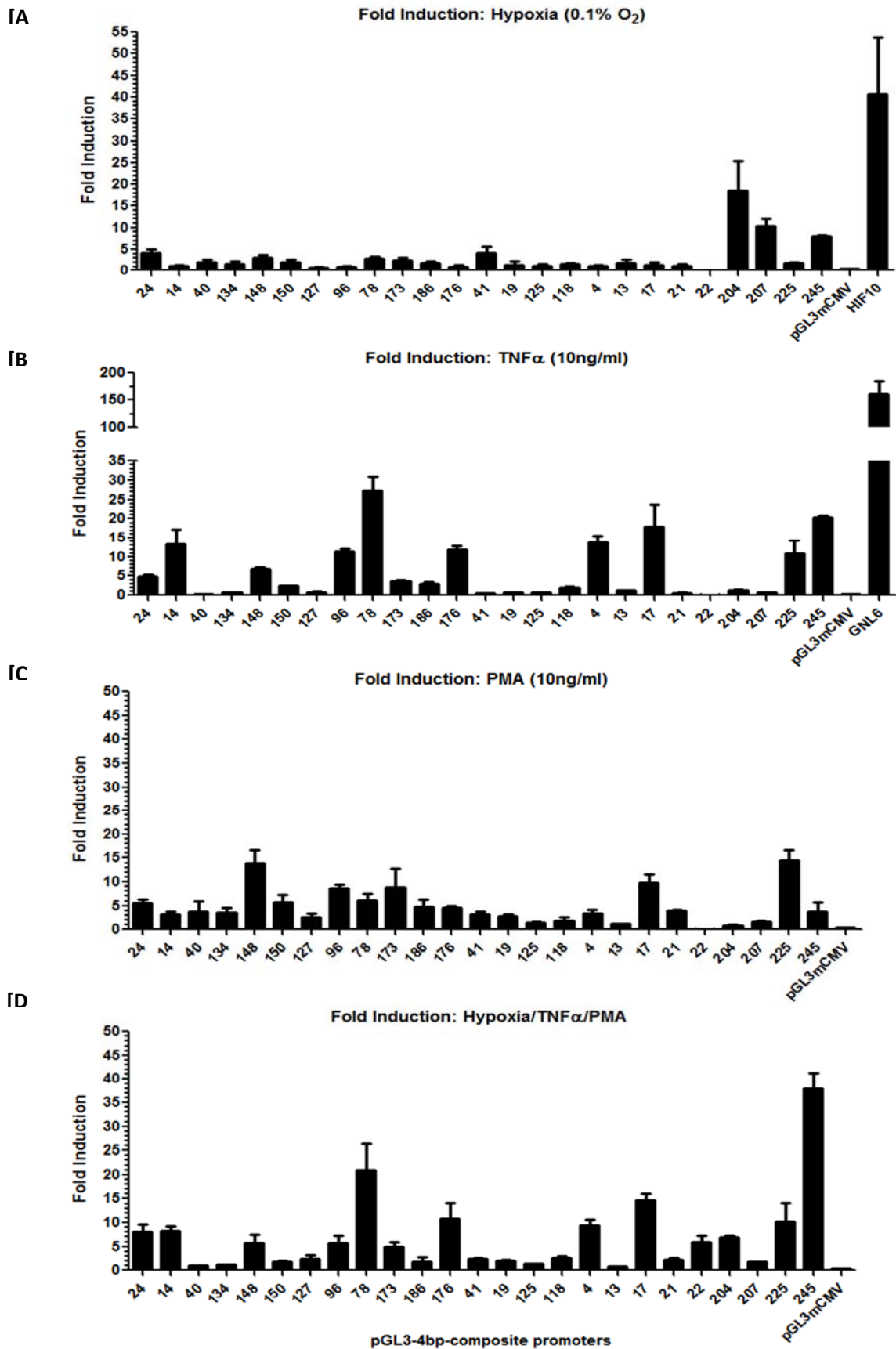
Figure 3.4. Fold inductions of the pGL3-4bp-composite promoters. The fold inductions were calculated by dividing the normalised induced luciferase values by the uninduced luciferase

Figure 3.3 confirmed the multi-responsiveness of the composite synthetic promoters, which was due to the presence of diverse TFBSs. Interestingly, when visually compared to the pGL3mCMV negative control construct, many of the synthetic promoters displayed high basal luciferase expression levels which were sometimes comparable to their induced luciferase gene expression (Fig 3.3). Consequently, such promoters displayed modest fold inductions (Fig 3.4).

DNA sequence analysis of the composite promoters confirmed that the responsiveness of the promoter to a particular stimulus was often reflected by its sequence composition. For example, synthetic promoters comprising the binding sites of inflammation-responsive TFs were significantly induced by TNFα and PMA inflammatory stimulation. However, it appeared that the presence of HIF-1α binding sites (HRE motifs) within the composite promoters did not guarantee hypoxia-responsiveness. For example, many composite promoters possessed HRE motifs (flanked by other TFBSs) yet failed to respond to hypoxic stimulation i.e. pGL3-14, -96, -176 (ns, $p > 0.05$). In contrast, the synthetic promoters which possessed only a single HRE motif were highly responsive to hypoxia i.e. pGL3-204 (* $p=0.04$) and pGL3-207 (** $p=0.006$), which displayed 18.6 fold and 10.3 fold induction, respectively (Fig 3.4 A).

Further comparisons between the level of hypoxia induction and the DNA sequences of the composite promoters revealed that the promoters unresponsive or marginally responsive to hypoxia often had HRE motifs flanked by the other TFBSs whereas the hypoxia-responsive promoters had HRE motifs proximal to or positioned close to the TATA box. For example, pGL3-204 (* $p=0.04$), and -207 (** $p=0.006$) possessed single proximal HRE motifs and pGL3-245 (*** $p=0.001$) was comprising of two HRE motifs and a pair of proximal NFκB sites and these promoters exhibited significantly greater hypoxia-responsiveness than the vast majority of composite promoters with HRE motifs flanked by other TFBSs (Fig 3.4 A) which suggests that the positioning of the HRE motif relative to the TATA box may be an important determinant of hypoxia-inducibility.

As anticipated, the vast majority of the composite synthetic promoters possessed the binding sites of inflammation-responsive NFκB, AP-1, C/EBPβ, Egr-1 and Ets-1 transcription factors. Consequently, these promoters were highly responsive to TNFα with fold inductions of upto 27 fold (displayed by pGL3-78, **** $p<0.0001$) and robust PMA inducibility i.e. pGL3-17, -96 and -148 (all ** $p< 0.01$) displayed 9.7, 8.5 and 13.9 fold induction in response to PMA, respectively (Fig 3.4 C). Interestingly, the DNA sequencing revealed that these promoters possessed either a proximal AP-1 or NFκB motif, which implies that the proximal TFBS influences the general responsiveness of the promoter to the corresponding stimulus (Fig 3.3 C and 3.4 C).

To assess the capability of additive/synergistic induction of luciferase gene expression from the composite promoters, the transfected 293T cells were stimulated with TNFα, PMA and hypoxia, to mimic the inflammatory and hypoxic environment within the RA joint. Interestingly, only pGL3-22 and -245 displayed synergistic luciferase expression whereas the other composite promoters (n=23) exhibited significant increases in luciferase gene expression following combined stimulation but unchanged or lower luciferase gene expression levels compared to when stimulated with a single stimulus (Fig 3.3 D and 3.4 D). This observation highlighted the possibility of spatial constraints of TFBSs which may have arisen due to the close proximity of the TFBSs (4 bp space). It is likely that during combined inflammatory and hypoxic stimulation, the binding of the TF to its binding site may have sterically hindered the binding of other TFs to the adjacent binding sites which consequently, impaired additive/synergistic induction.

Overall, the promoters were generally multi-responsive, however, the fundamental requirements of low basal and additive/synergistic induction of gene expression (to combined inflammatory and hypoxic stimulation) was not achieved by the majority of the composite promoters, which highlighted the need for an alternative, systematic cloning method to construct composite synthetic promoters with improved expression profiles.

3.3. Construction of inflammation-inducible composite synthetic promoters using the Assembly PCR Cloning Method

The important questions to emerge from the data in section 3.2 was whether the spacing between the TFBSs and also the distance between the proximal TFBS and the TATA box had any effect on the level of gene expression induced by the synthetic promoters. This was investigated using the Assembly PCR cloning method, which is a method traditionally used to synthesise genes from multiple short oligonucleotides. Increased spacing was introduced between the TFBSs relative to one another and also relative to the TATA box whilst retaining the random arrangement of the TFBSs within the composite synthetic promoters. This high-throughput and rational cloning approach represented an attractive alternative to the random ligation cloning method. The concept of the Assembly PCR cloning method is schematically depicted in Figure 3.5.

3.3.1.   Concept of the Assembly PCR cloning method

The Assembly PCR method was adapted to construct composite synthetic promoters with the binding sites of inflammation and/or hypoxia-responsive transcription factors (Fig 3.5)
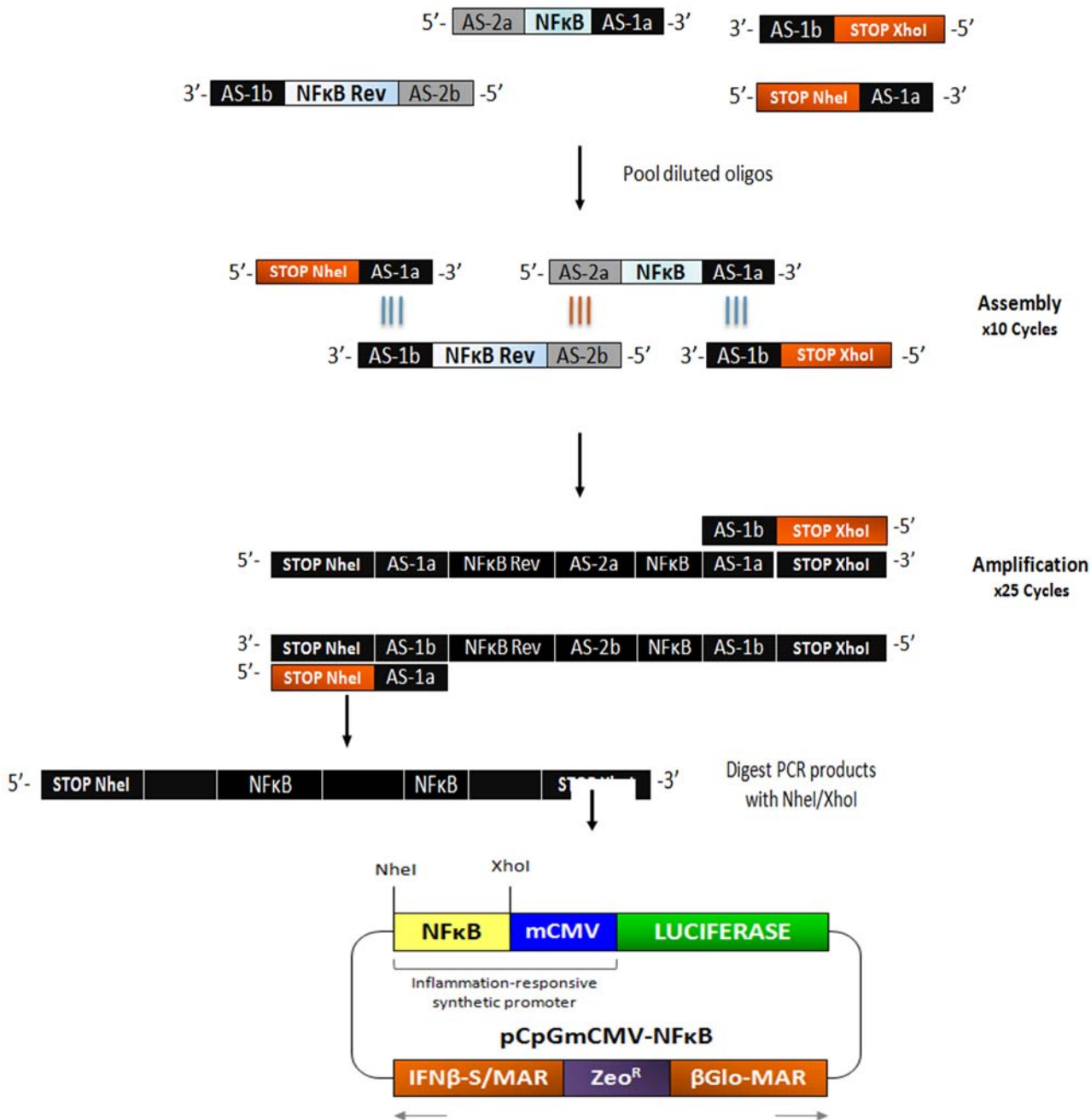
Figure 3.5. Schematic diagram illustrating the construction of inflammation-inducible synthetic promoters using the Assembly PCR cloning method.

The oligonucleotides used to generate the sense strand of the PCR product contained a TFBS flanked by two annealing sequences (AS-2a and AS-1a) whereas the anti-sense strand of the

PCR product consisted of oligonucleotides containing the reverse sequence of the TFBS flanked by the complementary annealing sequences (AS-2b and AS-1b). The 5'-'stop' oligonucleotide contained the NheI restriction enzyme site whereas the 3'-'stop' oligonucleotide contained the XhoI restriction enzyme site, for downstream cloning purposes.

The forward and reverse TFBS-oligonucleotides were pooled together with the 'stop' oligonucleotides and diluted. These oligonucleotides were assembled in a 10 cycle assembly PCR reaction to generate double stranded DNA which served as the template DNA for the subsequent 25 cycle amplification PCR reaction. The PCR products were then digested with NheI and XhoI and cloned upstream of the mCMV promoter within the pCpGmCMV vector to generate inflammation and/or hypoxia-inducible synthetic promoters.

3.3.2. Optimisation of the Assembly PCR method

There are several limitations of the Assembly PCR method which hinder the success of the cloning method. These constraints were overcome by designing unflawed oligonucleotides and optimising the oligonucleotide concentrations to develop a high-throughput method for generating synthetic promoters.

The first imperative step in the design of the Assembly PCR oligonucleotides was to generate annealing and spacer sequences which were devoid of mammalian TFBS. Inadequately designed annealing and spacer sequences containing mammalian TFBS could give rise to promoters with unnecessary TFBSs and therefore, negatively impact the promoter activity. The annealing and spacer sequences were screened using the TRANSFAC® database to verify the absence of mammalian TFBS.

The length of the annealing sequences and the use of spacer oligonucleotides determine the degree of spacing between the TFBS. Therefore, in order to investigate the effect of spacing between the TFBSs on gene expression, the oligonucleotides were designed to contain a TFBS flanked by annealing sequences of specified lengths to generate a library of promoters with varied spacing between the TFBS.

In order to determine the minimal length of the annealing sequence required for successfully assembling and amplifying PCR products using the Assembly PCR reaction, the oligonucleotides with 10 bp, 15 bp and 20 bp annealing sequences were used to generate PCR products which were visualised by agarose gel electrophoresis, shown in Figure 3.6.
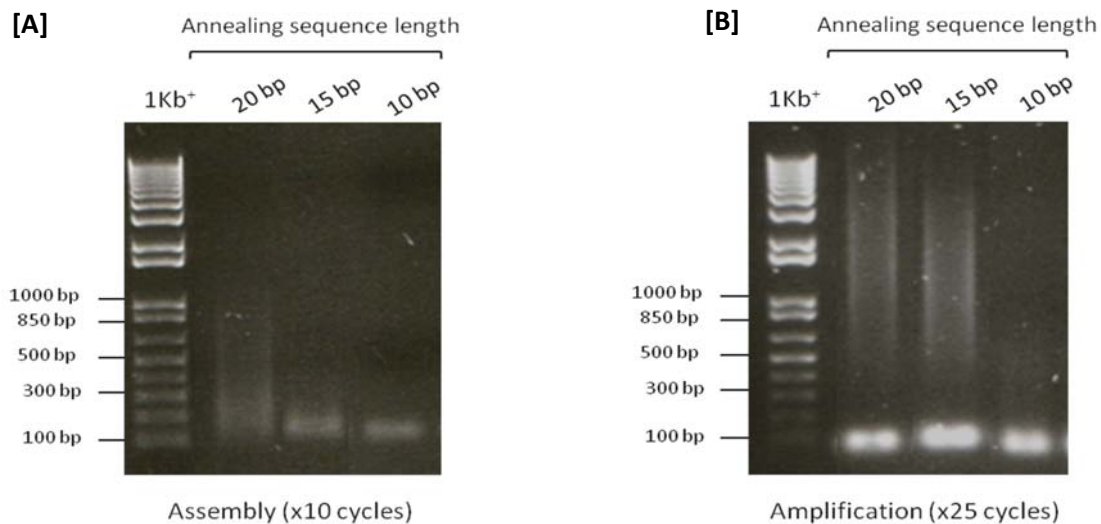
Figure 3.6. The synthesis of HRE-containing PCR products with different annealing sequence lengths. Standard PCR reaction components were combined with 5 µl of 100 µM forward and reverse HRE oligonucleotides (comprised of 20 bp, 15 bp and 10 bp annealing sequences) and 40 nM of 'stop' oligonucleotides, in individual PCR reactions. The oligonucleotides were assembled in a 10 cycle assembly PCR reaction [A] and 15 µl of the assembled dsDNA was amplified in a 25

A successful Assembly PCR reaction generates PCR products of different sizes which can be confirmed by the presence of a smear following gel electrophoresis, as demonstrated in Figure 3.6. The resulting PCR products containing the TFBSs can be isolated and cloned into the expression vector to generate the synthetic promoter construct.

Figure 3.6 B shows that the oligonucleotides with 15 bp and 20 bp annealing sequences formed PCR products that were large enough to isolate from the 'stop'-oligonucleotide primers. However, the PCR products generated using the oligonucleotides with 10 bp annealing sequences were very small and indistinguishable from the 'stop' oligonucleotide primers, which could introduce the risk of contaminating the PCR products with the 'stop'-oligonucleotides during the isolation of the PCR products. Therefore, 10 bp annealing sequences were not sufficient to enable the assembly method and 15 bp or 20 bp annealing sequences were used in the subsequent reactions.

The success of the Assembly PCR method is also heavily dependent on the 'stop' oligonucleotide concentration in the assembly reaction (x10 cycle reaction). These oligonucleotides serve three important functions:

1. Firstly, the 'stop' oligonucleotides prevent further assembly of the PCR product in the x10 cycle assembly reaction thereby truncating the PCR product to prevent indefinite oligonucleotide assembly.

2. Secondly, high concentrations of the 'stop' oligonucleotides (1 µM final) serve as amplification primers in the x25 cycle amplification reaction. The 'stop' oligonucleotides incorporated into the DNA template during the assembly reaction (x10 cycle) serve as binding sites for the amplification primers. The inefficient incorporation of 'stop' oligonucleotides in the initial assembly reaction will lower the efficiency of the subsequent amplification reaction due to the potential absence of the amplification primer binding sites.

3. Thirdly, the 'stop' oligonucleotides contain restriction enzyme sites (5'-NheI and   3'-XhoI) for downstream cloning of the PCR product into the expression vector. The inefficient incorporation of 'stop' oligonucleotides in the assembly reaction (x10 cycle) will also lower the efficiencies of restriction digestion and ligation reactions due to the potential absence of the NheI and XhoI restriction sites and overhangs within the PCR products, respectively, which consequently impairs the following cloning steps.

To confirm the importance of optimal 'stop' oligonucleotide concentrations on the success of the Assembly PCR method, forward and reverse oligonucleotides with HRE motifs, flanked by 20 bp annealing sequence, were randomly pooled together with increasing concentrations of 'stop' oligonucleotides. The resulting PCR products were visualised by gel electrophoresis on an ethidium-bromide containing agarose gel, presented in Figure 3.7.
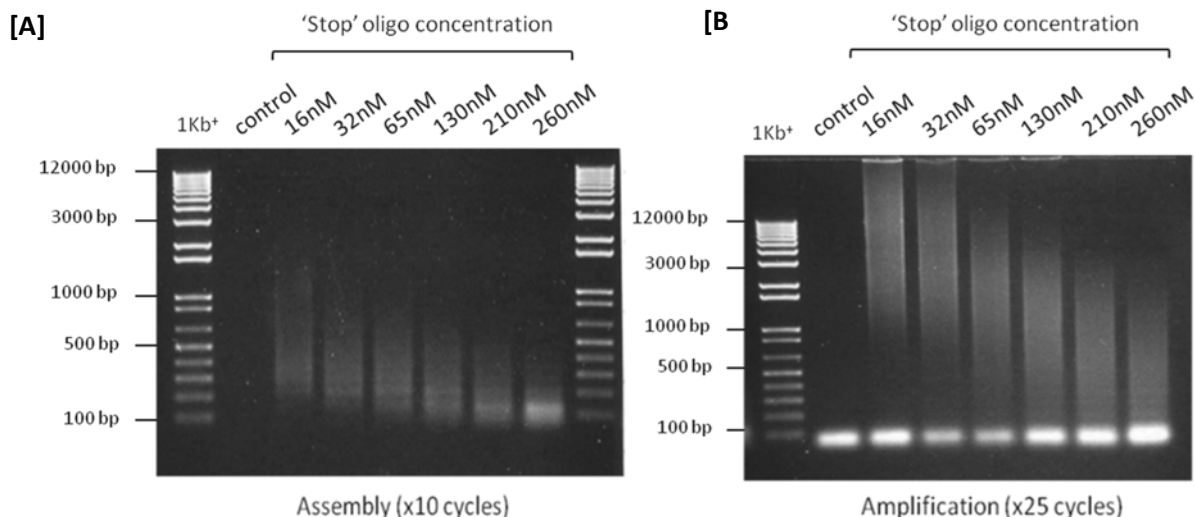
Figure 3.7. The assembly and amplification of HRE-containing PCR products. Standard PCR reaction components were combined with 5 µl of 100 µM forward and reverse HRE oligonucleotides (with 20 bp annealing sequence) and increasing concentrations of 'stop' oligonucleotides. No oligonucleotides were added to the negative control PCR reaction. The diluted HRE and 'stop' oligonucleotides were assembled in a 10 cycle assembly PCR reaction [A]. Fifteen microlitres of the assembled dsDNA was amplified with 1 µM final concentration of 'stop' oligonucleotides and PCR reaction components in a 25 cycle amplification PCR reaction [B]. Five microlitres of assembled [A] and amplified [B] PCR products were visualised by gel electrophoresis on an ethidium bromide containing 0.5 x TAE agarose gel and approximate sizes of the PCR products were determined by comparisons to the 1Kb⁺ DNA ladder.

As anticipated, the size of the PCR products decreased with increasing 'stop' oligonucleotide concentration in the initial assembly reaction (Fig 3.7). The amplified PCR products containing 16 nM of 'stop' oligonucleotides were very large and exceeded 12Kb in size confirming the inefficient truncation of the PCR products in the assembly reaction, due to the low concentration of 'stop' oligonucleotides (Fig 3.7 B). In contrast, the amplified PCR products containing 260 nM of 'stop' oligonucleotides were substantially smaller confirming the presence of a large proportion of truncated PCR products and therefore efficient primer binding and amplification (Fig 3.7 B). Consequently, 210 nM of 'stop' oligonucleotides was selected as the optimal concentration to be used in the assembly reaction (x10 cycles) which should generate PCR products of 300 bp - 2000 bp in size.

The presence of residual 'stop' oligonucleotides following the amplification reaction can also impede the efficiencies of restriction digests and ligation reactions due to the ability of the 'stop' oligonucleotides annealing together and potentially being cloned into the vector. Gel extraction of the PCR smear significantly decreased the yield of DNA which also decreased the efficiency of the ligation reaction and subsequently resulted in extremely low numbers of *E.coli* transformants. As an alternative, the amplified PCR products were purified using the high cut-off purification binding buffer B3 (Purelink PCR Purification Kit) which eliminated amplification primers and PCR products less than 300 bp and prevented the excessive loss of DNA. Two successive high cut-off purifications were performed and the PCR products were digested with NheI and XhoI and cloned into the equivalent site within the pCpGmCMV vector.

3.3.3. Construction of the pCpGmCMV cloning vector

The schematic diagram below (Fig 3.8) depicts the cloning procedure used to construct the pCpGmCMV cloning vector for the assembly PCR products.

A. pGL3mCMV was digested with NheI and AfeI to release the 2143 bp fragment containing the mCMV promoter and the luciferase gene (Fig 3.8 A).
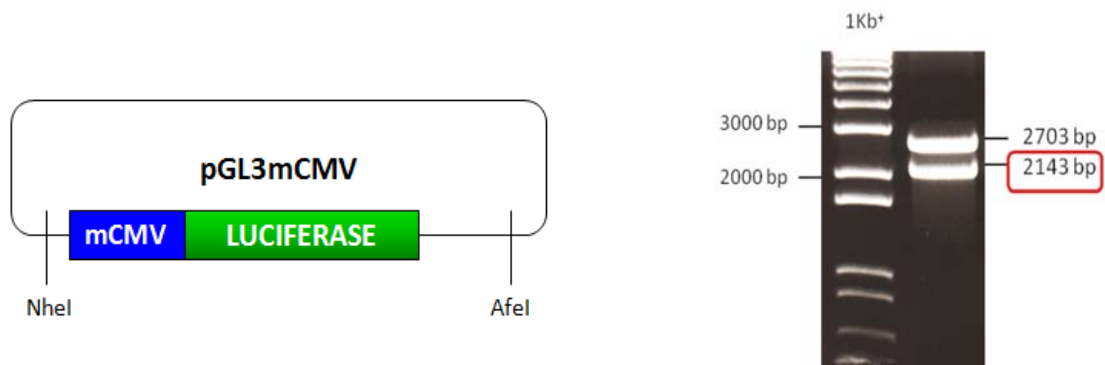
B. The AfeI enzyme produces blunt ended fragments, therefore pCpG-mSEAP was digested with SbfI and incubated with Klenow to produce a blunt ended fragment. The DNA was then digested with NheI and XbaI to release the 2194 bp fragment which contained the IFNβ-S/MAR and βGlo-MAR components, to generate the cloning vector pCpG (Fig 3.8 B).
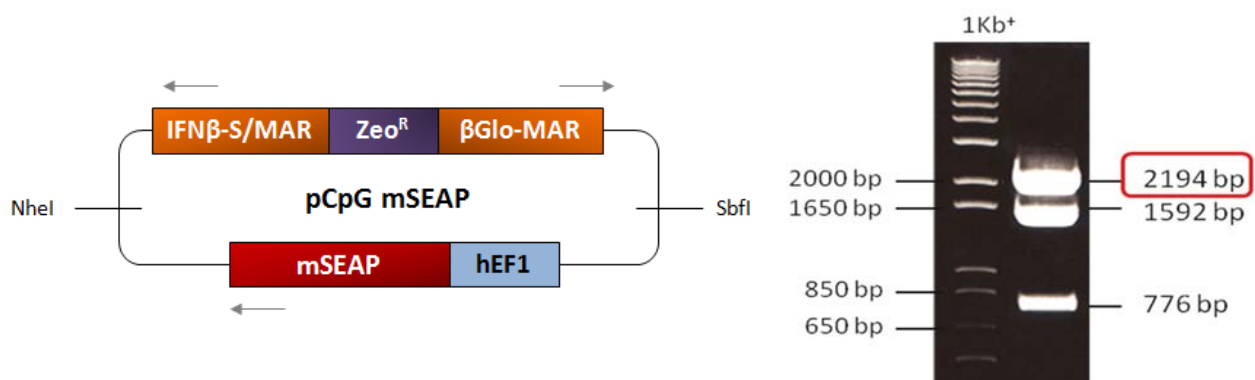
C. The mCMV and luciferase gene were cloned into the pCpG vector to generate pCpGmCMV-Luc⁺ (DNA sequence; Appendix 4), which was then digested with NheI and XhoI to serve as the cloning vector for the assembly PCR products (Fig 3.8 C).
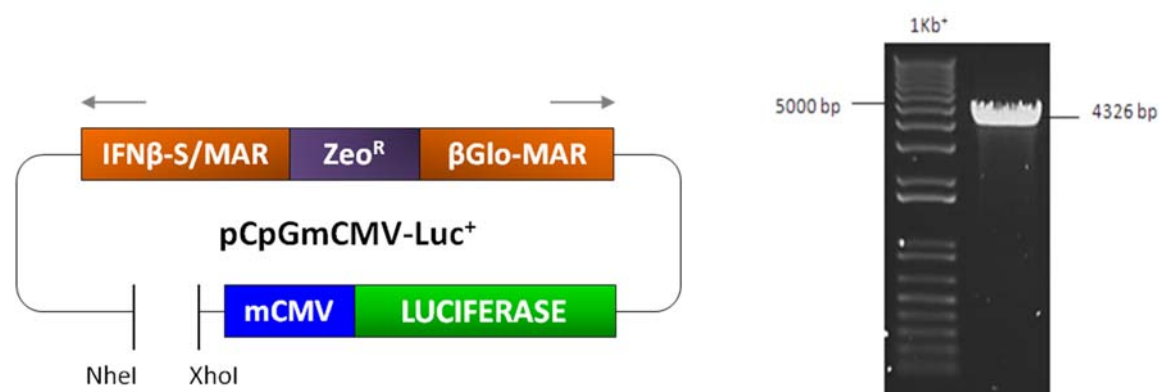
Figure 3.8. Schematic diagram illustrating the construction of the pCpGmCMV cloning vector. Sequential restriction digest reactions were visualised on an ethidium-bromide containing 0.5 x TAE agarose gel to confirm the correct size of the digested fragment prior to isolation and subsequent cloning.

The pCpGmCMV expression vector was generated by the ligation of two fragments from two different expression vectors; pCpG-mSEAP (InvivoGen, San Diego, CA, USA) and pGL3mCMV (section 3.2.1). The pCpG-mSEAP backbone plasmid contains two nuclear matrix attachment regions from the 5' region of the human IFN-β gene and the β-globin gene, which serve as insulators to form transcriptional barriers between independently regulated domains. The synthetic promoters and luciferase gene were cloned between these insulator elements to insulate the promoter from external influences on gene expression.

The PCR products with NheI and XhoI overhangs were cloned into the NheI/XhoI site within the pCpGmCMV cloning vector to generate a library of pCpGmCMV-TFBS plasmids, each harbouring multimerised TFBSs within their synthetic promoters. The constructed plasmids were verified by digestion with NheI and XhoI, which releases the PCR product containing the TFBSs and can be visualised by agarose-gel electrophoresis, as shown in Figure 3.9.
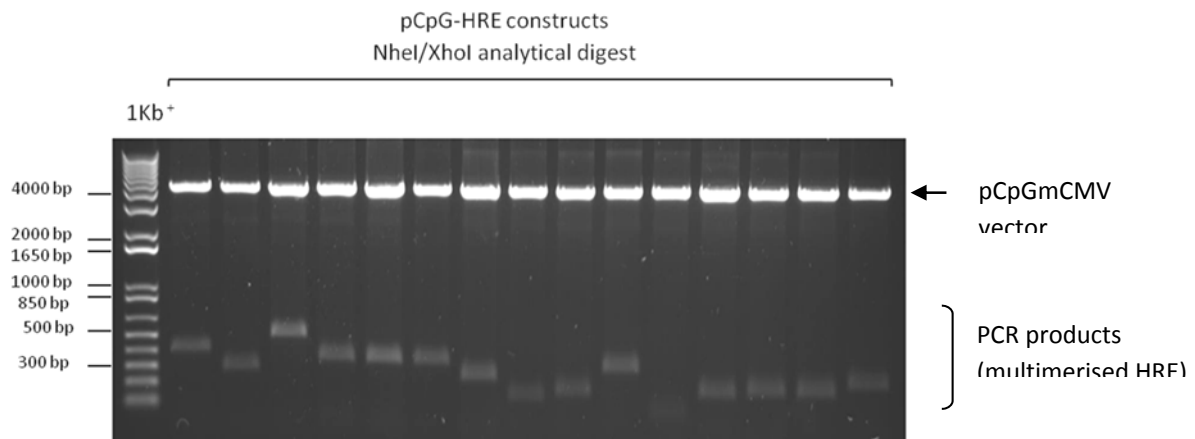
Figure 3.9. Analytical restriction enzyme digest of pCpG-HRE constructs. Each recombinant plasmid construct was digested with NheI and XhoI to release the PCR product. The digested DNA was visualised on an ethidium-bromide containing 0.5 x TAE agarose gel to confirm the successful incorporation of the PCR products. Released fragments larger than 100 bp confirmed the presence of multimerised HRE motifs within the pCpG-HRE constructs.

Interestingly, the released fragments in Figure 3.9 were much smaller than their amplified PCR products from the x25 cycle Assembly PCR reaction (Fig 3.7 B). This was probably due to the annealing of synthesised PCR products via the 5'-NheI and 3'-XhoI sequences which can bind to their complementary sequences to form a concatemer of several PCR products. Following digestion of the concatemeric PCR products with Nhe and XhoI, the individual PCR products were released and cloned into the pCpGmCMV vector. Therefore, analytical digests of the recombinant plasmid DNA constructs using NheI and XhoI should resemble the results in Figure 3.9.

3.3.4. Inflammation-inducible synthetic promoters exhibit varied transcription activities due to diversity of the consensus sequence

Divergence in gene expression between different cells and species is mediated by the ability of TFs to recognise numerous variations of their general consensus sequence and bind with different affinities to induce varied gene expression levels (Wittkopp, 2010). For example, various NFκB dimers can recognise and bind to the different variations of the general NFκB consensus sequence, GGGRNNYYCC (where R = A/G, Y= C/T and N= any nucleotide). The PCR component of the Assembly PCR method was exploited to incorporate non-specific nucleotides within the GGGRNNYYCC sequence to construct a variable NFκB-synthetic promoter library in order to investigate the effect of sequence diversity of the NFκB binding site on the functional activity of the synthetic promoters. In addition, synthetic promoters with fixed NFκB sequences were also generated to determine the effect of NFκB copy number on gene expression.

The fixed NFκB sequence, GGGACTTTCC, is preferentially bound by the p65/p50 NFκB dimer (Urban and Baeuerle, 1991) which is a highly expressed TF in synovial tissue and cells from RA patients (Handel *et al.,* 1995). Six copies of this sequence are present within the highly inflammation-responsive GNL6 construct (created by Khoury *et al.,* 2007), which served as a positive control for TNFα or PMA stimulation in the experiments within this thesis. The PCR oligonucleotides were comprised of the fixed or variable NFκB sequences (forward and reverse NFκB) flanked by 30 bp annealing sequences to introduce a 30 bp space between the fixed or variable NFκB motifs, respectively. The PCR products were digested with NheI and XhoI and cloned into the equivalent site within the pCpGmCMV vector. The plasmid DNA constructs were co-transfected with pRL-CMV into 293T cells which were unstimulated or stimulated with TNFα. Functional comparisons were made between the activities of the synthetic promoters with fixed and variable NFκB motifs to determine the effect of NFκB sequence diversity and fixed NFκB copy number on gene expression (Fig 3.10).
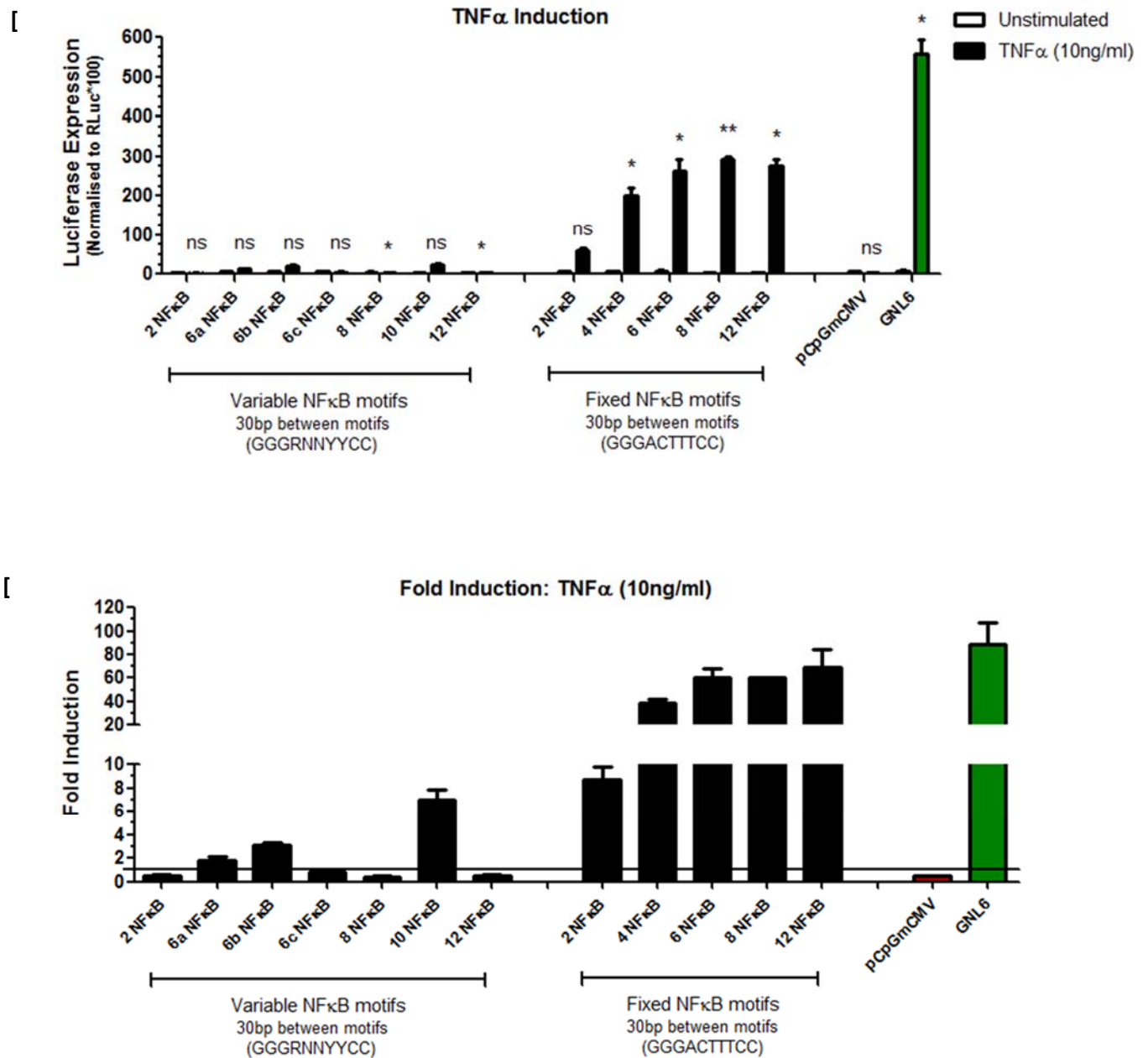
Figure 3.10. Comparative analysis of the functional activities of synthetic promoters comprised of variable or fixed NFκB binding sites. Oligonucleotides with the variable (GGGRNNYYCC) or fixed (GGGACTTTCC) NFκB motifs, flanked by 30 bp annealing sequences, were assembled and amplified using the Assembly PCR method. Following cloning, the plasmid DNA constructs (180 ng) were co-transfected with pRL-CMV (20 ng) in triplicate into 293T cells (20,000 cells in a 96-well plate) and were either unstimulated or stimulated with TNFα (10 ng/ml) for 18 hours [A] and the fold inductions were calculated [B]. The data represents the mean ± SD of triplicate values normalised to renilla luciferase. The statistical significance between unstimulated and TNFα-stimulated luciferase gene expression was calculated using the Student's t-test (ns= p>0.05, * = p≤0.05, ** = p≤0.01). The number of NFκB motifs in the promoters is stated e.g. 2NFκB.

Evidently, the synthetic promoters with fixed NFκB sequences were highly inflammation-inducible and displayed upto 70 fold induction in response to TNFα stimulation (pCpG-fixed-12NFκB) (Fig 3.10 B). In contrast, the synthetic promoters with variable NFκB motifs displayed very low and variable gene expressions, with a maximal fold induction of upto 7.2 fold (pCpG-variable-10NFκB). All of the synthetic promoters displayed similar basal gene expressions and therefore, the vast differences in fold inductions can be attributed to the robust TNFα-inducibility displayed by the fixed NFκB-promoters, in contrast to the marginal TNFα-induced gene expression by the variable NFκB-promoters.

Due to the use of the fixed NFκB sequences, it was possible to determine the effect of NFκB copy number on gene expression. As anticipated, the luciferase gene expression increased with increasing numbers of NFκB motifs, following TNFα stimulation. For example, pCpG-fixed-2NFκB displayed 9 fold induction, although the change in gene expression following TNFα stimulation was not statistically significant (ns, p=0.07). In contrast, pCpG-fixed-12NFκB displayed 70 fold induction to TNFα stimulation (* p=0.03), which was due to greater copy numbers of the fixed NFκB motifs within the latter promoter. Interestingly, the synthetic promoters with 6, 8 and 12 fixed NFκB motifs displayed 63, 62 and 70 fold inductions respectively, which suggests that the synthetic promoters reached a saturation point where increasing the TFBS copy number had minimal effects on the level of gene expression. The observation that very similar levels of gene expression can be induced by a promoter with fixed 6NFκB motifs (* p=0.05) or fixed 12NFκB motifs (* p=0.03) indicates that high-inducibility can be achieved in compact synthetic promoters, with minimal copy numbers of TFBSs (Fig 3.10 A)

In contrast, the diversity in gene expression induced by the variable NFκB-promoters was not solely attributed to the number of NFκB motifs within the promoters. For example, promoters 6a, 6b and 6c all possessed six variable NFκB motifs but display poor and varied fold inductions to TNFα stimulation, suggesting that the different NFκB sequences was accountable for the variation in their promoter activities, despite the changes in gene

expression being statistically insignificant (ns, p>0.05). Interestingly, the synthetic promoter within the pCpG-variable-10NFκB construct displayed 7.2 fold induction, although the change in gene expression following TNFα stimulation was not statistically significant (ns, p>0.05). The sequence analysis of this promoter revealed that some of the NFκB motifs had a very similar sequence to the fixed GGGACTTTCC NFκB sequence (DNA sequence: Appendix 5), which may have been accountable for the higher TNFα-induced gene expression by this promoter. These results also, to an extent, highlight the preferential binding of the TNFα-inducible p65/p50 NFκB dimer to the fixed NFκB binding sites and similar sequences.

Overall, minor variations within the NFκB binding site can have dramatic effects on the activity of the synthetic promoters and as shown in Figure 3.10, these changes can result in unpredictable promoter activities, which is undesirable. Synthetic promoters within the fixed NFκB library displayed consistently high responsiveness to TNFα and therefore, the fixed sequences of the NFκB (GGGACTTTCC), AP-1 (TGAGTCA) and HIF-1α (ACGTGC) binding sites will be used to generate the composite synthetic promoters.

3.3.5. Increased spacing between the TFBSs results in decreased basal and induced luciferase gene expression

Composite synthetic promoters were previously generated using the random cloning method (section 3.2). These synthetic promoters contained the core binding sites of AP-1, C/EBPβ, Egr-1, Ets-1, HIF-1α and NFκB transcription factors, with only a 4bp space between each TFBS and it is possible that the close proximity of the TFBSs introduced steric hindrance of the TFs which may have impaired the induction of synergistic gene expression following combined inflammatory and hypoxic stimulation. In order to determine whether increasing the spacing between TFBSs could alleviate the potential steric hindrance of the TFs to allow synergistic gene expression, it was important to first investigate the general effects of increased spacing between the TFBSs on gene expression.

In the Assembly PCR method, the length of the annealing sequence determines the degree of spacing between the TFBSs and this spacing can be extended by replacing the reverse TFBS-oligonucleotide with a reverse spacer-oligonucleotide. Figure 3.11 depicts the assembly of oligonucleotides with different lengths of the annealing and spacer sequences which were used to generate synthetic promoters with specified degrees of spacing between the TFBSs.

AP-1   15bp   AP-1

```
5'- TGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATC     TGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATC-3'
              3'-CACGGAGAATACTAGACTCAGTACGCTACTTGGAGTG-5'
```

AP-1 Reverse

AP-1   20bp   AP-1

```
5'- ATCTCTGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATCTGGAT     ATCTCTGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATCTGGAT-3'
              3'-CACGGAGAATACTAGACCTAACTCAGTTAGAGACGCTACTTGGAGTG-5'
```

AP-1 Reverse

AP-1   30bp   AP-1

```
5'- ATCTCTGCGATGAACCTCACCATGTTGAGTCAACAAGGTGCCTCTTATGATCTGGAT        ATCTCTGCGATGAACCTCACCATGTTGAGTCAACAAGGTGCCTCTTATGATCTGGAT-3'
              3'-CACGGAGAATACTAGACCTAGTACAACTCAGTTGTTCTAGAGACGCTACTTGGAGTG-5'
```

AP-1 Reverse

AP-1   35bp   AP-1

```
5'- TGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATC     TGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATC-3'
              3'-CACGGAGAATACTAGGGTACACGCTACTTGGAGTG-5'
```

5nt Spacer

AP-1   40bp   AP-1

```
5'- TGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATC       TGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATC-3'
              3'-CACGGAGAATACTAGGGTACAGACAACGCTACTTGGAGTG-5'
```

10nt Spacer

AP-1   45bp   AP-1

```
5'- ATCTCTGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATCTGGAT       ATCTCTGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATCTGGAT-3'
              3'-CACGGAGAATACTAGACCTAGGTACTAGAGACGCTACTTGGAGTG-5'
```

5nt Spacer

AP-1   50bp   AP-1

```
5'- ATCTCTGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATCTGGAT        ATCTCTGCGATGAACCTCACTGAGTCAGTGCCTCTTATGATCTGGAT-3'
              3'-CACGGAGAATACTAGACCTAGGTACAGACATAGAGACGCTACTTGGAGTG-5'
```

10nt Spacer

AP-1   55bp   AP-1

```
5'- ATCTCTGCGATGAACCTCACCATGTTGAGTCAACAAGGTGCCTCTTATGATCTGGAT        ATCTCTGCGATGAACCTCACCATGTTGAGTCAACAAGGTGCCTCTTATGATCTGGAT-3'
              3'-CACGGAGAATACTAGACCTAGGTACTAGAGACGCTACTTGGAGTG-5'
```

5nt Spacer

AP-1   60bp   AP-1

```
5'- ATCTCTGCGATGAACCTCACCATGTTGAGTCAACAAGGTGCCTCTTATGATCTGGAT          ATCTCTGCGATGAACCTCACCATGTTGAGTCAACAAGGTGCCTCTTATGATCTGGAT-3'
              3'-CACGGAGAATACTAGACCTAGGTACAGACATAGAGACGCTACTTGGAGTG-5'
```

10nt Spacer

Figure 3.11. Assembly of oligonucleotides to generate PCR products with specified spacing between the TFBSs. Oligonucleotides comprising the AP-1 (and fixed NFκB) binding sites flanked by annealing sequences of specified lengths were assembled using the Assembly PCR method.

The oligonucleotides with different annealing sequence lengths and spacer nucleotides were assembled in specific combinations to generate synthetic promoter libraries with 15 bp space to 60 bp space between the AP-1 and NFκB motifs (n=6 per group).

The assembly of oligonucleotides with forward and reverse TFBSs flanked by 15 bp, 20 bp and 30 bp annealing sequences generated promoters with 15 bp, 20 bp and 30 bp space between the TFBSs, respectively. The annealing of a forward TFBS oligonucleotide with 15 bp annealing sequence to either a reverse 5 nt spacer-oligonucleotide or a reverse 10 nt spacer-oligonucleotide generated promoters with 35 bp and 40 bp space between the TFBSs, respectively. A spacing of 45 bp and 50 bp between the TFBSs was achieved by annealing a forward TFBS oligonucleotide with 20 bp annealing sequence to either a reverse 5 nt spacer-oligonucleotide or a reverse 10 nt spacer-oligonucleotide, respectively. Finally, the promoters with TFBSs separated by 55 bp and 60 bp space were generated by annealing of a forward TFBS oligonucleotide with 25 bp annealing sequence to either a reverse 5 nt spacer-oligonucleotide or a reverse 10 nt spacer-oligonucleotide, respectively.

The synthetic promoters with the specified spacing between the AP-1 binding sites (Fig 3.12 A) or the NFκB binding sites (Fig 3.12 B) were constructed to investigate the effect of spacing between the TFBSs on the transcriptional activity of the synthetic promoters. The PCR products were cloned into the pCpGmCMV vector and the recombinant plasmid constructs were co-transfected with pRL-CMV into 293T cells. Transfected cells were either unstimulated or stimulated with PMA (10ng/ml) or TNFα (10ng/ml) to activate AP-1 or NFκB, respectively.
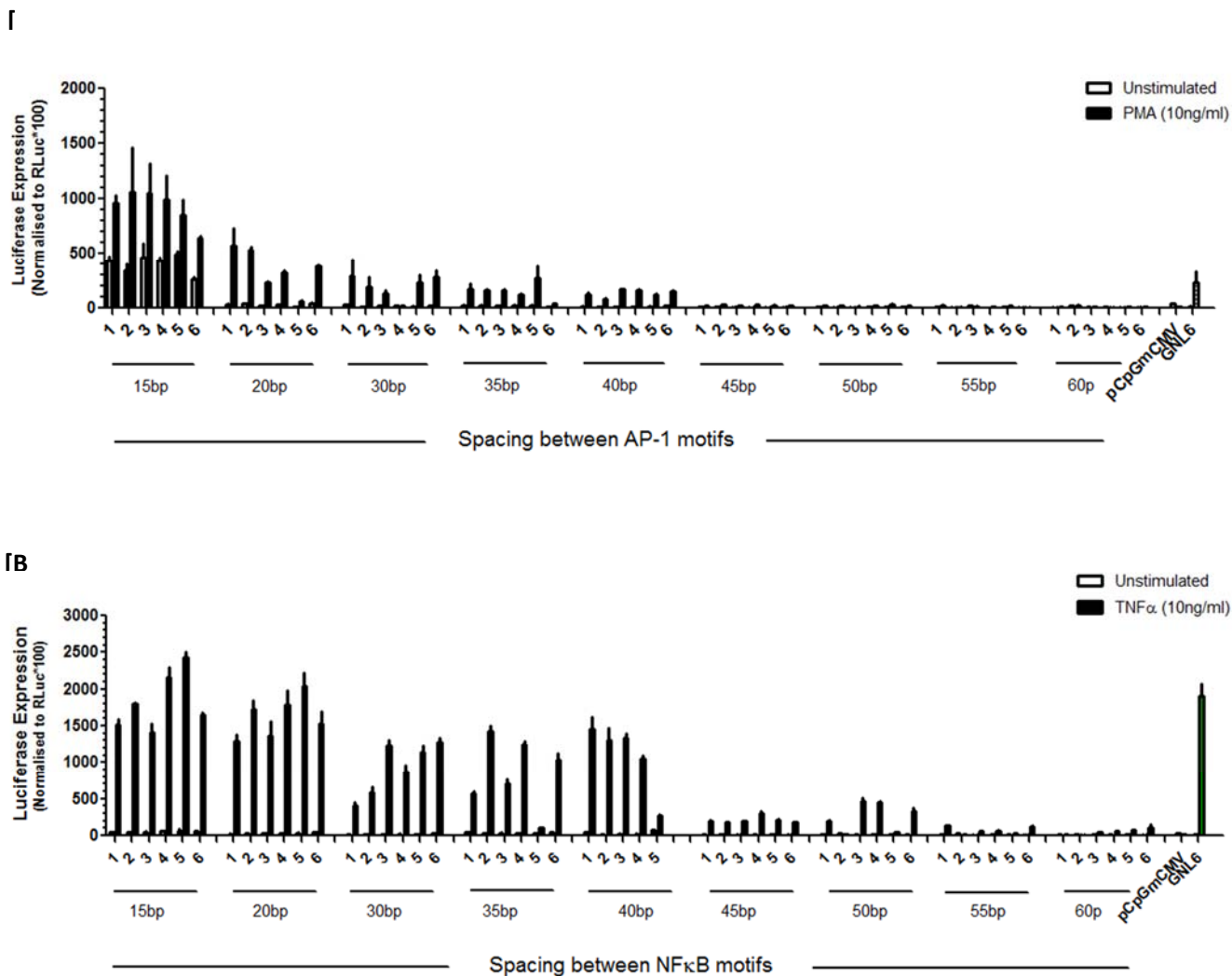
Figure 3.12. Increased spacing between the TFBSs negatively impacts the transcriptional activity of the synthetic promoters. Increased spacing was introduced between the AP-1 [A] and NFκB motifs [B] within the AP-1 and NFκB-responsive synthetic promoters (n=6, per group). The promoter constructs (180 ng) were co-transfected with pRL-CMV (20 ng) into 293T cells (20,000 cells in a 96-well plate). Transfected cells were unstimulated or stimulated with PMA (10 ng/ml) or TNFα (10 ng/ml) for 18 hours. The data represents the mean ± SD of triplicate values normalised to renilla luciferase.

Interestingly, increasing the spacing between the TFBSs had a profound effect on the transcriptional activity of the synthetic promoter. A clear trend can be observed; with increased spacing between the TFBSs, there was a gradual decrease in the basal (unstimulated) and induced (PMA and TNFα stimulated) luciferase gene expression. This trend was more

noticeable when the basal and induced luciferase gene expression data were plotted separately in Figure 3.13.

A

AP-1: Basal Expression

B

NFκB: Basal Expression

C

AP-1: PMA Induction
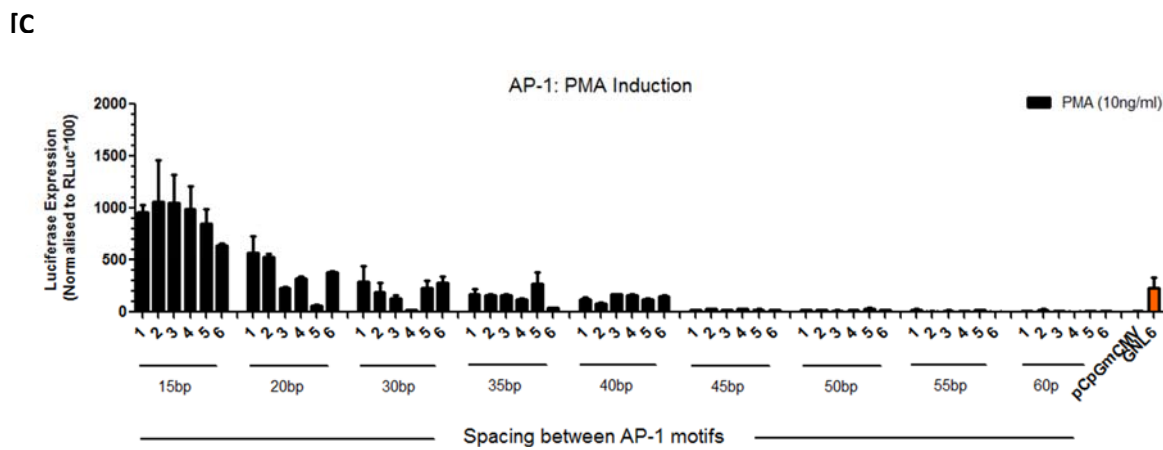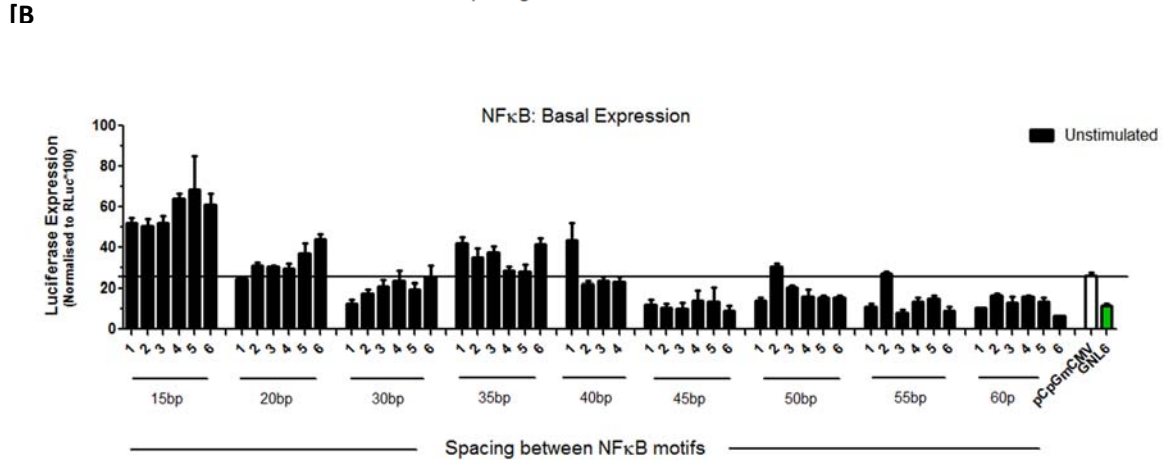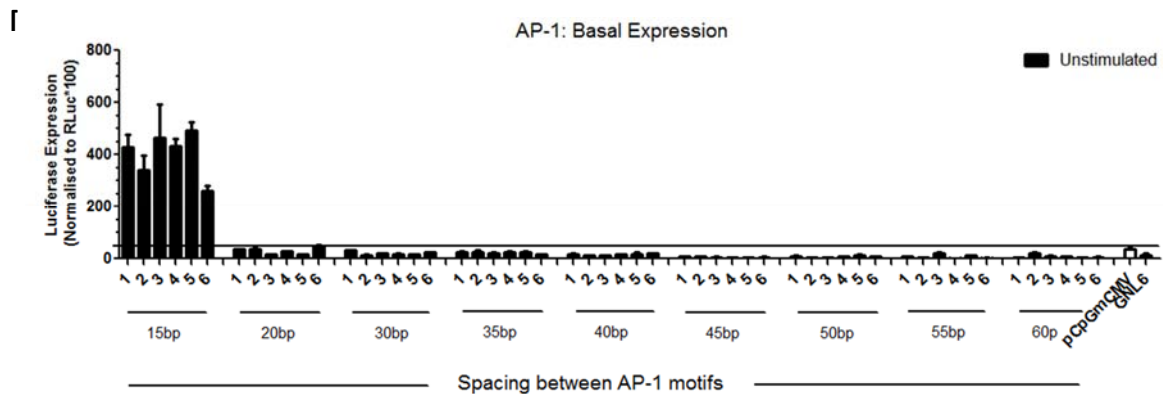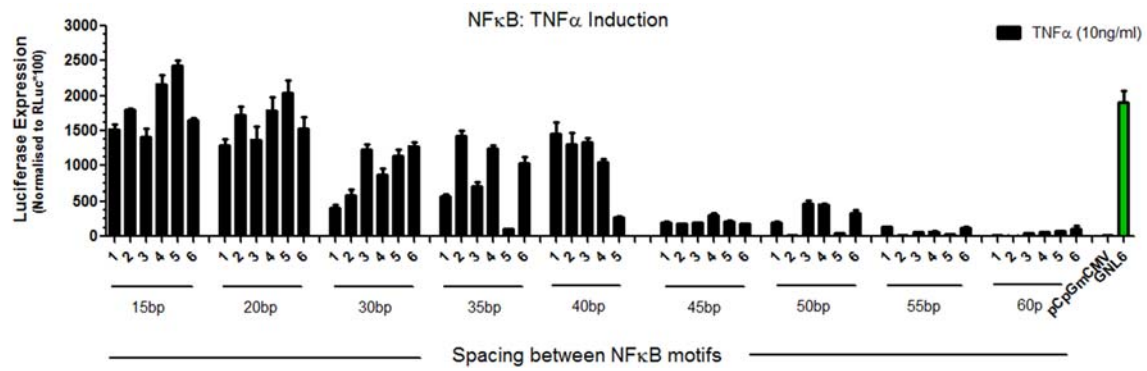
Figure 3.13. The basal and induced luciferase gene expression decreases with increased spacing between the TFBSs. The basal and induced gene expression data from AP-1-responsive synthetic promoters [A] and [C], respectively and from NFκB-responsive synthetic promoters

As demonstrated in the preceding section 3.3.4, the number of TFBSs within the synthetic promoters can impact the transcriptional activity of the promoter. Therefore, to eliminate the influence of TFBS copy number on the gene expression, the basal and induced luciferase gene expression from synthetic promoters with 8NFκB motifs (selected from 5 spacer groups) were re-plotted from Figure 3.12 (DNA sequencing in Appendix 6). In this way, all variables were kept constant except for the degree of spacing between the TFBSs, in order to determine the effect of spacing between the TFBSs on promoter activity (Fig 3.14).
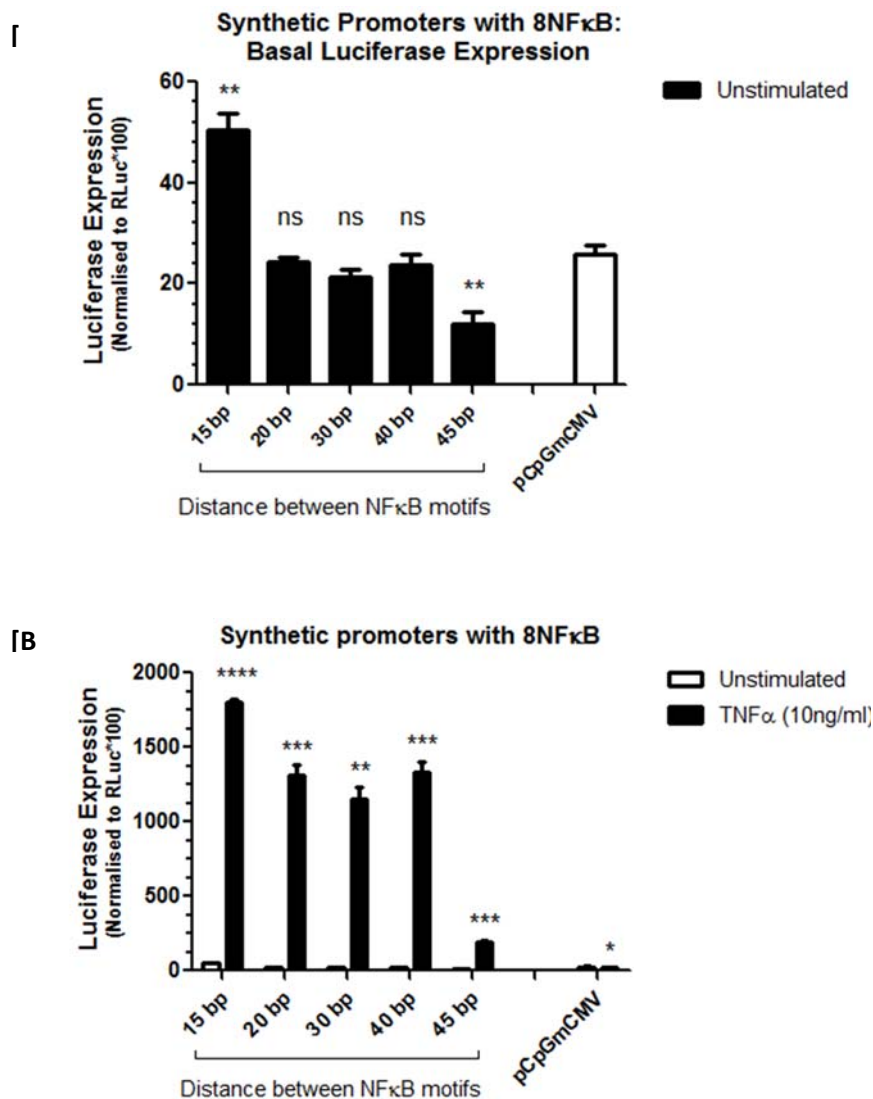
Figure 3.14. Synthetic promoters with clustered TFBSs have higher basal and induced gene expression than synthetic promoters with sparse TFBSs. Synthetic promoters with 8NFκB motifs were selected from 5 spacer groups. The basal alone [A] and basal and TNFα-induced [B] luciferase gene expressions were re-plotted from the data in Figure 3.12. The data represents the mean ± SD of triplicate values normalised to renilla luciferase. The statistical significance between the basal

Figure 3.14 demonstrated that the synthetic promoter with 15bp space between NFκB motifs had a significantly higher basal expression than pCpGmCMV (** p=0.002) and the highest TNFα-induced gene expression (**** p=0.0001). In contrast, the synthetic promoter with 45bp space between NFκB motifs had a significantly lower basal expression than pCpGmCMV (** p=0.002) and the lowest TNFα-induced gene expression (*** p=0.001). The results in Figure 3.14 comply with the trend observed in Figure 3.12, therefore, the effect of TFBS number on luciferase gene expression was negated (in this experiment), and the general activities of the synthetic promoters with diverse TFBS spacing were comparatively analysed to determine the effect of TFBSs spacing on gene expression induced by the synthetic promoters.

Interestingly, the synthetic promoters with a 15 bp space between the AP-1 and NFκB motifs had higher basal luciferase expressions than the negative control backbone plasmid pCpGmCMV, which is undesirable. Generally, the synthetic promoters with a 15bp space between TFBSs had higher basal luciferase gene expression than the synthetic promoters with sparse TFBSs (Fig 3.13 A and B). Figures 3.13 C and 3.13 D showed that with increased spacing between the TFBSs, there was also a decrease in the PMA and TNFα-induced luciferase gene expression by the AP-1 and NFκB-responsive synthetic promoters, respectively. Interestingly, the synthetic promoters with 45-60 bp spacing between the TFBSs had dramatic lower induced luciferase gene expression levels than the other promoters.

Overall, the observed trend in Figure 3.12 provided a filtering parameter to select the spacing between the TFBSs which permitted the sought-after characteristics of inflammation-inducible synthetic promoters. Promoters with a 20 bp space between the TFBSs displayed low basal gene expression (comparable to the pCpGmCMV control) and high induced luciferase gene expression and therefore, composite synthetic promoters will be constructed with a 20 bp space between the TFBSs, which may alleviate the proposed steric hindrance associated with the close proximity of TFBSs.

3.3.6. Construction of cloning vectors used to investigate the effect of spacing between the proximal TFBS and the TATA box

During transcription, activated TFs bind to their respective binding sites and with the aid of general transcription factors, RNA polymerase II is positioned on the TATA box to form a preinitiation complex (PIC), for transcription to proceed. The experiment in section 3.3.2 highlighted that the pGL3-4bp-composite promoters with proximal HRE motifs had greater hypoxia-inducibility than promoters with HRE motifs flanked by other TFBSs, which may have been due to the facilitated accessibility of the proximal HRE to interact with the closely located RNA polymerase II and PIC on the TATA box.

To investigate the effect of spacing between the proximal TFBS and the TATA box on gene expression, the Assembly PCR method was implemented to generate synthetic promoters with increased spacing between these two components. This was achieved by incorporating specified degrees of spacing between the XhoI and SalI sites within the various 3'-XhoI 'stop' primers, listed in Table 3.1.

3'-XhoI 'stop' primers with 20bp annealing sequence	Distance between proximal TFBS and TATA box within the promoter
XhoI-primer (current)	55 bp (current)
0bp-SalI primer	60 bp
5bp-SalI primer	66 bp
9bp-SalI primer	70bp
14bp-SalI primer	74 bp

Table 3.1. List of 3'-XhoI 'stop' PCR primers, with 20bp annealing sequence, used to create PCR products with specified spacing between the proximal TFBS and the TATA box. The specified spacings between the proximal TFBS and the TATA listed above will only be generated with 3'-XhoI 'stop' primers which have <u>20bp</u> annealing sequence. If the annealing sequence length is changed, the distance between the TFBS and the TATA box will also change.

Figure 3.15 schematically depicts the construction of synthetic promoters with various spacing between the proximal TFBS and the TATA box, using the Assembly PCR method.
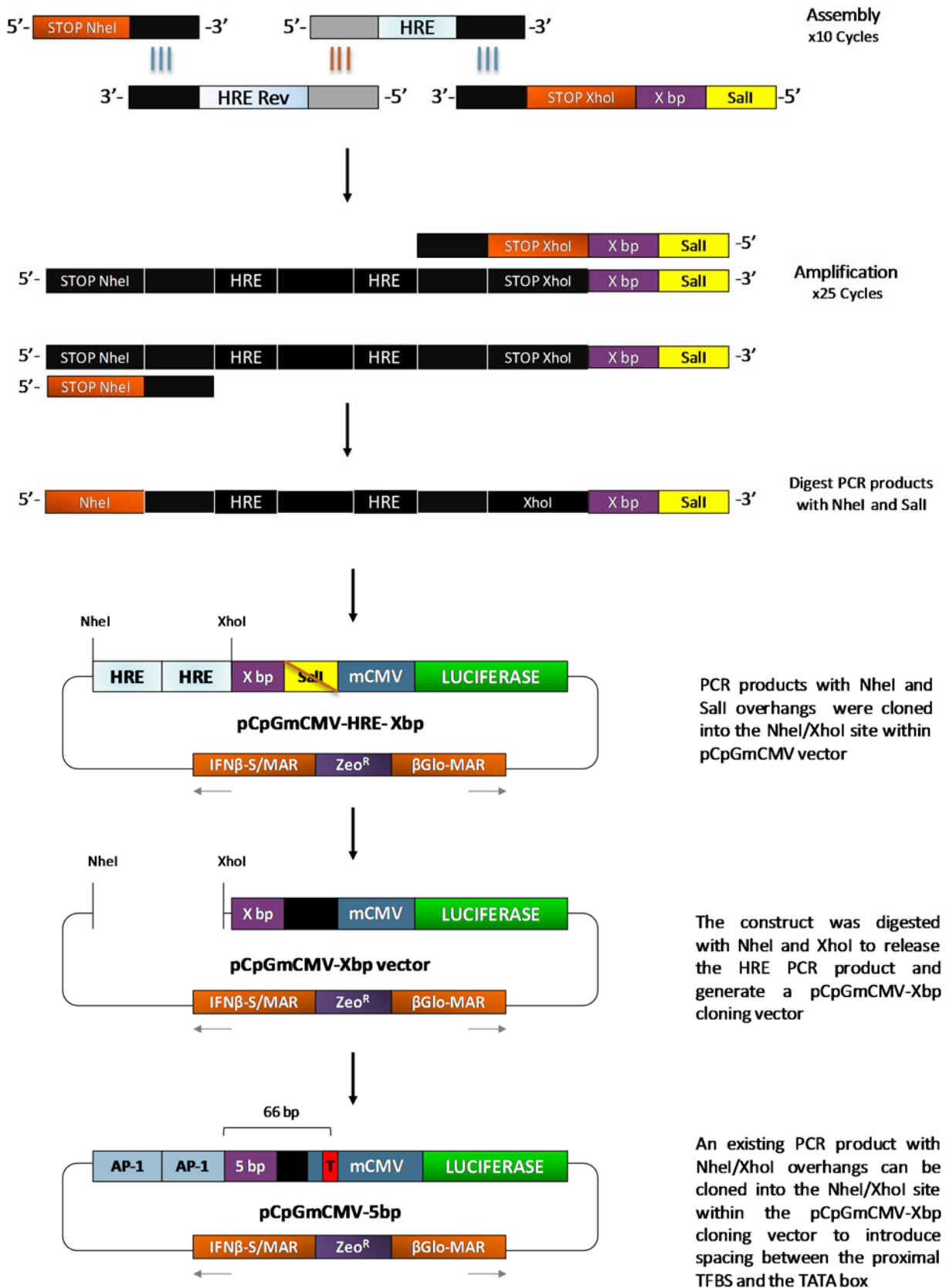
Figure 3.15. Cloning strategy used to construct synthetic promoters with varied spacing between the TFBS and the TATA box. Specified spacing was introduced between the proximal TFBS and the TATA box (T) by cloning existing PCR products into various

Using Assembly PCR, four individual PCR reactions were set up with the standard PCR components, forward and reverse oligonucleotides with HRE motifs flanked by 20 bp annealing sequences, 5'-NheI 'stop' primer and the four different 3'-XhoI 'stop' primers to generate PCR products with a 20 bp space between HRE motifs and specified spacings between the XhoI and SalI restriction enzyme sites. Following purification, the four different PCR products were digested with NheI and SalI and cloned into the NheI/XhoI digested pCpGmCMV vector. The ligation of compatible SalI and XhoI overhangs generated the hybrid GTCGAG sequence, which is unrecognisable by any restriction enzyme, thereby destroying the XhoI overhang within the vector. The resulting constructs (DNA sequencing; Appendix 7) were digested with NheI and XhoI to release the HRE PCR product and generate four different pCpG-Xbp cloning vectors, where X is 0 bp, 5 bp, 9 bp or 14 bp between the XhoI overhang and XhoI/SalI destroyed restriction enzyme site. The existing PCR products with a 20 bp space between the TFBSs were cloned into the four different cloning vectors to introduce specific spacings (listed in Table 3.1) between the proximal TFBS and the TATA box.

Successful assembly and amplification of the PCR products was confirmed by visualising on an ethidium bromide containing 0.5 x TAE agarose gel (Fig 3.16).
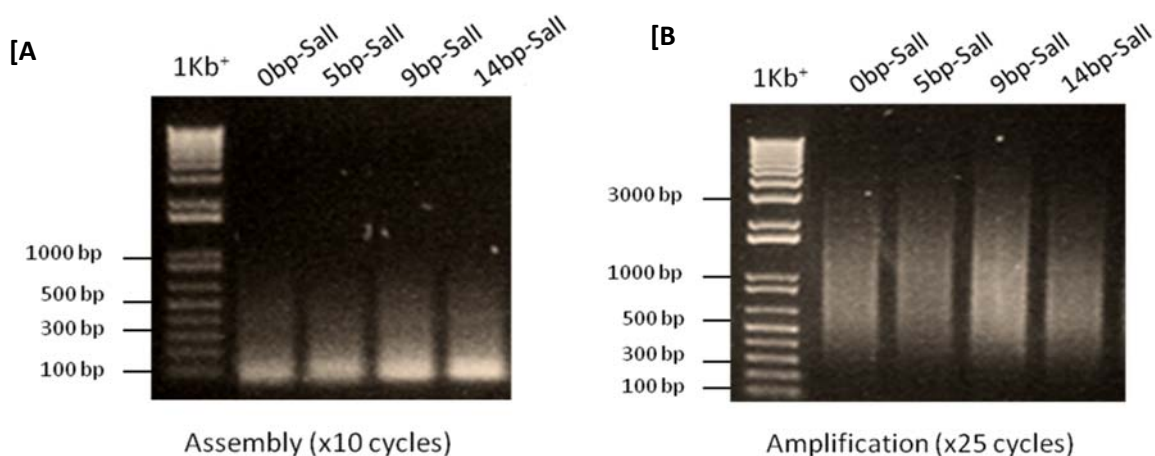
Figure 3.16. Assembly and amplification of PCR products used to generate cloning vectors with varying distances between the TFBS and the TATA box. Five microlitres of assembled [A] and amplified [B]

3.3.7. Proximal TFBSs exhibit positional preferences relative to the TATA box

Synthetic promoters with increased spacing between proximal AP-1 and NFκB motifs and the TATA box were generated by cloning existing PCR products with 8AP-1 or 6NFκB motifs into the various cloning vectors constructed in section 3.3.6 (DNA sequencing; Appendix 8). The resulting plasmid DNA constructs were co-transfected with pRL-CMV into 293T cells and were either unstimulated or stimulated with PMA or TNFα to activate AP-1 or NFκB, respectively (Fig 3.17).

[A
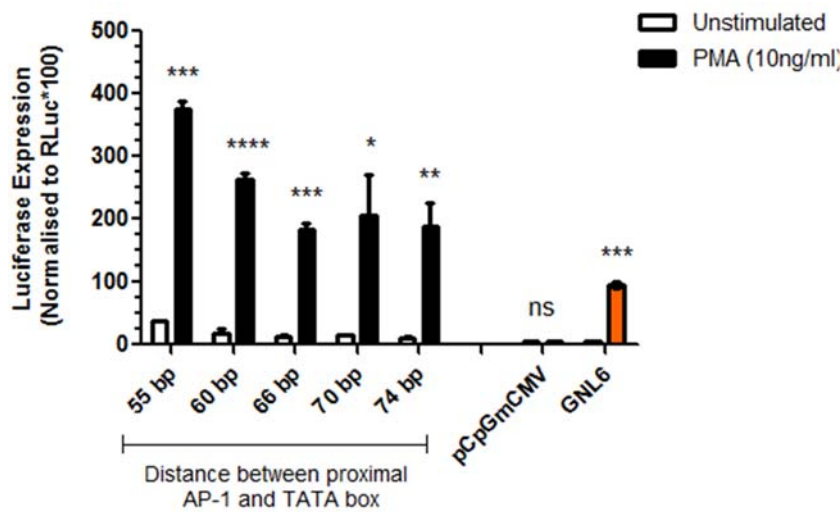
[B
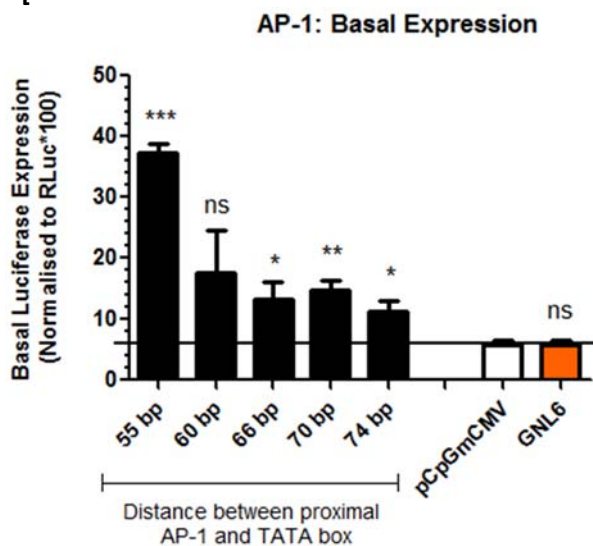
[C
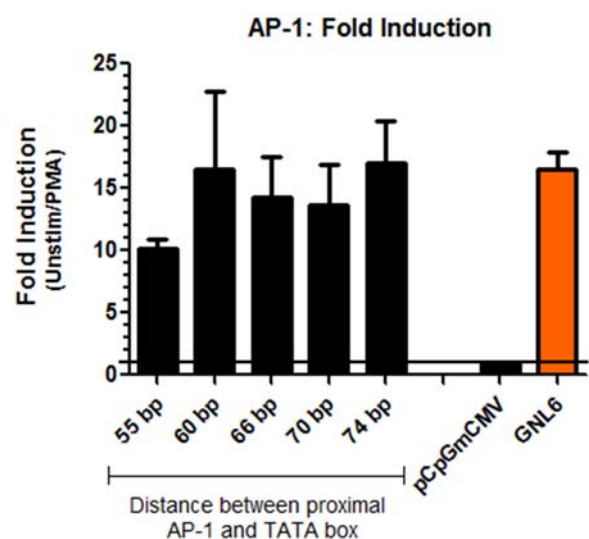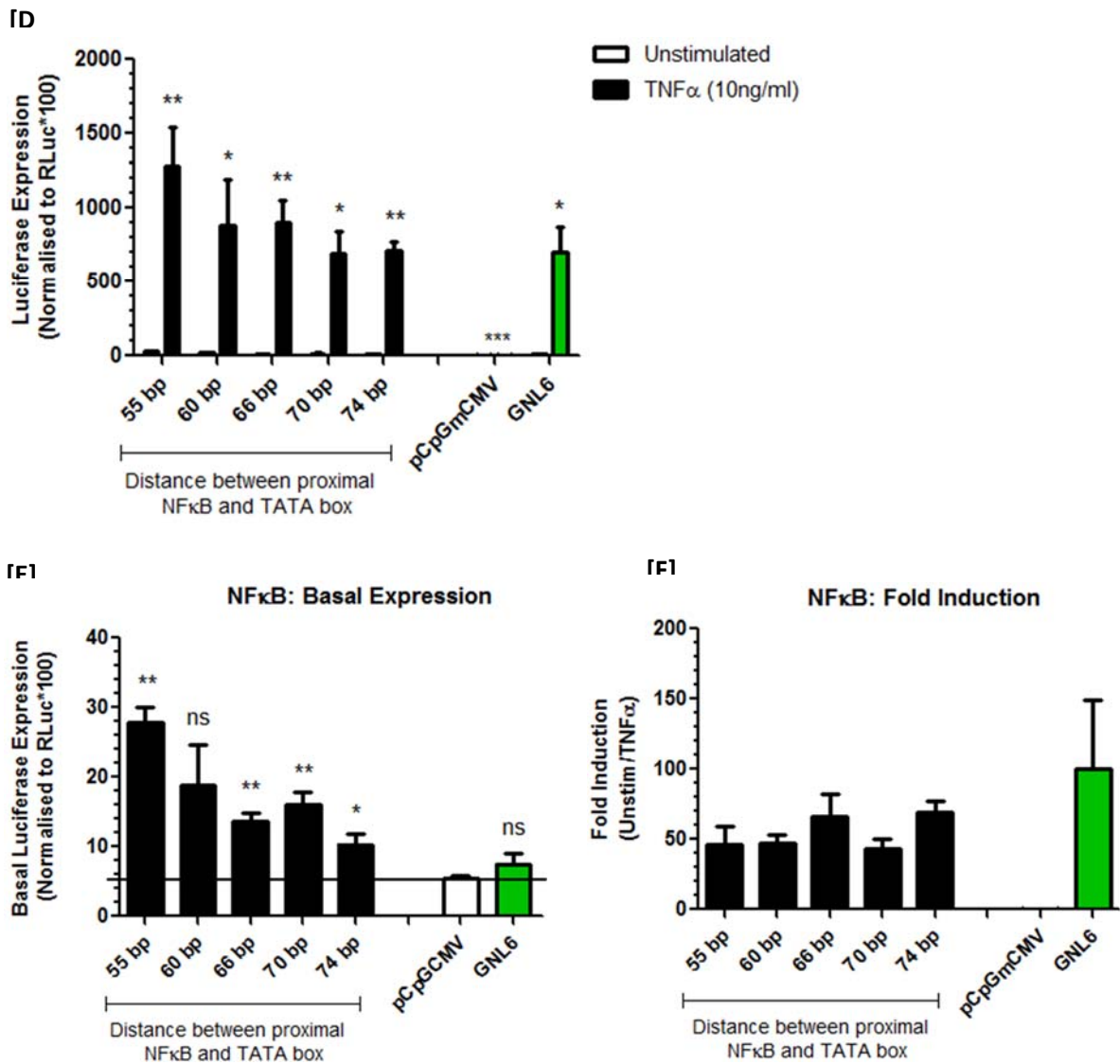
Figure 3.17. Increased spacing between the proximal TFBS and the TATA box negatively affects the basal and induced gene expression induced by the synthetic promoter. PCR products with 8AP-1 [A], [B] and [C] and 6NFκB motifs [D], [E] and [F] were cloned into different expression vectors which allowed 55 bp to 74 bp spacing between the proximal TFBS and the TATA box. The resulting constructs (180 ng) were co-transfected with pRL-CMV (20 ng) into 293T cells (20,000 cells in a 96-well plate). Transfected cells were unstimulated or stimulated with PMA (10 ng/ml) or TNFα (10 ng/ml) for 18 hours to activate AP-1 or NFκB, respectively. The data represents the mean ± SD of triplicate values normalised to renilla luciferase. The statistical significance was calculated using the Student's t-test (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p≤0.0001). The statistical differences between the unstimulated and stimulated luciferase gene expressions are shown in [A] and [D], whilst the statistical differences between the basal expressions of the constructs compared to that of pCpGmCMV are shown in [B] and [E]. Fold inductions are shown in [C] and [F].

Interestingly, the synthetic promoters with a proximal TFBS located closer to the TATA box had the highest basal and induced luciferase gene expression levels. As the proximal TFBSs was positioned further from the TATA box, the basal and induced gene expression gradually decreased. This trend was observed in both the AP-1-responsive promoters (Fig 3.17 A and B) and the NFκB-responsive promoters (Fig 3.17 D and E). Due to equivalent decreases in both the basal and induced luciferase gene expressions levels, the overall fold inductions were very similar for the AP-1-responsive promoters (Fig 3.17 C) and the NFκB-responsive promoters (Fig 3.17 F).

Overall, the promoters with a 66 bp space generally displayed low basal and highly induced luciferase gene expression and will be applied to construct the composite synthetic promoters.

3.4. pCpG-4bp-composite synthetic promoters exhibit impaired synergistic gene expression, potentially due to steric hindrance of TFs

For simplification purposes, a revised pCpG-4bp-composite promoter library was generated so that all of the composite promoters possess the same pCpG vector, the same degree of spacing between the proximal TFBS and the TATA box (66 bp space) and the same TFBS sequences (NFκB, AP-1 and HRE motifs), to eliminate the influences of dissimilar components. In this way, all factors remained constant, with the exception of the degree of spacing between the TFBS in order to confidently determine whether increasing the spacing between the TFBS can permit synergistic gene expression. The construction of the revised pCpG-4bp-composite promoters is schematically depicted in Figure 3.18.

A. A spacer oligonucleotide was ordered with 5'-XbaI and 3'-XhoI overhangs to allow a 66bp space between the proximal TFBS and the TATA box. The XbaI overhang 5'-CTAG, is compatible to the NheI overhang, and the ligation of XbaI and NheI overhangs generates the sequence 5'-GCTAGA, which is unrecognisable by any restriction enzyme and destroys the NheI site. Therefore, the annealed spacer oligonucleotide was cloned into the NheI/XhoI site, upstream of the mCMV promoter within the pCpGmCMV vector to generate the pCpGmCMV-66bp cloning vector (Fig 3.18 A).

B. The spacer oligonucleotide contains an internal NheI site,

206
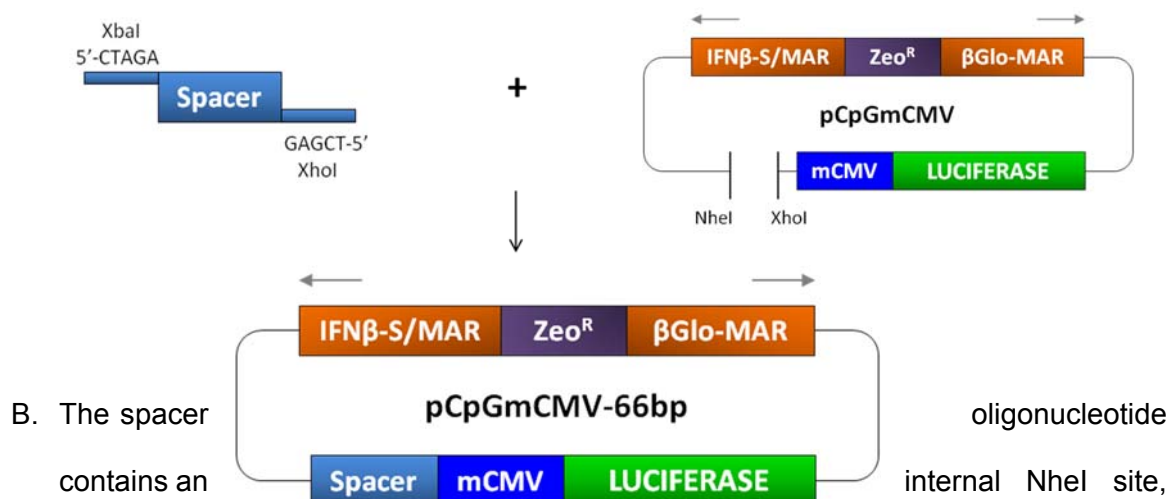
therefore the pCpGmCMV-66bp construct was digested with NheI to generate the linear pCpGmCMV-66bp cloning vector. The linear vector was incubated with CIP to dephosphorylate the ends of the DNA to prevent re-ligation of the vector (Fig 3.18 B).
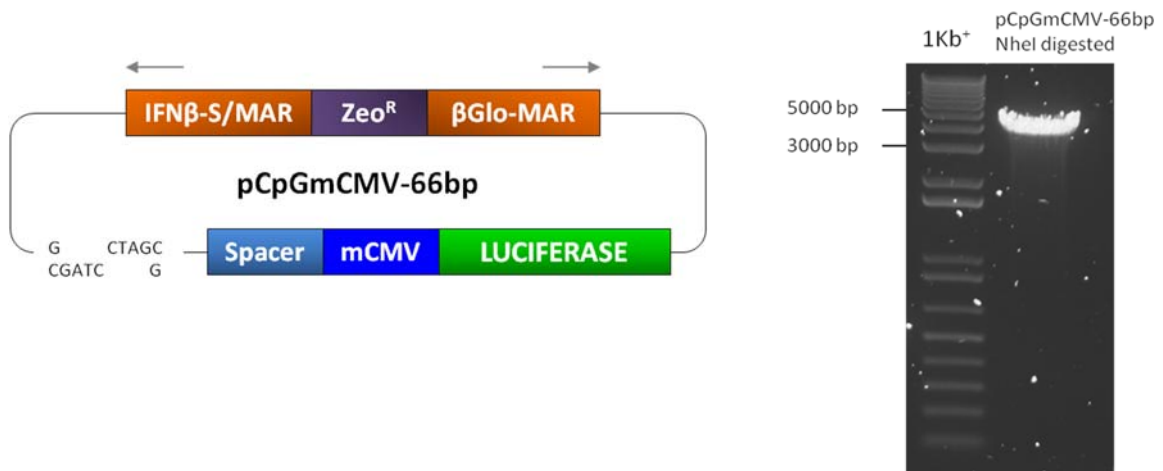
C. The oligonucleotides comprised of the fixed NFκB, AP-1 and HIF-1α binding site sequences were ordered with phosphorlyated 5'-CTAG overhangs, which are compatible to the NheI overhang. The annealed oligonucleotides were cloned at random into the pCpGmCMV-66bp cloning vector to generate a library of composite promoters with varying compositions and numbers of TFBSs, with a 4bp space between each TFBS and a 66bp space between the proximal TFBS and the TATA box (Fig 3.18 C). The DNA sequences of selected pCpG-4bp-composite promoters are presented in Appendix 9.
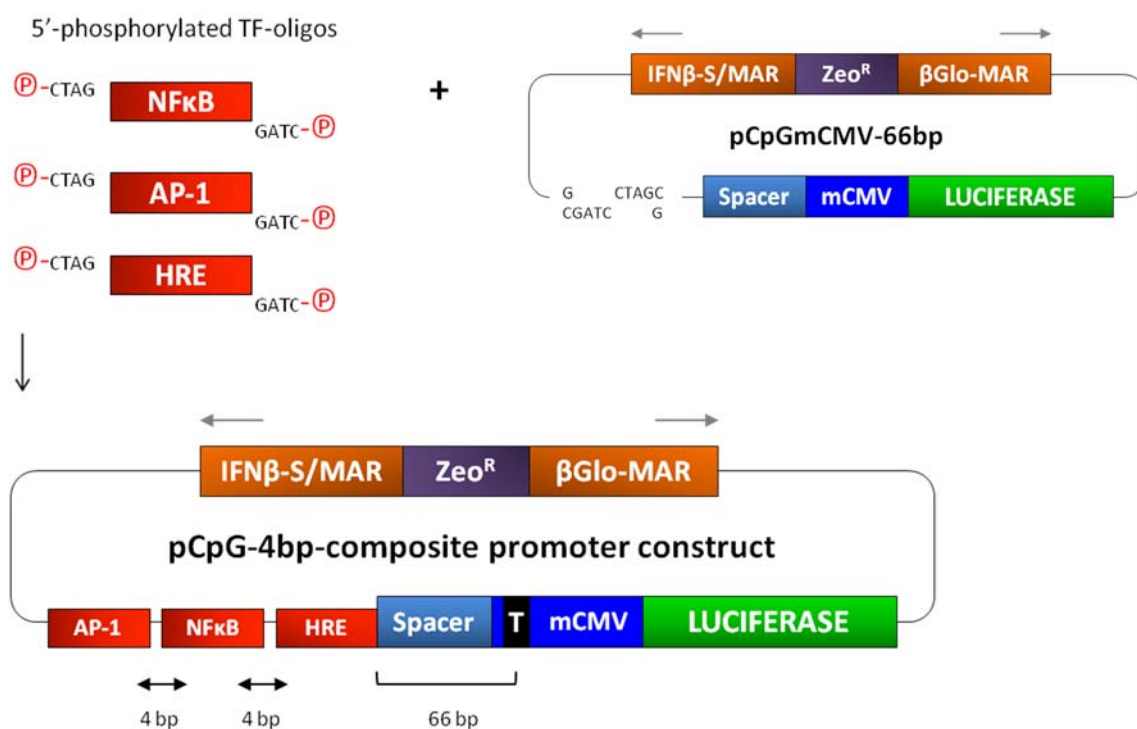
Figure 3.18. A schematic diagram of a pCpG-4bp-composite synthetic promoter. Using the random ligation method, each TFBS was separated by a 4bp space. The proximal TFBS and the TATA box (T) were separated by a 66bp space.

The pCpG-4bp composite promoter constructs (n=20) were co-transfected with pRL-CMV into 293T cells. Transfected 293T cells were unstimulated, incubated in hypoxia (0.1% O_2) or stimulated with TNFα (10 ng/ml), PMA (10 ng/ml) or their combination. The normalised luciferase expression was plotted in Figure 3.19.
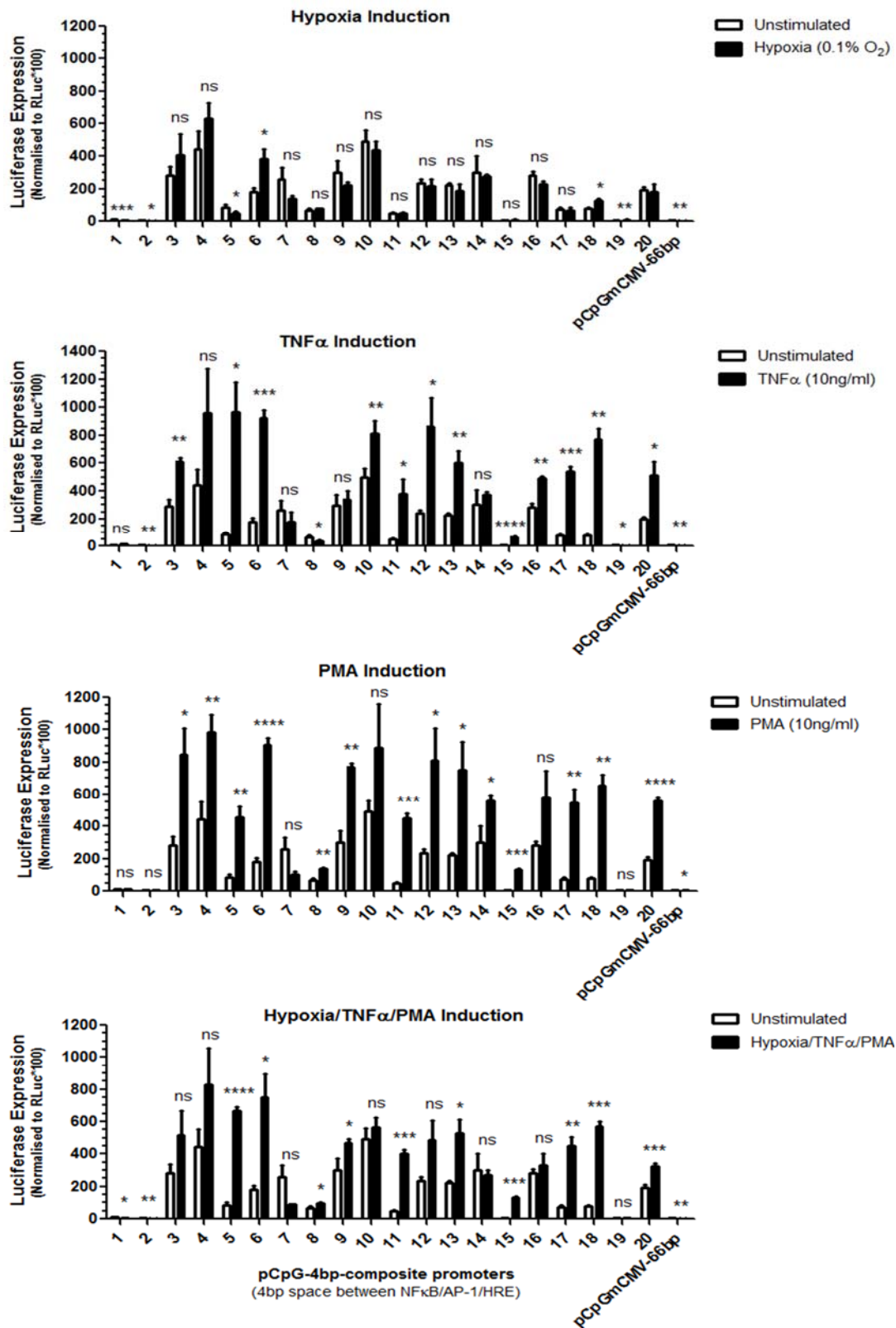
Figure 3.19. The pCpG-4bp-composite promoters display multi-responsiveness and impaired synergistic gene expression. pCpG-4bp-composite promoters (180 ng) were co-transfected with pRL-CMV (20 ng) into 293T cells (20,000 cells in a 96 well-plate) which were either unstimulated, incubated in hypoxia (0.1% O_2) [A], or stimulated with TNFα (10 ng/ml) [B], PMA (10 ng/ml) [C] or their combination [D] for 18 hours. The data represents the mean ± SD of triplicate values normalised

Similar to the pGL3-4bp-composite promoters (section 3.2.2), the pCpG-4bp composite promoters also displayed very high basal luciferase gene expression which were almost comparable to their induced luciferase gene expression. Both of these data strongly support the results in section 3.3.5 (Fig 3.12) which demonstrated that synthetic promoters with closely located TFBSs induce high basal and high induced luciferase gene expression.

Despite the significant increases in gene expression following combined inflammatory and hypoxic stimulation, the majority of the pCpG-4bp-composite promoters exhibited impaired synergistic gene expression when compared to their level of gene expression following stimulation with a single stimulus, which strengthens the notion that the cloning of TFBSs in close proximity impedes the induction of additive/synergistic gene expression.

Overall, the expression profiles of the pGL3-4bp-composite promoters (section 3.2.2) and the revised pCpG-4bp-composite promoters (described in this section), were very similar and both promoter libraries displayed high basal, high induced gene expression, multi-responsiveness and impaired additive/synergistic induction. Therefore, the ability to induce synergistic gene expression by increasing the spacing between the TFBSs was assessed by constructing composite promoters with a 20 bp space between the NFκB, AP-1 and HIF-1α binding sites.

3.5. pCpG-20bp-composite synthetic promoters display differential, multi-responsive and synergistically-inducible luciferase gene expression

Using the Assembly PCR method, the forward and reverse oligonucleotides with NFκB, AP-1 or HIF-1α binding sites, flanked by 20 bp annealing sequences, were pooled together with the standard 5'-NheI and 3'-XhoI 'stop' primers and PCR reaction components. The PCR products were digested with NheI and XhoI and cloned into the equivalent site within the pCpG-5bp cloning vector to introduce a 66 bp space between the proximal TFBS and the TATA box. The resulting pCpG-20bp-composite synthetic promoters (n=20) each had different compositions and numbers of the NFκB, AP-1 and HRE motifs, as schematically depicted in Figure 3.20.
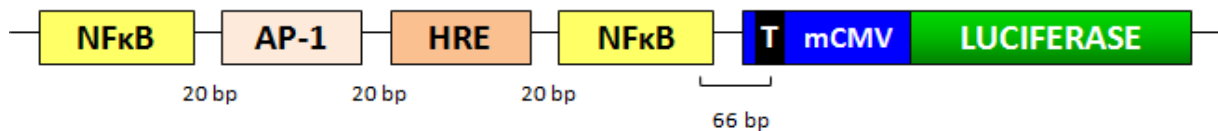
Figure 3.20. A schematic diagram of a pCpG-20bp-composite synthetic promoter. Each TFBS was separated by a 20 bp space. The proximal TFBS and the TATA box (T) were separated by a 66 bp space. The DNA sequences of selected pCpG-20bp-composite promoters are presented in Appendix 10.

The structural components comprising the three distinct composite promoter libraries are compared in Table 3.2.

	pGL3-4bp composite promoters	pCpG-4bp composite promoters	pCpG-20bp composite promoters
Results section	Section 3.3.2	Section 3.4	Section 3.5
Backbone vector	pGL3mCMV	pCpGmCMV	pCpGmCMV
Spacing between the TFBSs	4 bp	4 bp	20 bp
Distance between proximal TFBS and TATA box	45 bp	66 bp	66 bp
NFκB sequence	GGGAATTTC	GGGACTTTCC	GGGACTTTCC
HIF-1α sequence	ACGTGG	ACGTGC	ACGTGC
AP-1 sequence	TGAGTCA	TGAGTCA	TGAGTCA
C/EBPβ, Egr-1 and Ets-1 sites included	Yes	No	No

Table 3.2. Structural comparisons of three composite synthetic promoter libraries. The structures of the pGL3-4bp-composite promoters, pCpG-4bp-composite promoters and the pCpG-20bp-composite promoters are compared.

The pCpG-4bp-composite promoters and the pCpG-20bp-composite promoters possessed the same components with the exception of different spacing between the TFBSs. Therefore, the functional activities of the composite promoters within these two libraries can be reliably compared to determine the effect of increased TFBS spacing on synergistic induction. The pCpG-20bp-composite promoter constructs (n=20) were co-transfected with pRL-CMV into 293T cells and were either unstimulated, incubated in hypoxia (0.1% O_2) or stimulated with TNFα (10 ng/ml), PMA (10 ng/ml) or their combination. The normalised luciferase gene expression was plotted in Figure 3.21.
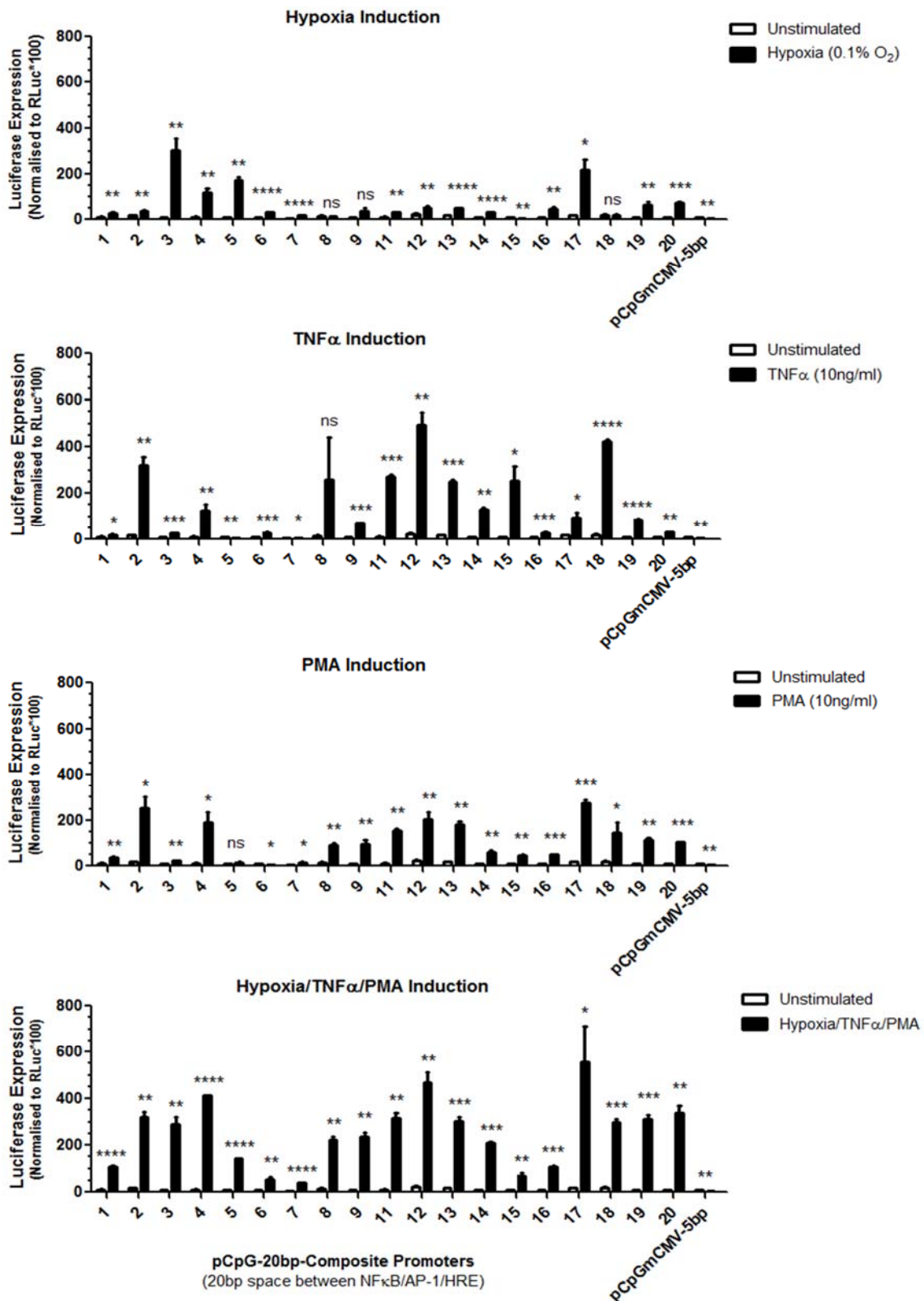
Figure 3.21. pCpG-20bp-composite synthetic promoters display multi-responsiveness and additive/synergistic gene expression. The pCpG-20bp composite promoters (180 ng) were co-transfected with pRL-CMV (20 ng) into 293T cells (20,000 cells in a 96-well plate) which were either unstimulated, incubated in hypoxia (0.1% O_2) [A], or stimulated with TNFα (10ng/ml) [B], PMA (10ng/ml) [C] or their combination [D] for 18 hours. The data represents the

Encouragingly, the synthetic promoters displayed low basal and relatively high luciferase gene expression in response to individual stimuli, which was the expression profile particular to promoters with a 20 bp space between the TFBS and a 66 bp space between the proximal TFBS and the TATA box (as identified in sections 3.3.5 and 3.3.7, respectively). As anticipated, the vast majority of the composite synthetic promoters displayed significant multi-responsiveness to individual stimuli ($p \leq 0.05$ to $p \leq 0.0001$), which can be attributed to the diversity in their sequence compositions. Importantly, all of the promoters exhibited significant increases in luciferase gene expression in response to combined inflammatory and hypoxic stimulation ($p \leq 0.05$ to $p \leq 0.0001$), where numerous promoters displayed additive/synergistic luciferase gene expression. For example, promoters 1, 4, 9, 17, 19 and 20 were multi-responsive to the individual stimuli and were synergistically induced with combined stimuli. Similarly, several promoters displayed additive luciferase gene expression following combined stimulation e.g. promoters 2, 6, 7, 11, 13, 14 and 16. These observations strongly suggest that the steric hindrance associated with the close proximity of TFBSs had been alleviated by increasing the spacing between the TFBSs from a 4bp to 20bp space, which potentially permitted the induction of additive/synergistic gene expression.

As previously speculated in section 3.2.2, the proximal TFBS appeared to influence the overall responsiveness of the promoter to the corresponding stimulus. However, comparisons between the functional analysis data (Fig 3.21, above) and the DNA sequences of the composite synthetic promoters (Fig 3.22, below) revealed that the presence of a single proximal TFBS did not always ensure high responsiveness to the equivalent stimulus. For example, promoter 5 possessed a proximal AP-1 motif and according to my initial proposal, this promoter should have induced the greatest gene expression in response to PMA and/or TNFα stimulation, irrespective of the presence of other TFBSs. Contrary to this prediction, the promoter was more inducible by hypoxia than to PMA or TNFα stimulation, which may have been due to the presence of 3 consecutive HRE motifs followed by a single proximal AP-1 motif.

A noticeable observation was that the synthetic promoters with pairs or triplicates of the same TFBS often had very high responsiveness to the corresponding stimulus but displayed unchanged or slightly decreased gene expression following combined inflammatory and hypoxic stimulation. For example, promoter 18 was comprised of two separate pairs of NFκB motifs (amongst the other TFBSs) and a proximal HRE motif and this promoter was significantly induced by TNFα and PMA but was unresponsive to hypoxia and displayed slightly reduced gene expression in response to combined stimulation (Fig 3.22). Similarly, promoter 3 possessed a pair of HRE motifs (flanked by unpaired AP-1 and NFκB motifs) and also a pair of proximal HRE motifs. This promoter was significantly responsive to hypoxia and marginally, but significantly inducible, by inflammatory stimulation and displayed unchanged gene expression in response to combined stimulation, suggesting that the presence of unevenly distributed clusters of the same TFBS permit high-responsiveness to corresponding stimuli but hinder the induction of strong additive/synergistic gene expression.

In contrast, the synthetic promoters with relatively evenly distributed single TFBSs generally displayed low responsiveness to individual stimuli but exhibited strong synergistic gene expression e.g. promoters 9, 14 and 19. Similarly, the even distribution of paired/triplicate TFBSs appeared to permit high gene expression in response to individual stimuli as well as strong synergistic gene expression e.g. promoter 4 which possessed multiple pairs and/or triplicates of NFκB, AP-1 and HRE motifs (Fig 3.22), which highlighted the importance of evenly distributed and potentially clustered TFBSs on high gene expression in response to individual and combined stimulation.

For simplification, the sequences of the promoters discussed in this section are schematically presented in Figure 3.22 (DNA sequencing data in Appendix 10).

Figure 3.22. Schematic representation of the sequences of selected pCpG-20bp composite promoters. The DNA sequences of the selected composite synthetic promoters from Appendix 10 are schematically presented above.

Overall, the majority of the promoters exhibited the required differential, highly-inducible and multi-responsive gene expression. The introduction of a 20 bp space between the TFBSs appeared to alleviate the steric hindrance associated with a 4 bp space between the TFBSs and permit the induction of additive/synergistic gene expression.

3.6. The pCpG-clustered composite synthetic promoters exhibit stimuli-specific induction of luciferase gene expression

DNA sequence analysis and functional characterisation of the pCpG-20bp composite synthetic promoters (previous section 3.5) highlighted that the synthetic promoters with evenly distributed clusters of the same TFBSs were highly-inducible to individual and combined stimuli and clustered TFBSs appeared to dominate the overall responsiveness of the promoter to the corresponding stimulus. To verify these observations, composite synthetic promoters with randomly cloned clusters of 6NFκB, 8AP-1 or 6HRE and systematically cloned proximal 6NFκB, 8AP-1 or 6HRE motifs were constructed, as depicted in Figure 3.23.

A. The plasmid DNA constructs pCpG-6NFκB, pCpG-8AP-1 and pCpG-6HRE were digested with NheI to serve as cloning vectors with proximal 6NFκB, 8AP-1 and 6HRE motifs, respectively. Re-ligation of the linear vector was prevented by dephosphorylating the DNA ends using CIP enzyme (Fig 3.23 A).
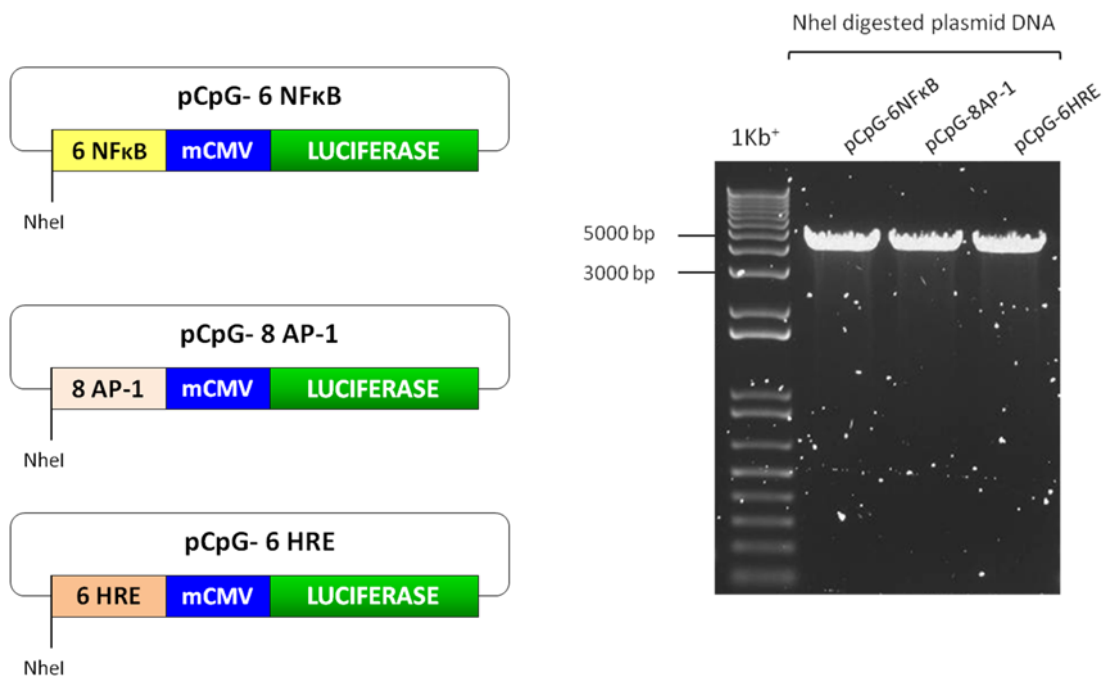
B. The plasmid DNA constructs pCpG-6NFκB, pCpG-8AP-1 and pCpG-6HRE were also digested with NheI and XhoI to isolate the 6NFκB, 8AP-1 and 6HRE fragments (Fig 3.23 B).
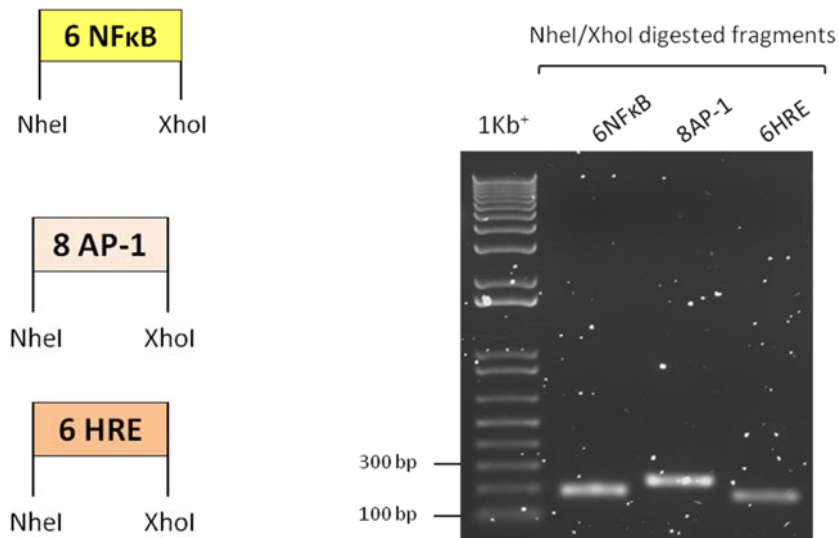
C. Equimolar concentrations of the digested 6NFκB, 8AP-1 and 6HRE fragments were cloned at random into the NheI site within the different vectors (proximal 6NFκB, 8AP-1 and 6HRE vectors) to generate the pCpG-clustered composite promoter constructs (Fig 3.23 C).
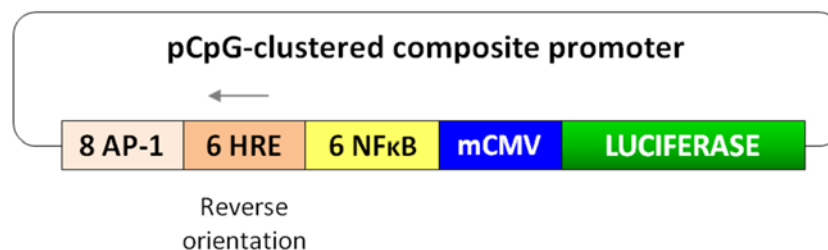
Figure 3.23. Schematic diagram of a pCpG-clustered composite synthetic promoter with randomly cloned clusters of 8AP-1, 6HRE and 6NFκB. Each cluster had a 20 bp space between the TFBSs and the distance between the proximal TFBS and the TATA box was 66bp. The reverse orientation of the middle TFBS is indicated by a left-pointing arrow.

The presence of multimerised clusters within the pCpG-clustered composite promoters was confirmed by an EcoRI/StuI analytical digest and gel electrophoresis (Fig 3.24).

[
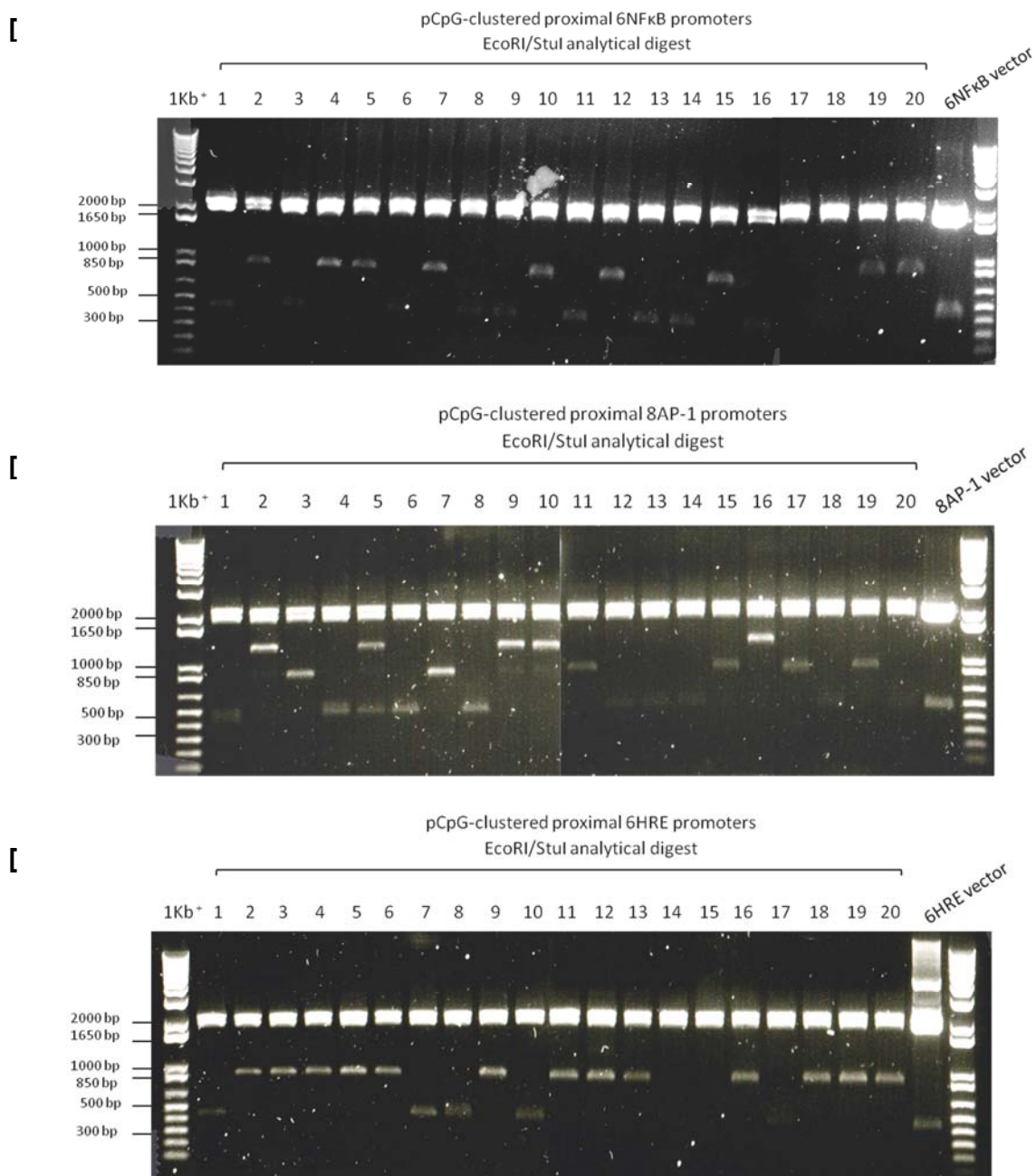
Figure 3.24. Analytical restriction enzyme digest of pCpG-clustered composite promoters. The plasmid DNA constructs with proximal 6NFκB [A], proximal 8AP-1 [B] and proximal 6HRE [C] motifs were digested with EcoRI and StuI to release the PCR product (and an excess of 247 bp). Therefore, the size of the promoter was calculated by subtracting 247 bp from the size of the excised fragment using the

Figure 3.24 shows the EcoRI and StuI restriction digests of the pCpG-clustered composite promoters with proximal 6NFκB, 8AP-1 and 6HRE motifs (Fig 3.24 A-C, respectively). Due to the cloning method, the release of EcoRI/StuI digested fragments from the pCpG-clustered composite promoters larger than the fragments isolated from their respective vector, indicated that the promoter contained at least 3 clustered motifs (6NFκB, 8AP-1 and/or 6HRE), two of which differed from the proximal TFBS (DNA sequencing: Appendix 11). The pCpG-clustered composite promoters containing at least 3 clustered motifs were selected and co-transfected with pRL-CMV into 293T cells (n=8, per group). Transfected cells were unstimulated, incubated in hypoxia (0.1% O_2) or stimulated with TNFα (10 ng/ml), PMA (10 ng/ml) or their combination. The normalised luciferase gene expressions were plotted in Figure 3.25.
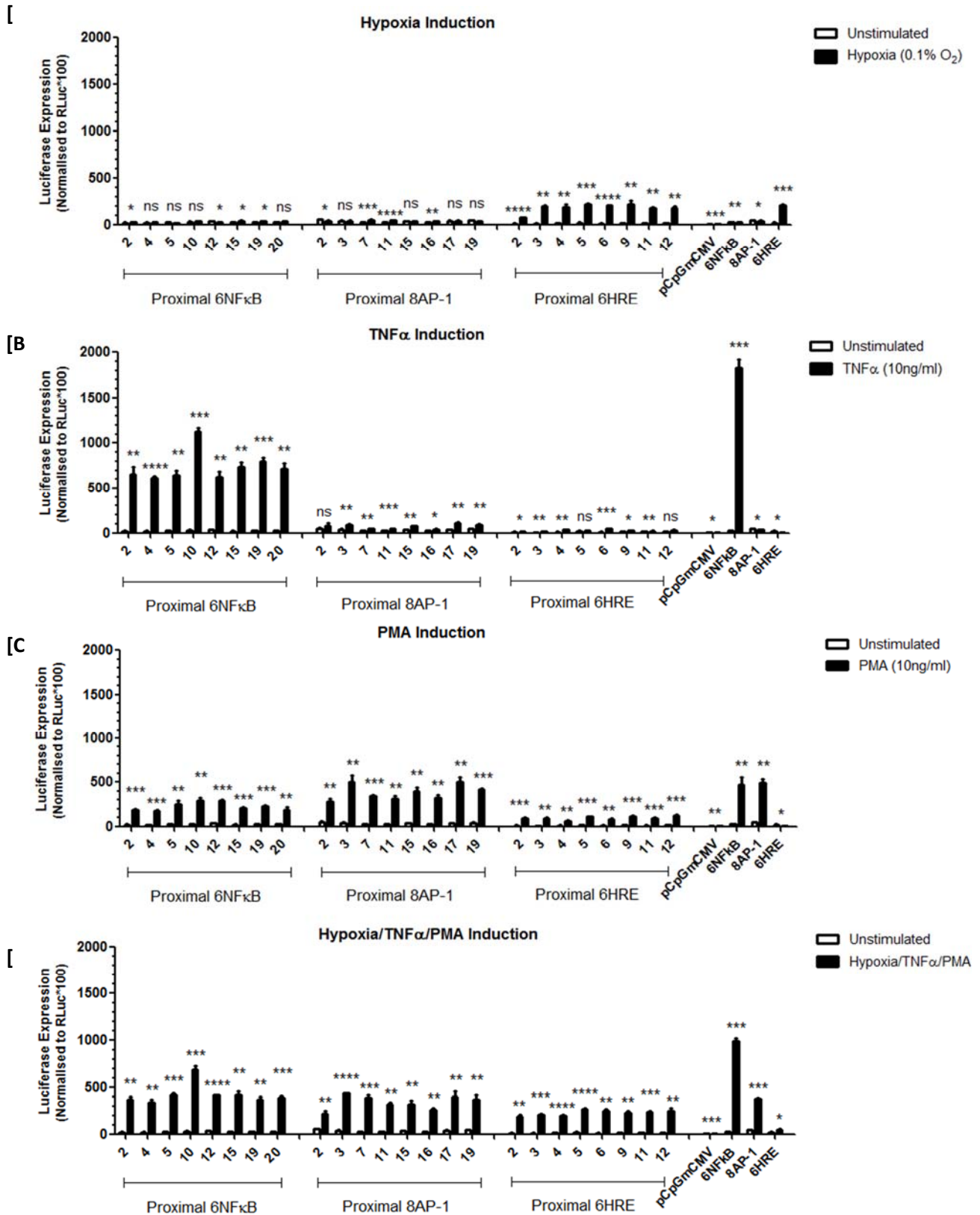
Figure 3.25. Functional analysis of pCpG-clustered composite promoters. The pCpG-clustered composite promoters (180 ng) were co-transfected with pRL-CMV (20 ng) into 293T cells (20,000 cells in a 96-well plate), which were either unstimulated, incubated in hypoxia (0.1% O_2) [A] or stimulated with TNFα (10 ng/ml) [B], PMA (10 ng/ml) [C] or their combination [D] for 18 hours. The data represents the mean ± SD

221

Generally, the promoters exhibited significant increases in gene expression to different individual stimuli but importantly, the proximal TFBSs appeared to dictate the overall responsiveness of the composite promoter to the corresponding stimulus e.g. composite promoters with proximal 6HRE displayed the greatest hypoxia-inducibility ($p \leq 0.01$ to $p \leq 0.0001$), irrespective of the presence of HRE motifs within the proximal-6NFκB and -8AP-1 promoters. Similarly, proximal-6NFκB and -8AP-1 promoters were significantly responsive to TNFα and PMA respectively, but marginally inducible by hypoxia.

Interestingly, despite the statistically significant increases in luciferase gene expression following combined stimulation, the composite synthetic promoters with proximal 6NFκB and 8AP-1 had impaired synergistic gene expression which is in agreement with the previous observation that composite promoters with paired/triplicate TFBSs displayed unchanged or slightly reduced gene expression in response to combined inflammatory and hypoxic stimulation (section 3.5). In contrast, the composite promoters with clustered proximal 6HRE displayed marginally increased luciferase gene expression following the same treatment. Strikingly, the promoters with proximal 6NFκB motifs were highly and significantly responsive to TNFα but displayed markedly lower gene expression (upto 50% reduced fold induction) in response to the combined inflammatory and hypoxic stimulation. An important question to emerge from these observations is 'which TFBS should be positioned closest to the TATA box?'. This is an important consideration when constructing multi-responsive composite synthetic promoters because the promoter will need to be constructed according to the most dominant pathological feature of the disease under study. With focus on RA, inflammation is the main pathological feature and therefore, the composite promoters are required to be highly inflammation-inducible which can be achieved by positioning AP-1 or NFκB, preferably clustered motifs, proximal to the TATA box. However, as demonstrated in the results above, the presence of clustered proximal NFκB and AP-1 motifs hinders the induction of synergistic gene expression and unless optimised, precludes the further application of composite promoters with clustered proximal NFκB and AP-1 motifs.

3.7. Discussion

An integral component of this thesis was to engineer and functionally characterise composite synthetic promoters, which were comprised of randomly cloned core TFBS for inflammation and/or hypoxia-responsive TFs. The experiments in Chapter 3 expanded on the concept of exploiting the endogenous activity of diverse TFs which are activated during inflammation and/or hypoxia, to transcriptionally regulate therapeutic gene expression using composite synthetic promoters. The functional activities of the pGL3-4bp-composite synthetic promoters revealed several structural parameters which negatively impacted the transcriptional activity of the promoter and highlighted the requirement to optimise the composite promoter structure. An alternative Assembly PCR method was implemented to control the spatial arrangement of TFBSs and the functional activities of the resulting promoters revealed that increased spacing between the TFBS and also increased spacing between the proximal TFBS and the TATA box resulted in decreased basal and induced luciferase gene expression. Application of the optimal spatial preferences of TFBSs, relative to one another and also to the TATA box, generated improved, highly-inducible, multi-responsive and synergistically-inducible composite synthetic promoters, which are promising novel candidates for local RA gene therapy.

The feasibility of synthetic promoters for disease-regulated gene therapy has been successfully demonstrated by a number of publications, for example, an NFκB-responsive synthetic promoter has been used to control anti-TNFα therapeutic gene expression in an inflammation-inducible manner (Khoury *et al.,* 2007) and HIF-1α-responsive synthetic promoters are commonly used for cancer gene therapy (Shibata *et al.,* 2000). However, these synthetic promoters respond to limited stimuli which restrict their application to the treatment of diseases characterised by a single dominant pathological condition. The concept of combining TFBSs within composite promoters has significant advantages over single-responsive promoters for RA gene therapy. Notably, the use of composite promoters for local RA gene therapy represents a means to harness the endogenous activity of different TFs to regulate therapeutic gene expression in response to multiple pathological stimuli in the RA

joint and the potential to induce synergistic therapeutic gene expression, which is an ideal system for RA gene therapy.

3.7.1. Composite synthetic promoters with a 4bp space between TFBS exhibit impaired synergistic gene expression, potentially due to steric hindrance of TFs

Section 3.2.2 proceeded with the construction of the initial pGL3-4bp-composite synthetic promoter library, which were comprised of the core binding sites of NFκB, HIF-1α, AP-1, C/EBPβ, Egr-1 and Ets-1, known to be activated in RA. The design of the composite synthetic promoters adhered to the requirements necessary for eukaryotic gene transcription; the synthetic promoters were comprised of TFBSs and a minimal cytomegalovirus (mCMV) promoter which consisted of a TATA box positioned in the forward orientation to direct RNA polymerase II transcription (Wang and Stumph, 1995).

The conventional method for constructing synthetic promoters, as demonstrated by others, involves PCR amplification of the region of an endogenous promoter which contains the required TFBSs and cloning the fragment upstream of the transgene. For example, Geurts *et al.,* (2009) rationally designed synthetic promoters suitable for RA gene therapy by combining computational analysis and experimental verification which identified that the binding sites of NFκB, AP-1 and C/EBPβ were over-represented within the proximal promoters (region containing TFBSs) of endogenous genes in the synovial tissues of mice with collagen-induced arthritis (CIA). Their studies led to the identification of the inflammation-inducible serum amyloid A3 (*Saa3*) proximal promoter (containing 3 C/EBPβ motifs and possibly other TFBSs), which was PCR cloned upstream of the IL-1 receptor antagonist (IL-1ra) therapeutic gene to transcriptionally regulate IL-1ra expression during inflammation. Based on their striking observation that numerous endogenous promoters of genes upregulated in CIA mice were comprised of different TFBSs i.e. NFκB, AP-1 and C/EBPβ, the studies of Geurts *et al.,* (2009)

strengthened our concept of engineering composite synthetic promoters, using the above candidate TFBSs, for RA gene therapy.

However, repeated PCR cloning of endogenous proximal promoters would not be a feasible method to generate compact composite promoters with multiple TFBSs, as this would require numerous PCR amplifications of multiple genes containing each of the TFBSs under study. This cloning strategy would result in extremely large synthetic promoters that could potentially hinder further cloning applications into limited capacity vectors. Furthermore, this cloning method would amplify the TFBSs as well as the intervening sequences between the TFBSs, which may exert unwarranted regulatory effects on the synthetic promoter activity. Therefore, oligonucleotides containing only the core TFBSs of the candidate TFs were randomly ligated via the NheI overhangs, which resulted in the construction of compact composite promoters with diverse arrangements and numbers of the TFBSs. The synthetic promoters were cloned upstream of the firefly luciferase reporter gene, which enabled an efficient functional readout of the promoter activity.

The pGL3-4bp-composite promoters (section 3.2.2) were screened for the required expression profiles; low basal and highly induced gene expression, multi-responsiveness and synergistic-inducibility. The vast majority of the composite promoters demonstrated high basal luciferase gene expression and the gene therapy application of these synthetic promoters would result in 'leaky' therapeutic gene expression during the remission phase of the disease, which is an undesirable attribute. Unwarranted high concentrations of the therapeutic protein during remission could increase the risk of immunosuppression and other adverse effects thereby decreasing the safety and efficacy of the therapy. However, low/moderate basal therapeutic gene expression is essential to sustain remission and to prevent the disease from rebounding, which demonstrates the complexity of generating tightly-regulated synthetic promoters.

The induction of synergistic gene expression by the composite synthetic promoters is crucial to the development of a novel gene therapy for RA, however, the vast majority of the pGL3-4bp-composite promoters displayed unchanged or reduced luciferase gene expression following combined inflammatory and hypoxic stimulation, which may have been due to the close proximity of the TFBSs. The relevance of optimal spacing between the TFBSs within the composite synthetic promoters can be logically explained by the endogenous TF-DNA interactions during transcription. Upon activation, TFs rapidly locate and bind to their TFBS along the DNA to regulate expression of the downstream gene(s). Published thermodynamic principles of TF-DNA interactions report that the search mechanism of TFs involves a combination of 3D diffusion through the cell volume and 1D sliding along the DNA, which has been a widely accepted model over the past 30 years (Berg *et al.,* 1981; Halford and Marko, 2004). These principles of TF-DNA interactions can be applied to interpret the interactions between the candidate TFs and the composite synthetic promoters with a 4bp space between the TFBSs. During combined hypoxic and inflammatory stimulation, the TFs become activated and diffuse through the cell volume and bind to the DNA where they adopt one of two conformations; the TFs loosely bind to the DNA and slide along the DNA by 1D diffusion in search of their respective binding sites or the TFs become immobile by tightly associating with the DNA (Hu *et al.,* 2008). During the sliding search process, it is likely that DNA-bound TFs impede the efficient search processes of other TFs and loosely bound TFs are possibly dislodged. Once the TF has successfully located its binding site within the synthetic promoter, it undergoes an irreversible conformational transition to a final bound state (Hu *et al.,* 2008). Therefore, the impaired synergistic expression by pGL3-4bp-composite promoters may have been a consequence of cloning the TFBSs in close proximity where the irreversibly bound TF could have sterically hindered the binding of another TF to the adjacent binding site, rendering some TFBSs inaccessible.

The impaired synergistic induction may have also been a consequence of DNA torsional stress induced by the arrangement of the TFBSs within the synthetic promoters. One turn of the DNA helix is approximately 10.5 bp (Wang, 1979) and a 4bp space between the TFBS, equates to less than half a turn of the DNA helix. This positions the TFBSs on alternating sides of the DNA which can induce additional DNA torsional stress and impede efficient transcription (Schleif, 1992). The topological state of the DNA is a major regulator of the transcriptional process and it is well known that RNA polymerase II introduces a degree of torsional stress in DNA during transcription which effects the expression of downstream gene(s) (Mirkin, 2001). During transcription, the DNA loops to allow remote regulatory elements to interact with the transcription initiation complex assembled on the TATA box (Ptashne, 1986). In response to combined inflammatory and hypoxic stimulation, the accessible TFBSs within the composite synthetic promoters become occupied by their respective TFs. Looping and repeated twisting to allow the interaction of each TF-TFBS complex with the TATA box may have induced excessive DNA torsional stress which potentially decreased the efficiency of the transcriptional process and consequently impaired the induction of synergistic gene expression. DNA torsional stress and subsequent supercoiling of closed circular plasmid DNA are energetically unfavourable, therefore the changes to relax the local DNA become favourable (Benham, 1979). A possible means to relax the DNA can be to avoid repeated twisting by excluding energetically unfavourably positioned TF-TFBS complexes and only allowing the TF-TFBS complexes positioned every 10-12bp (equal to integral multiples of DNA helical repeats) to interact with the TATA box. This may explain why the vast majority of the pGL3-4bp-composite promoters containing one or more HRE motifs (flanked by other TFBS) failed to display hypoxia-inducibility.

Interestingly, the pCpG-4bp-composite promoters (section 3.4) displayed very similar expression profiles to the pGL3-4bp-composite promoters, as discussed above. Both composite promoter libraries were constructed using the random ligation cloning method and demonstrated high basal and high induced luciferase gene expression, multi-responsiveness

to individual stimuli and impaired synergistic gene expression following combined stimulation, irrespective of dissimilar structural components between the two promoter libraries. These observations supported the implication of a 4bp space between TFBSs as the cause of hindered synergistic gene expression and encouraged the reconstruction of the synthetic promoters using an alternative cloning method.

3.7.2. Optimisation and application of the Assembly PCR method to fine-tune constraining parameters of gene expression

Section 3.3 provided a detailed account on the optimisation of the Assembly PCR method and its application to generate synthetic promoters which demonstrated that the spatial organisation of the TFBSs had profound effects on the luciferase gene expression induced by synthetic promoters.

The Polymerase Chain Assembly (PCA) method is a two-step PCR technique traditionally used to synthesise genes from oligonucleotides comprising the sequence of the gene of interest: the overlapping oligonucleotides are assembled using PCA and then amplified using PCR to create the synthetic gene (Stemmer *et al.* 1995). The Assembly PCR method is an adaptation of the PCA method and research conducted by Team Heidelberg (International Genetically Engineered Machine (iGEM) Competition, 2009) has demonstrated the feasibility of the Assembly PCR method to construct single-responsive synthetic promoters. However, the functional activities of their NFκB-responsive synthetic promoters varied considerably, with some promoters displaying minimally induced GFP reporter gene expression. Also, the gene expression induced by the most responsive promoter did not exceed 2.6 fold induction in response to 2.5 μM TNFα in U2OS cells. The studies of Team Heidelberg also reported that their HIF-1α and p53 synthetic promoters failed to induce sufficient gene expression. The incorporation of random spacer oligonucleotides and oligonucleotides with random TFBS sequences generated libraries of promoters with extremely variable TFBS organisation and

TFBS sequence composition which was reflected by their highly variable promoter strengths. Their concept of constructing spatially diverse promoters, from a single PCR reaction, represented an inefficient method of identifying the optimal spacing between TFBS, which was one of my objectives. Therefore, the Assembly PCR method described in section 3.3 adhered to the method by Team Heidelberg but with systematic and optimised modifications to rationally design single- and composite-synthetic promoters.

The oligonucleotides used in this thesis were modified to incorporate the TFBSs flanked by two annealing sequences of specified lengths, in individual PCR reactions. This rational approach to generating different libraries of promoters with specific TFBS spacing facilitated the investigation on the effect of increased spacing between TFBSs on gene expression. Section 3.3 described the optimisation of the Assembly PCR method, which expanded on the observations by Wu *et al*. (2006) and TerMaat *et al.,* (2009) who showed that the efficiency of the PCA reaction is affected by several parameters i.e. the concentrations of assembly oligonucleotides and amplification primers and the choice of polymerase enzyme. These parameters provided a guideline for the optimisation of the Assembly PCR reaction described in this thesis.

The use of Phusion DNA polymerase enzyme, which is a high-fidelity enzyme with 50 times greater fidelity than Taq polymerase (www.neb.com), ensured correct and efficient assembly and amplification of the PCR products, which was fundamental for the error-free construction of synthetic promoters. The 'stop' oligonucleotide concentration in the initial assembly reaction was also optimised and demonstrated that the PCR product size decreased with increasing 'stop' oligonucleotide concentration. In turn, the successful incorporation of the 'stop' oligonucleotides increased the efficiencies of the subsequent amplification reaction, restriction enzyme digestion, ligation and cloning steps, which strongly corroborated the results of Wu *et al.,* (2006) and TerMaat *et al.,* (2009).

The PCR products were cloned into the pCpGmCMV vector which had more suitable properties for *in vitro* and potential *in vivo* gene expression analysis than the pGL3mCMV vector. The pCpGmCMV vector is a modified construct of the pCpG free-mSEAP plasmid which is completely devoid of CpG dinucleotides. It has been demonstrated that CpG-devoid plasmid DNA has significantly increased long-term gene expression compared to CpG-replete analogues (Hodges *et al.,* 2004) where the presence of unmethylated CpG dinucleotides can elicit strong immunogenic responses (Krieg *et al.,* 1995) which may hinder the safety and efficacy of the gene therapy. The pCpGmCMV vector also possesses two matrix attachment regions (MARs) from the 5'-region of the human IFN-β gene and the β-globin gene, which function as insulators by preventing external enhancers from influencing the promoter when placed between them, to enhance long-term gene expression and prevent expression variability (Recillas-Targa *et al.,* 2002). The absence of CpG dinucleotides and the action of 'enhancer blocking' properties of the pCpGmCMV vector limits toxicity and increases the durability of gene expression. This vector was intended for *in vitro* promoter analysis however, due to the aforementioned attributes of the pCpGmCMV vector, the pCpG-promoter constructs have the potential to be delivered *in vivo*, thereby serving as a reserve plan should the promoters fail to induce significant gene expression in the lentiviral constructs.

Endogenous promoters consist of many conserved properties therefore it was important to adhere to the favourable characteristics of endogenous promoters as well as incorporating additional properties when constructing the synthetic promoters. One common feature of endogenous promoters is diversity within the TFBSs among different genes, which mediates divergence in gene expression between different cells and species, to control the gene activity levels globally (Wittkopp, 2010). The PCR component of the Assembly PCR reaction was exploited to introduce sequence diversity within the general NFκB consensus sequence, GGGRNNYYCC, to determine the effect of NFκB sequence diversity on the activity of promoters with variable NFκB motifs. As anticipated, the variable NFκB promoters displayed diverse expression profiles in response to TNFα, which can be explained by the promiscuous

nature of TFs, which bind to the different variations of their consensus sequence with different affinities to induce diverse gene expression (Bradley *et al.,* 2010). Noticeably, the number of variable NFκB sites had no effect on gene expression and a striking observation was that the two most active variable NFκB promoters possessed NFκB sequences which were very similar to the fixed NFκB sequence, which highlighted that the TF binding sequence is more important for achieving high-level gene expression than TFBS copy number, as also demonstrated by Sharon *et al.,* (2012).

Interestingly, the fixed NFκB promoter consistently induced high luciferase gene expression following TNFα stimulation where the magnitude of gene expression was dependent on fixed NFκB copy number within the promoter. However, saturation was observed for promoters with 6 or more fixed NFκB motifs where further addition of NFκB motifs within the promoters marginally increased gene expression. A similar observation was described by Shibata *et al.,* (2000) who demonstrated that increasing the number of HRE motifs within the synthetic promoters increased hypoxia-responsiveness upto 5HRE motifs, at which point the luciferase gene expression reached a maximal gene expression plateau. Taken together, the observations from this experiment led to the incorporation of the 'high affinity' fixed NFκB binding sites in the composite promoters to favour high inflammation-inducibility.

3.7.3. Increased spacing between the TFBSs and the proximal TFBSs relative to the TATA box negatively impacts gene expression of synthetic promoters.

The Assembly PCR method was implemented to explore the functional significance of increased spacing between the TFBSs and also to determine whether the changes in the position of the proximal TFBS relative to the TATA box influences gene expression. The experiments in section 3.3.5 confirmed the importance of optimal structural arrangements of the TFBSs and its effect on the activity of the synthetic promoter.

Nine NFκB and AP-1 promoter libraries with 15 bp-60 bp spaces between the NFκB or AP-1 binding sites were generated using the Assembly PCR method. The functional activities of these promoters confirmed the influence of TFBS spatial arrangement on gene expression and displayed a clear trend; synthetic promoters with closely located TFBSs displayed high basal and high induced luciferase gene expressions compared to lower gene expressions induced by promoters with sparse TFBSs.

Many *in silico* studies have developed computational models which correlate the structural properties of active endogenous promoters to their gene expression profiles, for example, Gotea *et al.,* (2010) revealed that endogenous promoters often possess clusters of the same type of TFBSs (homotypic clusters), which appear to play an important role in the gene regulation in human and other vertebrae genomes (Gotea *et al.,* 2010). Other studies have suggested that the presence of homotypic TFBS within locally dense clusters might enhance TF recruitment and the TF search process (Brackley *et al.,* 2012) and increase gene expression (Sharon *et al.,* 2012), which is consistent with the expression profile trend presented in Figure 3.12. For example, the synthetic promoters with a 15bp space between the TFBSs displayed the highest basal and induced luciferase gene expressions which could be attributed to the high local density of the clustered homotypic TFBSs (NFκB or AP-1 binding sites) and facilitated the TF search process which induced higher levels of gene expression compared to promoters with sparse TFBS (i.e. 45-60 bp space).

Encouragingly, this appeared to be a general trend among NFκB and AP-1, which belong to two distinct transcription factor families. In addition, the pGL3-4bp- and pCpG-4bp-composite promoters also comply with this trend, as these promoters had a 4 bp space between the TFBSs and exhibited very high basal and induced luciferase gene expression levels. The experiment described in section 3.3.5 confirmed that changes in promoter architecture could affect gene expression and the identified trend in Figure 3.12 served as a reliable filtering parameter to exclude promoters with unfavourable attributes, i.e. promoters which induced high basal 'leaky' expression and also promoters which failed to induce sufficient gene

expression following stimulation. Therefore, a 20 bp space was selected as the optimal spacing between the TFBSs as these promoters display the required low basal and high induced luciferase gene expression.

During transcription, RNA Pol II is recruited to the TATA box (Parker and Topol, 1984) and ample evidence supports the fundamental role of the TATA box during transcription (Breathnach and Chambon, 1981). This section described the construction and functional analysis of NFκB- or AP-1-responsive promoters with 55 bp – 74 bp spacer insertions between the proximal TFBS and the TATA box. Interestingly, Geurts *et al.,* (2009) observed that the majority of over-represented NFκB and AP-1 binding sites present in the proximal promoters of genes upregulated in the tissues of CIA mice were spatially conserved and typically located -170/-50 bp and -290/-30 bp upstream of the TATA box, respectively, which is in line with the location of the proximal NFκB and AP-1 within the synthetic promoters described in section 3.3.7.

Although the NFκB and AP-1 sites were positioned in the conserved spatial window, these TFBSs were removed from their normal genomic context and the respective TFs may have performed differently within the synthetic promoters. Interestingly, the results in section 3.3.7 showed that systematically increasing the distance between the proximal TFBS (NFκB or AP-1) and the TATA box within the synthetic promoters resulted in reduced basal and induced luciferase gene expression levels. These results are in line with many studies that demonstrated, using synthetic promoter variants, that expansion of the distance between upstream elements i.e. TFBSs and the TATA box can reduce transcriptional activation (Guarente and Hoar, 1984; McKnight, 1982; Takahashi *et al.,* 1986, Wu and Berk, 1988; Smith *et al.,* 1995; Dobi and Winston, 2007; Sharon *et al.,* 2012). Although their observations may depend on the origin of the tested promoter, taken together, these publications suggest that the transcriptional initiation process requires specific alignments between TFBSs and the proteins assembled on the TATA box. Therefore, it is likely that positioning the proximal TFBS closer to the TATA box facilitates their interactions, which favour high gene expression.

Overall, the optimisation and application of the Assembly PCR method to construct synthetic promoters has offered the clear advantage of a high-throughput, inexpensive and efficient method to systematically design and construct synthetic promoters. The experiments discussed in this section have confirmed the biological implications of TFBS sequence diversity and TFBS spatial preferences on gene expression and have enabled the identification of 20bp as the optimal spacing between the TFBS and 66bp as the optimal distance between the proximal TFBS and the TATA box, which were applied to construct improved composite synthetic promoters.

3.7.4. Comprehensive analysis of gene expression induced by pCpG-4bp-composite promoters, pCpG-20bp-composite promoters and pCpG-clustered composite promoters.

Section 3.5 described the modifications applied to the composite promoter architecture which introduced a 20bp space between randomly cloned NFκB, HRE and AP-1 motifs with a 66bp space between the proximal TFBS and the TATA box. Encouragingly, the functional activities of these pCpG-20bp-composite promoters demonstrated that the vast majority of these promoters were significantly multi-responsive, displayed low basal and high induced gene expression and additive/synergistic gene expression following combined inflammatory and hypoxic stimulation, all of which are fundamental requirements of composite synthetic promoters for RA gene therapy. The induction of additive/synergistic gene expression strongly suggests that the spatial hindrance associated with a 4bp space between the TFBSs, (as demonstrated in the pGL3-4bp and pCpG-4bp-composite promoters) had been alleviated by increasing the spacing between the TFBSs from 4bp to 20bp which permitted synchronous binding of the TFs to induce additive/synergistic gene expression.

Interestingly, the DNA sequencing data of the pCpG-20bp-composite promoters revealed that promoters displaying high induction to a particular stimulus possessed pairs/triplicates of the

same TFBS (homotypic cluster), which supported the results of Gotea *et al.,* (2010) and also, the results in section 3.3.2 demonstrating that promoters with dense homotypic TFBSs (within clusters) induced greater gene expression than promoters with sparse TFBSs. The sequencing data also highlighted that the presence of paired proximal TFBSs governed the overall responsiveness of the promoter to a particular stimulus, irrespective of the presence of other TFBSs. However, the presence of unevenly distributed paired/triplicate TFBSs appeared to confer high inducibility to an individual stimulus but hindered the induction of strong additive/synergistic gene expression. In contrast, the composite promoters with evenly distributed paired/triplicate TFBSs displayed similar expression levels to each individual stimulus and exhibited robust synergistic gene expression. These observations comply with the finding that clustered TFBSs induce high gene expression in response to the corresponding stimulus.

The subsequent results in section 3.6 confirmed that the systematic positioning of 6NFκB, 8AP-1 or 6HRE clusters proximal to the TATA box dominated the overall responsiveness of the pCpG-clustered composite promoter to TNFα, PMA and hypoxia stimulation, respectively. However, composite promoters with clustered proximal 6NFκB and 8AP-1 exhibited impaired synergistic gene expression whilst the proximal 6HRE appeared to permit modest additive gene expression, which strengthen the observations in section 3.5.

Overall, hypoxia-inducibility is a favourable characteristic but not a prerequisite of composite synthetic promoters for RA gene therapy. However, high inflammation-inducibility and the induction of synergistic gene expression are undoubtedly required. The induction of high inflammation-inducible gene expression at the expense of synergistic induction by the composite promoters with clustered proximal 6NFκB and 8AP-1 precluded the further application of clustered composite promoters. Nevertheless, the introduction of optimal TFBS spatial arrangements permitted additive/synergistic gene expression by the pCpG-20bp-composite synthetic promoters with randomly arranged TFBSs.

CHAPTER 4:

In Vitro and *In Vivo* Translational Studies of

Luciferase Gene Expression from Lentiviral

Integrated Synthetic Promoters

4.1. Introduction

The *in vitro* characterisation of transiently transfected synthetic promoter constructs in Chapter 3 enabled the selection of eight candidate promoters which displayed favourable gene expression profiles: promoters 2, 9, 11, 12 and 14 (pCpG-20bp composite promoters), promoters 4 and 6 (pCpG-clustered proximal 6HRE) and promoter 245 (pGL3-4bp-composite promoter). The synthetic promoters were further characterised by cloning into lentiviral vectors (LV) which offer the advantage of allowing integration of the synthetic promoter and transgene into the genome of dividing and non-dividing cells to confer long-term and stable transgene expression, which is ideal for experimental gene therapy strategies.

4.1.1. Construction of Lentiviral Composite Synthetic Promoters for Regulation of Luciferase Gene Expression

The self-inactivating (SIN) lentiviral plasmid, pLV.CMV.eGFP (Addgene plasmid 30471; Barde *et al.,* 2011) was used as the primary lentiviral cloning vector. The schematic diagram below illustrates the cloning strategy used to clone the candidate composite synthetic promoters into the lentiviral plasmid DNA (Fig 4.1.1.).

A. pLV.CMV.eGFP was digested with BamHI and SalI to release the eGFP gene. The resulting 8357 bp fragment served as the pLV.CMV cloning vector (Fig 4.1.1 A).
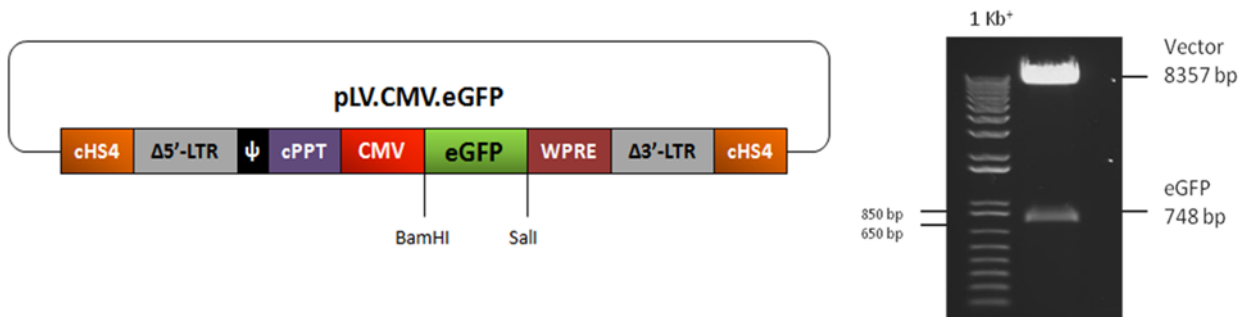
B. The luciferase gene was PCR amplified from pGL3mCMV using forward and reverse PCR primers with BamHI and SalI restriction sites, respectively (Fig 4.1.1 B).
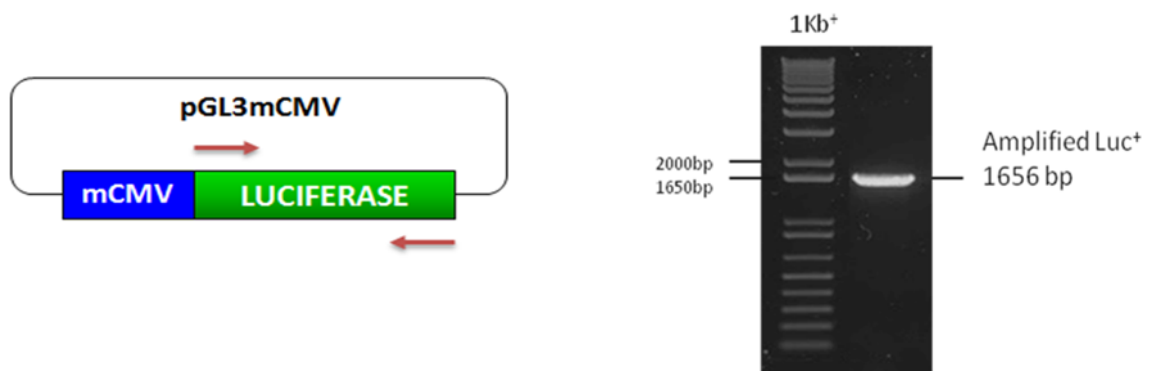
C. The luciferase PCR product was digested with BamHI and SalI and cloned into the equivalent site within the pLV.CMV vector to generate pLV.CMV.Luc$^+$ (DNA sequence in Appendix 13.1). The resulting plasmid was digested with PmeI and BstBI to release the CMV promoter and a 5'-portion of the luciferase gene. The 7882 bp fragment served as the cloning vector, pLV.Luc$^+$(Fig 4.1.1 C).
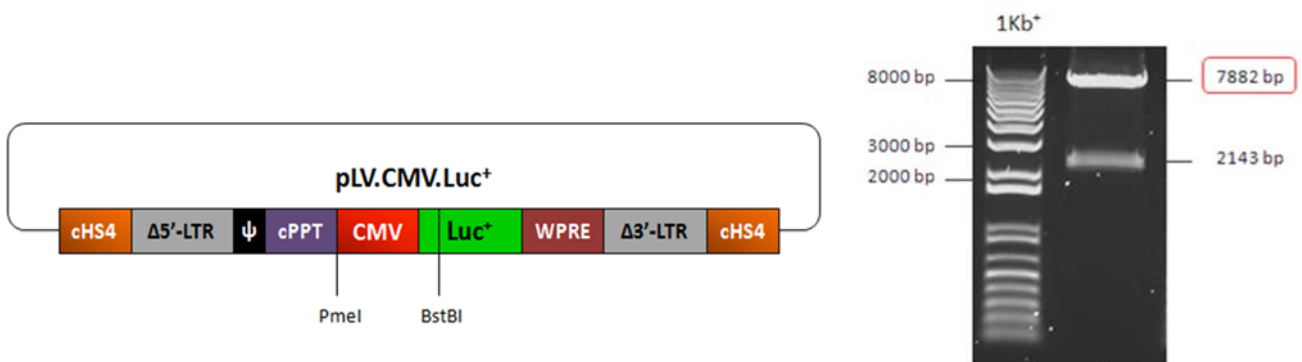
D. The candidate synthetic promoters (n=8) were amplified from their equivalent pCpG- or pGL3-composite promoter constructs using forward ClaI and reverse BstBI PCR primers. The synthetic promoters and the 5'-portion of the luciferase gene were amplified (Fig 4.1.1 D).
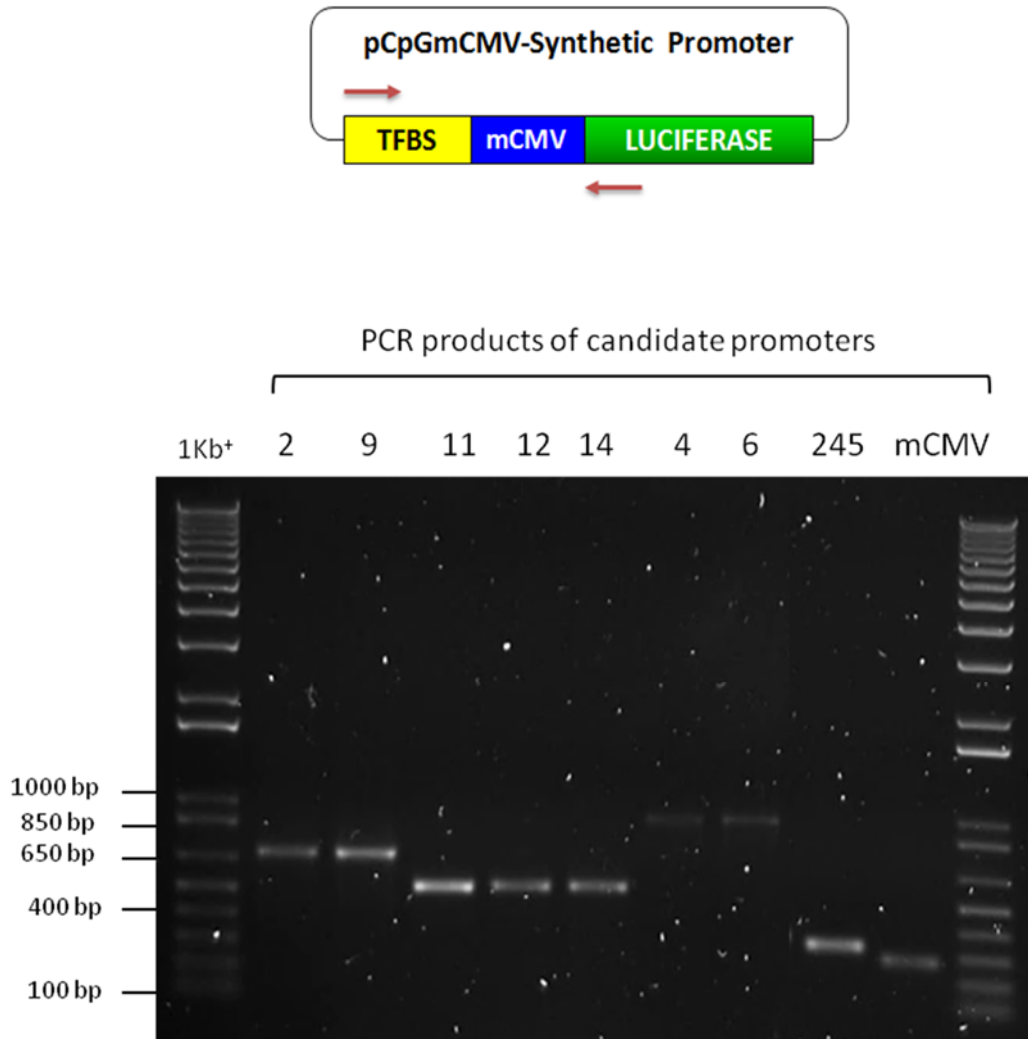
E. The PmeI restriction enzyme produces blunt ended DNA fragments. Therefore the PCR products were digested with ClaI and the 5'-overhang was blunted using Klenow enzyme to generate blunt ended PCR products. Following Klenow treatment, the PCR products were digested with BstBI and ligated with the pLV.Luc⁺ vector to generate a lentiviral plasmid containing the composite synthetic promoter and the restored luciferase gene (Fig 4.1.1 E). The resulting lentiviral-luciferase constructs (transfer plasmids) were confirmed by DNA sequencing (Appendix 13.2- 13.10).
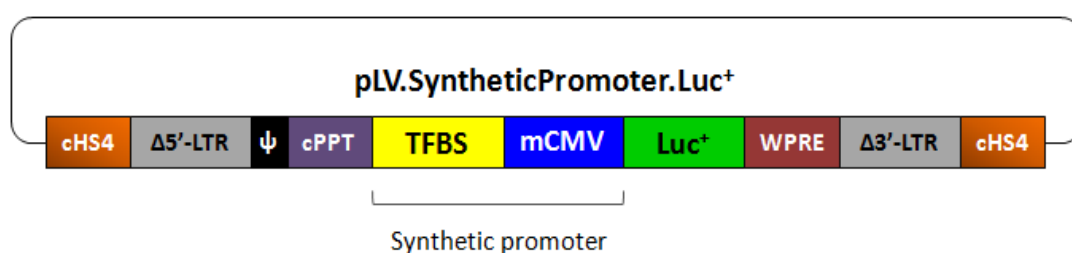
F. Following the successful construction of the lentiviral synthetic promoter constructs, lentiviral particles were produced using a second generation packaging system. This involved a three-plasmid transient transfection into 293T cells with the transfer, packaging and envelope plasmids to generate lentiviral particles pseudotyped with the VSV-G envelope glycoprotein. The VSV-G pseudotyped lentiviral particles were transduced into 293T cells to generate a stable 293T cell line with the integrated synthetic promoter and luciferase gene.

Figure 4.1.1. Schematic diagram illustrating the cloning method used to generate the lentiviral composite synthetic promoter constructs expressing the luciferase gene. The pLV.CMV.eGFP lentiviral plasmid was systematically modified to generate the lentiviral constructs expressing the luciferase gene under the control of inflammation-inducible composite synthetic promoters.

4.1.2. Construction of lentiviral NFκB-responsive and SFFV synthetic promoters (controls) for luciferase gene expression

The constitutive SFFV-promoter and the inflammation-inducible 4NFκB-promoter were included in the *in vivo* experiment as positive controls. The 4NFκB promoter contained a 20 bp space between the motifs and a 66 bp space between the proximal NFκB and the TATA box and was cloned into the lentiviral vector using the method described in section 4.1.1. (above). The schematic diagram below illustrates the cloning strategy used to clone the SFFV promoter into the lentiviral plasmid DNA (Fig 4.1.2.).

A. The pUCL-Luc$^+$ plasmid was digested with EcoRI, purified and then incubated with the Klenow enzyme to generate a 5'-blunt end which can ligate to the PmeI blunt end of the pLV.Luc$^+$ vector. The fragment was subsequently digested with BstBI to release the SFFV promoter and 5'-portion of the luciferase gene (730 bp fragment) (Fig 4.1.2 A).
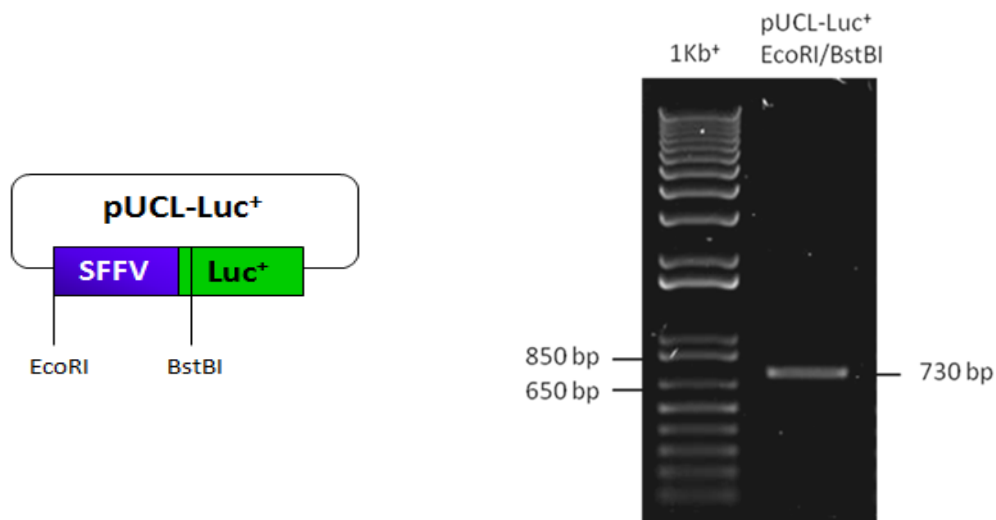
B. The SFFV-Luc$^+$ fragment (730 bp) was cloned into the PmeI/BstBI site within the pLV.Luc$^+$ vector to restore the luciferase gene and generate pLV.SFFV.Luc$^+$ (Fig 4.1.2 B), which was confirmed by DNA sequencing (Appendix 13.11).
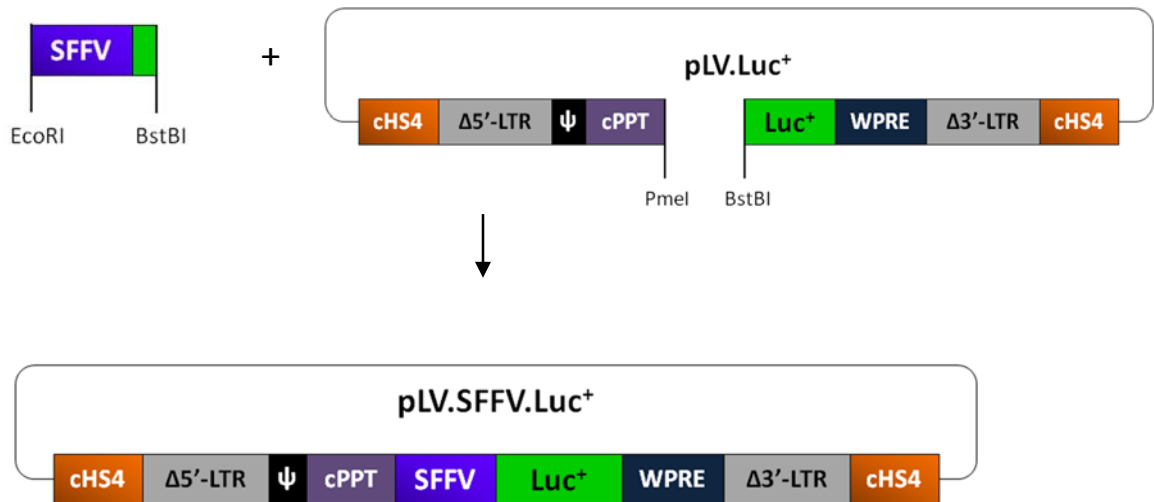
Figure 4.1.2. Schematic diagram illustrating the cloning method used to generate the lentiviral SFFV-promoter positive control construct. The SFFV promoter from the pUCL-Luc$^+$ construct was cloned into the lentiviral vector to generate the LV-SFFV-Luc$^+$ construct expressing the luciferase gene under the control of the constitutive SFFV promoter.

4.2. ***In Vitro* functional analysis of luciferase gene expression induced by lentiviral integrated composite synthetic promoters**

4.2.1. Comparative analysis of the transcriptional activities of transiently transfected plasmid DNA synthetic promoters and lentiviral integrated synthetic promoters

The luciferase expression induced by the eight candidate promoters in plasmid vectors were combined in a single graph for comparative purposes and presented in Figure 4.2. The lentiviruses encoding the expression cassettes were used to transduce 293T cells to generate stable 293T cell lines expressing the luciferase gene under the control of the integrated composite synthetic promoter. An initial screen of the magnitude of luciferase gene expression from the stable 293T cells was performed to identify three composite promoters demonstrating the greatest inflammation-inducibility and synergistic gene expression (Fig 4.3).
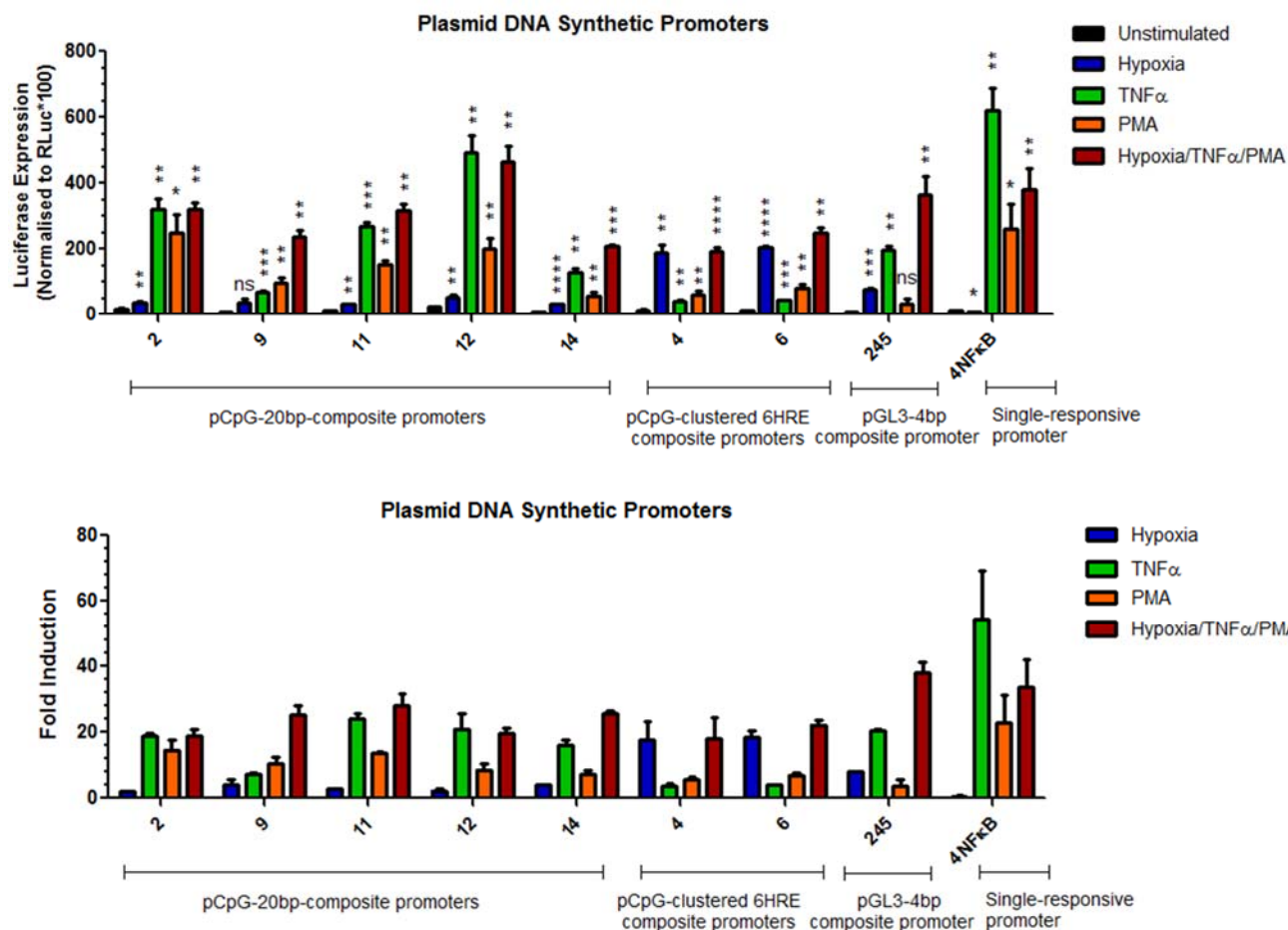
Figure 4.2. Functional analysis of transiently transfected inflammation-inducible synthetic promoters, from various promoter libraries. The luciferase gene expression of the promoters [A] and luciferase fold induction data [B], from separate experiments, were combined in one graph to compare the expression profiles of the candidate synthetic promoters. The data represents the mean ± SD of triplicate values from independent experiments described in Chapter 3. The statistical difference between the unstimulated and stimulated luciferase gene expression was calculated using the Student's t-test (ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$).
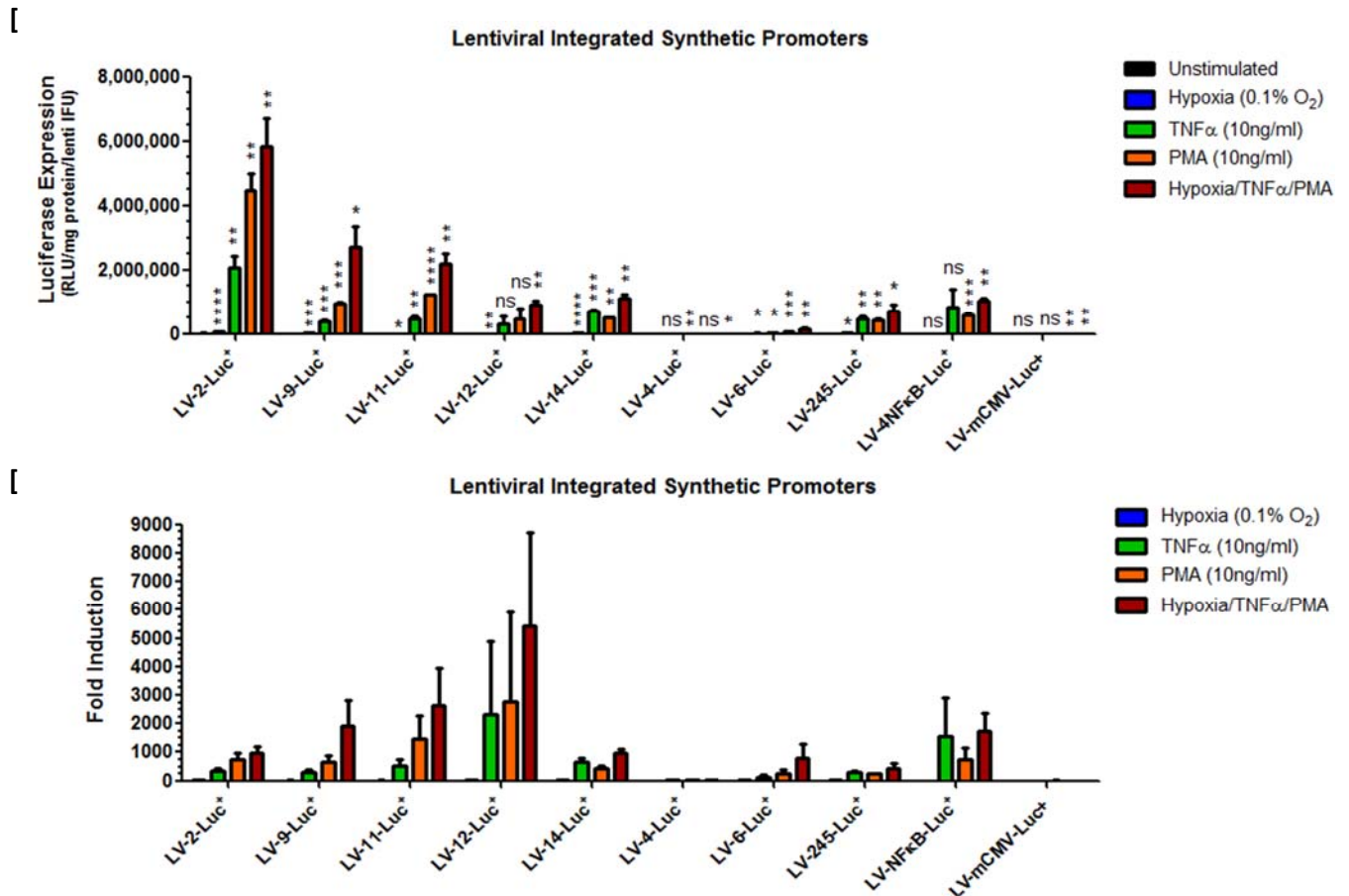
Figure 4.3. Luciferase gene expression induced by lentiviral (LV) composite synthetic promoters. Stable 293T cells expressing the luciferase gene under the control of each of the eight composite promoters were seeded at 20,000 cells per well in a 96 well plate. After 24 hours, the stable 293T cells were unstimulated, incubated in hypoxia (0.1% O_2) or stimulated with TNFα (10 ng/ml), PMA (10 ng/ml) or their combination [A] for 18 hours. Fold inductions were calculated by dividing the induced luciferase values by the uninduced luciferase values [B]. Luciferase expression was normalised to the protein content in the cell lysate and lentiviral titre of the corresponding lentiviral preparation (expressed as RLU/mg protein/lenti IFU). The data represents the mean ± SD of triplicate values, from the same experiment. The statistical difference between the unstimulated and stimulated luciferase gene expression was calculated using the Student's t-test (ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$).

Figure 4.2 graphically presents the combined gene expression profiles of the composite synthetic promoters selected in Chapter 3 (n=8). These promoters displayed differential and multi-responsive luciferase gene expression in response to the individual stimuli and exhibited diverse magnitudes of additive/synergistic gene expression following combined inflammatory and hypoxic stimulation.

Comparisons between the expression profiles of transiently transfected promoter constructs (Fig 4.2) and the lentiviral integrated promoters (Fig 4.3) demonstrated that the stable integration of the composite promoters significantly improved the magnitude of induced luciferase gene expression and exhibited low basal luciferase gene expression. Also, the vast majority of the composite promoters displayed greater additive/synergistic gene expression in the stable 293T cells than in plasmid transfected cells.

Interestingly, the lentiviral promoter LV-2 displayed synergistic luciferase gene expression in response to combined inflammatory and hypoxic stimulation, with high fold induction of 988 fold (Fig 4.3 B). However, the basal expression of promoter LV-2 (6193 RLU/mg protein/lenti IFU) was higher than the basal gene expression levels displayed by the other promoters. Consequently, the overall fold induction displayed by promoter LV-2 was lower than many of the other promoters, e.g. promoter LV-12 displayed an extremely low basal luciferase expression (232 RLU/mg protein/lenti IFU) and so that the induced expression resulted in a maximal fold induction of 5450 fold following combined inflammatory and hypoxic stimulation.

In contrast, the luciferase gene expression induced by the lentiviral promoter LV-4 was the lowest compared to the other promoters (Fig 4.3), which was not observed following the transient transfection of the plasmid DNA equivalent (Fig 4.2). Lentiviral integration of promoter LV-4 resulted in a modest 27.5 fold induction in response to combined inflammatory and hypoxic stimulation (* p= 0.04, increase in gene expression), which was considerably lower than the high fold inductions displayed by the other promoters. Nevertheless, on the basis of the expression profile demonstrated with the transiently transfected promoter

construct (Fig 4.2A), the plasmid DNA promoter 4 displayed greater gene expression in response to hypoxia than to inflammatory stimulation, which is not ideal for RA gene therapy, as inflammation is the dominant pathological characteristic.

The single-responsive promoter 4NFκB in both the plasmid DNA (Fig 4.2) and lentiviral construct (Fig 4.3) was responsive to inflammatory stimulation but exhibited reduced or unchanged luciferase gene expression following combined inflammatory and hypoxic stimulation, respectively. These observations further support our strategy of utilising composite promoters to achieve multi-responsive and synergistic gene expression, compared to the use of single-responsive synthetic promoters.

Interestingly, there were slight variations in the expression profiles of the synthetic promoters in response to individual stimuli. For example, plasmid DNA promoters 2, -11 and -12 were more responsive to TNFα than to PMA (Fig 4.2). However, the lentiviral promoters LV-2, LV-11 and LV-12 induced greater gene expression to PMA than to TNFα (Fig 4.3). Despite these discrepancies, these promoters retained their ability to respond highly to various inflammatory stimuli and induce synergistic gene expression in response to combined stimulation, which was a fundamental requirement of the composite synthetic promoters.

In general, the fold change of luciferase gene expression induced from lentiviral integrated composite synthetic promoters greatly exceeded that demonstrated by transiently transfected promoter constructs. The composite promoters generally displayed similar expression profiles and some had improved synergistic gene expression e.g. LV-2, LV-9 and LV-12. Promoters LV-2 and LV-9 exhibited significant increases in gene expression in response to all treatments and promoter LV-12 displayed robust fold inductions following individual and combined stimulation. Therefore, these three composite promoters were selected for further detailed analysis of their induction profile.

4.2.2.   Dose-response and time-course kinetic assays of luciferase gene expression of lentiviral integrated composite synthetic promoters

293T cells, which stably expressed the luciferase gene under the control of the composite promoters, facilitated further functional characterisation in time-course and dose-response assays. These experiments provided knowledge on the sensitivity, and the rate and duration of promoter activation, to facilitate the selection of the optimal composite promoter for regulating gene expression *in vivo.*

4.2.2.1.   Lentiviral composite synthetic promoters are highly sensitive and display dose-dependent induction

The changes in the inflammatory status of the RA joint promote the different relapse and remission phases of RA. Therefore, it is imperative that the composite promoter for RA gene therapy responds accordingly to the magnitude of disease activity within the RA joint: the candidate synthetic promoters require a high degree of sensitivity to the inflammatory environment in order to induce the appropriate level of therapeutic gene expression and subsequently reduce joint inflammation.

Dose-response experiments were conducted to evaluate the sensitivity of lentiviral composite promoters to TNFα or PMA stimulation over the concentration range 0.001 – 100 ng/ml (Figure 4.4) and the fold inductions are presented in Figure 4.5.
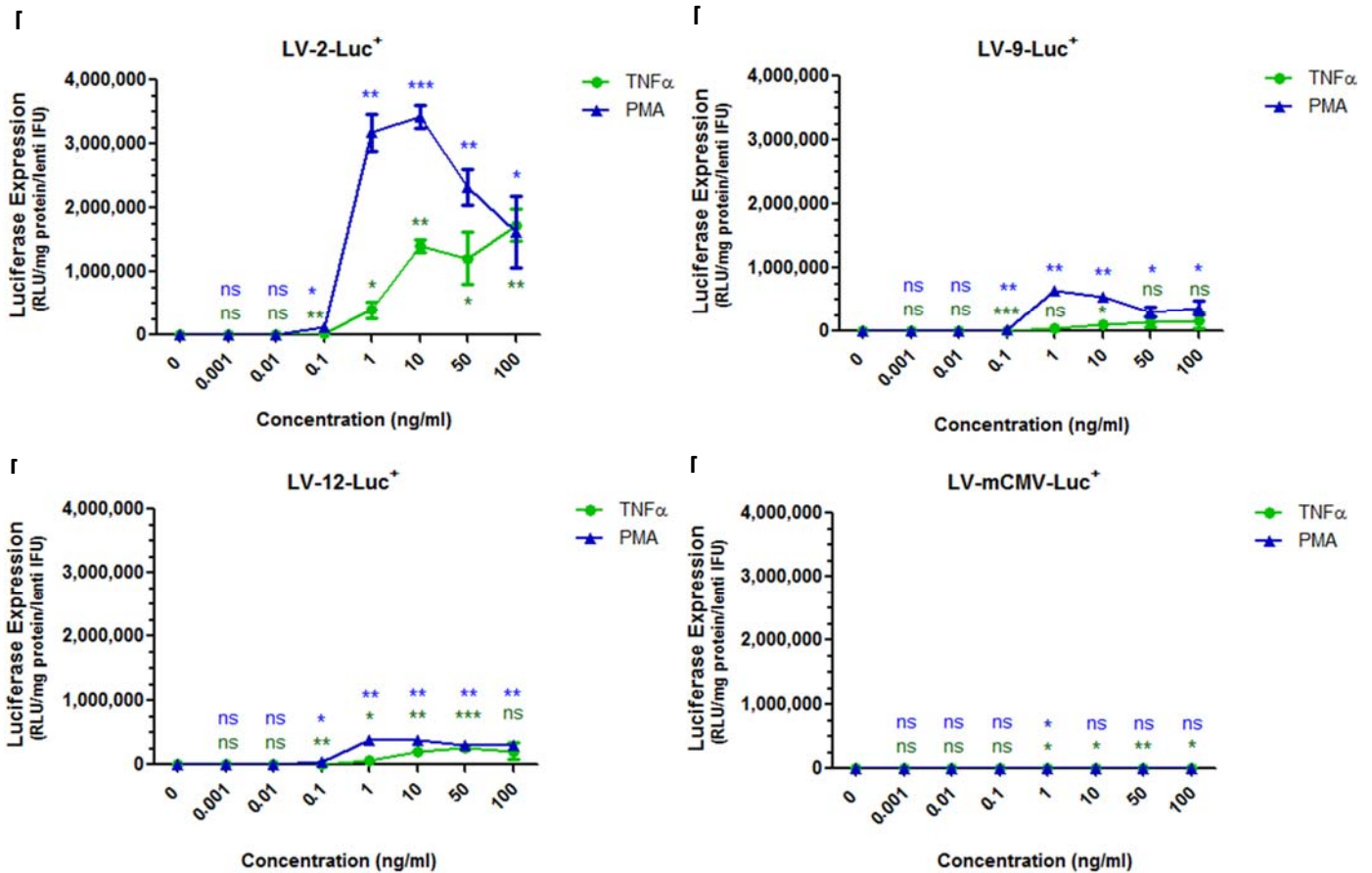
Figure 4.4. Responsiveness of lentiviral composite promoters to inflammatory mediator stimulation. Stable 293T cells expressing luciferase under the control of the synthetic promoters LV-2 [A], LV-9 [B], LV-12 [C] and LV-mCMV [D] were seeded at 20,000 cells per well in a 96-well plate in triplicate wells. After 24 hours, the cells were unstimulated (0 ng/ml) or treated with 0.001 ng/ml –100 ng/ml of TNFα and PMA to generate a dose-response curve. The cells were harvested 18 hours after stimulation and the luciferase expression was normalised to the protein content in the cell lysate and the lentiviral titre of the corresponding lentiviral preparation (expressed as RLU/mg protein/ lenti IFU). The data represents the mean ± SD of triplicate values. The statistical significance compared to the unstimulated luciferase gene expression (at 0 ng/ml) was calculated using the Student's t-test (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p≤0.0001, where green and blue asterisks correspond to TNFα and PMA datasets, respectively).

250
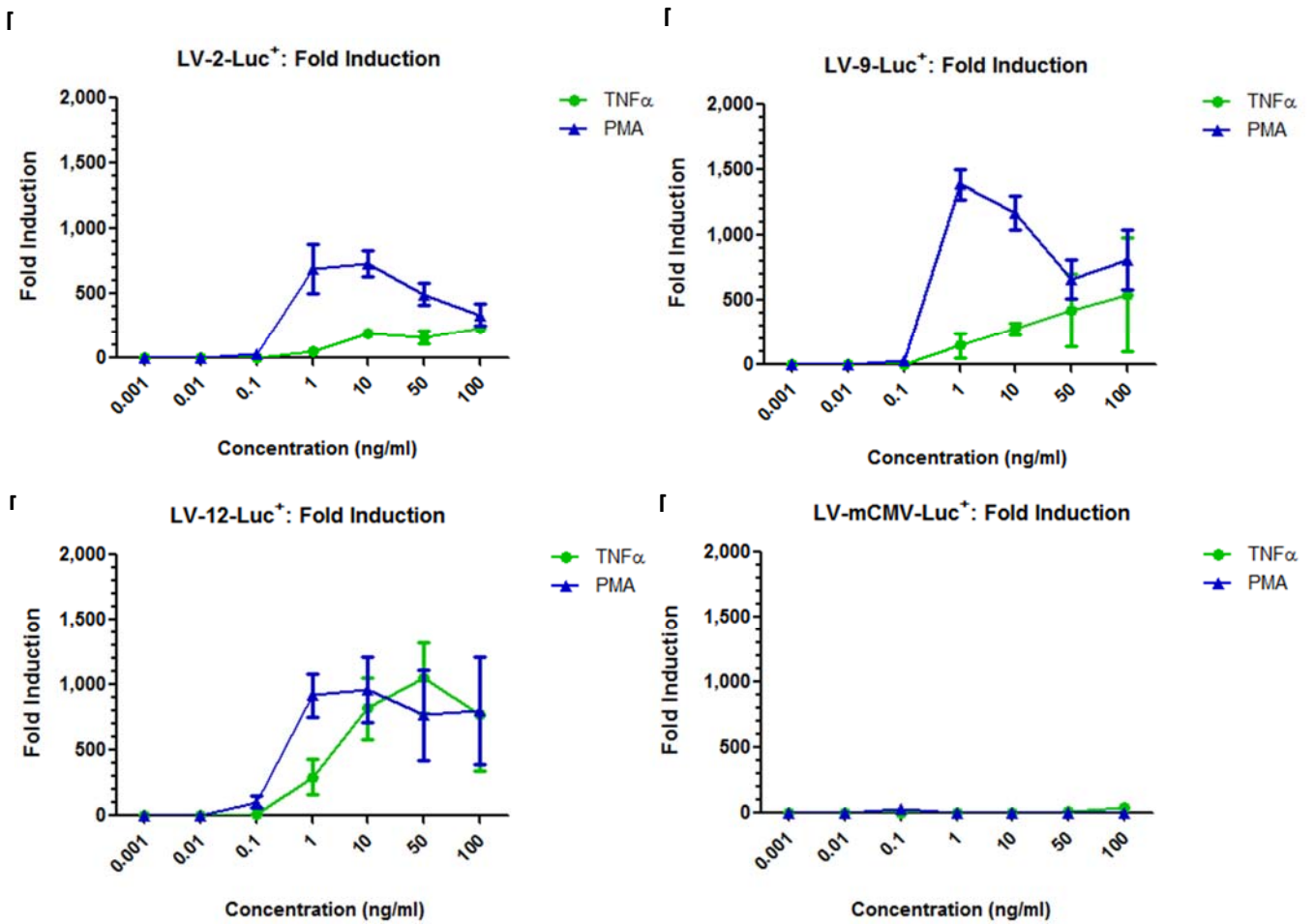
Figure 4.5. Fold induction of lentiviral composite promoters in response to inflammatory mediator stimulation. The changes from basal luciferase gene expression of the composite promoters LV-2 [A], LV-9 [B], LV-12 [C] and LV-mCMV [D] were calculated using the data in Figure 4.4. The data represents the mean ± SD of triplicate values.

Encouragingly, the composite promoters displayed dose-dependent increases of TNFα-induced luciferase gene expression and to an extent, with PMA-induction (Fig 4.4 A-C and Fig 4.5 A-C). In contrast, the negative control promoter LV-mCMV (Fig 4.4 D and 4.5 D) was generally unresponsive to inflammatory stimulation, although marginal increases in TNFα-induced gene expression was observed.

The composite promoters were highly sensitive to low concentrations of TNFα and PMA and induced significant increases in luciferase gene expression ($p \leq 0.05$ to $p \leq 0.001$) in response to 0.1 ng/ml of TNFα or PMA stimulation. Generally, the induced luciferase gene expression peaked in response to 1-10 ng/ml of TNFα or PMA stimulation. However, noticeable differences can be observed in the induced luciferase expression profiles of the composite promoters e.g. the luciferase gene expression of promoter LV-2 peaked in response to 10 ng/ml PMA (*** $p=0.001$) but gene expression significantly decreased with higher PMA concentrations (50-100 ng/ml). In contrast, LV-2-induced-gene expression increased with increasing TNFα concentration and the luciferase gene expression induced by promoters LV-9 and LV-12 also displayed similar trends with both TNFα and PMA stimulation.

4.2.2.2. Lentiviral composite synthetic promoters are rapidly activated and display time-dependent increases in luciferase gene expression

Time-course kinetics of luciferase gene expression induced by the composite promoters was evaluated by stimulating stable 293T cell lines with 10 ng/ml of TNFα or PMA or their combination and monitoring luciferase gene expression for 2, 8, 12, 24 or 48 hours (Fig 4.6). The fold inductions are presented in Figure 4.7.
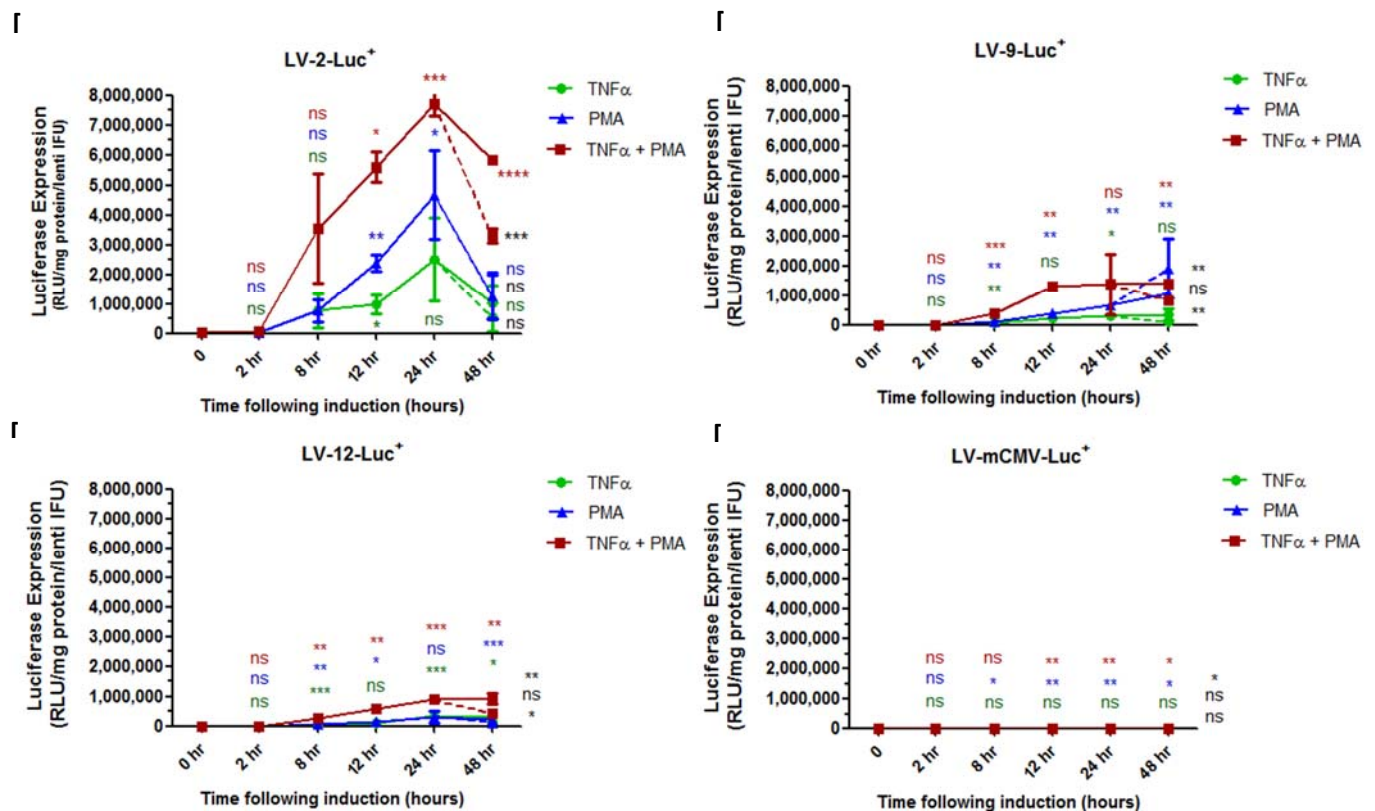
Figure 4.6. Lentiviral composite promoters are rapidly activated and display peak activation 24 hours after inflammatory mediator stimulation. Stable 293T cells expressing luciferase under the control of the synthetic promoters LV-2 [A], LV-9 [B], LV-12 [C] and LV-mCMV [D] were seeded at 20,000 cells per well in a 96-well plate in triplicate wells. After 24 hours, the cells were either unstimulated (0 hours) or stimulated with TNFα (10 ng/ml), PMA (10 ng/ml) or a combination of TNFα and PMA for 2, 8, 12, 24 or 48 hours. At 24 hours, the inflammatory stimulus was removed from the selected wells (indicated by the dotted line) and replaced with low serum medium after which the luciferase expression was monitored for a further 24 hours (48 hour time-point). Cell lysates were harvested at the stated time points and the luciferase expression was quantified and normalised to the protein content in the cell lysate and the lentiviral titre of the corresponding lentiviral preparation (expressed as
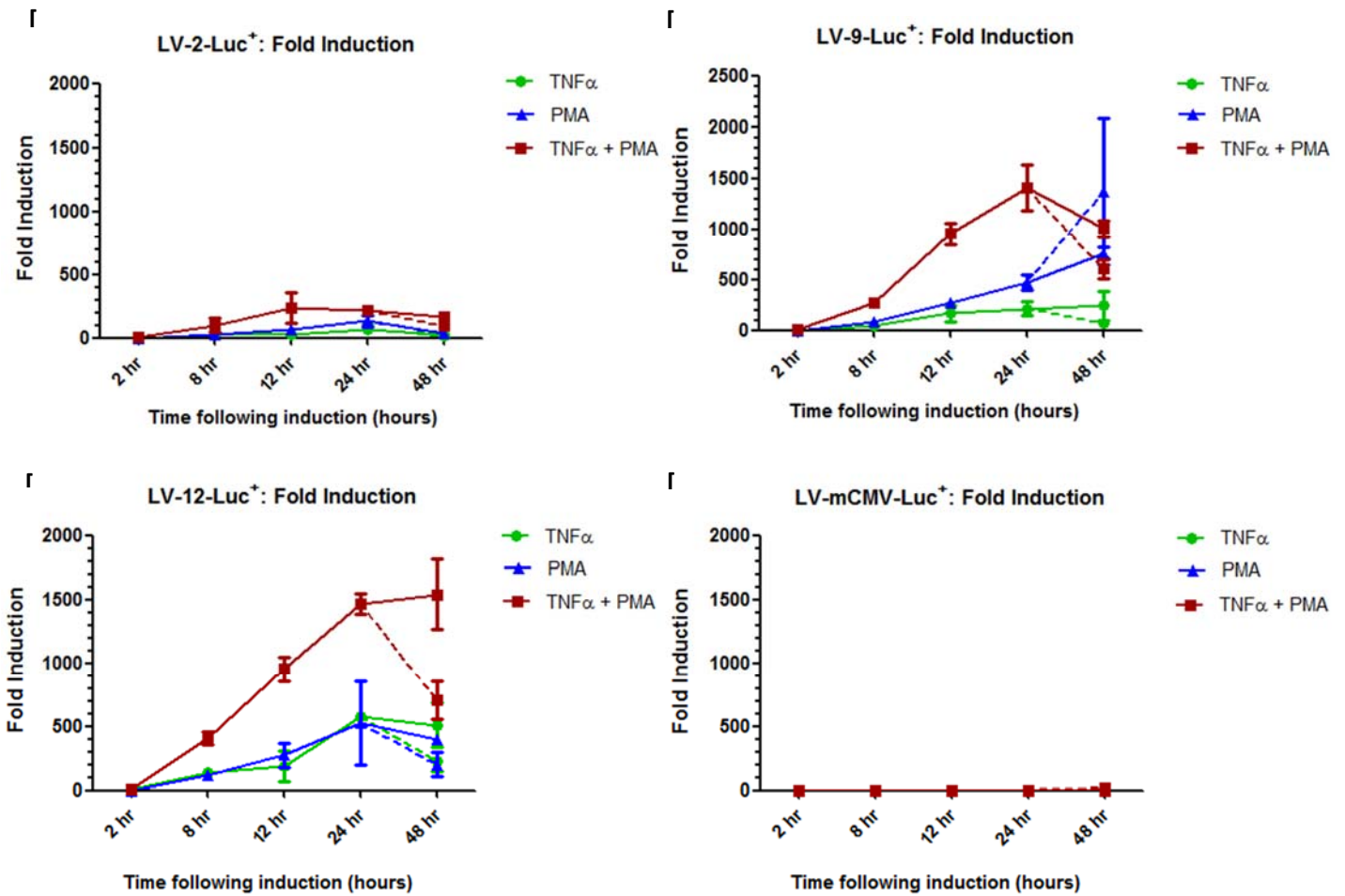
Figure 4.7. Kinetics of lentiviral composite promoters expressed as fold inductions. The changes from the basal luciferase gene expression induced in the composite promoters LV-2 [A], LV-9 [B], LV-12 [C] and LV-mCMV [D] from the data in Figure 4.6 was calculated as fold inductions. The data represents the mean ± SD of triplicate values.

Figures 4.6 and 4.7 confirmed the rapid activation of composite promoters, which generally peaked after 12 hours of exposure to inflammatory stimulation ($p \leq 0.05$ to $p \leq 0.001$). Interestingly, LV-9 and LV-12 promoter activation was steadily maintained from 24-48 hours in contrast to the luciferase expression induced by promoter LV-2 which sharply decreased beyond 24 hours. The composite synthetic promoters displayed significant synergistic activation, at most time-points, in response to combined inflammatory stimulation. It is possible that the variation of values within specific datasets may have been the cause for the low or absent statistical significance in datasets which clearly had increased gene expression following stimulation (Fig 4.6 A-C and Fig 4.7 A-C).

To mimic the fluctuations of inflammation within the RA joint, the inflammatory stimulus was removed from cells and replaced with low serum medium (indicated by the dotted line, statistical significance indicated by black asterisks). Removal of the inflammatory stimulation at 24 hours generally resulted in decreased luciferase gene expression in comparison to relatively sustained luciferase gene expression in response to continued inflammatory stimulation which corroborates the high sensitivity and rapid responsiveness of the composite synthetic promoters to the presence, removal and various concentrations and durations of exposure to inflammatory stimulation. Also, similarities between the induced luciferase expressions at the 24 hour time point in Figure 4.6 to the results obtained from the same stable 293T cells presented in section 4.2.1 (Fig 4.3) shows that there was no compromise in promoter activity after four weeks in cell culture.

The dose-response and time-course kinetics of luciferase gene expression confirmed the appropriate responsiveness, high-sensitivity and rapid activation induced by the composite promoters LV-2, LV-9 and LV-12. These studies provide encouraging evidence to progress to monitoring promoter activation during paw inflammation.

4.3. *In vivo* functional analysis of luciferase gene expression from lentiviral integrated synthetic promoters

The activation profiles of the lentiviral composite promoters were assessed in a carrageenan-induced paw inflammation mouse model. The lentiviruses expressing the luciferase gene under the control of the composite promoters LV-2, LV-9, and LV-12, the inflammation-inducible promoter LV-4NFκB, the positive control promoter LV-SFFV, and the negative control promoter LV-mCMV, were prepared at 260,000 lentiviral IFU in 25 µl. Male adult CD1 mice (n=3 per group) received an intraplantar injection of 25 µl of the candidate lentiviral particles into both hind paws 7 days before carrageenan administration to allow sufficient time for integration of the vector genome. After 7 days, paw inflammation was induced by an intraplantar injection of 50 µl of 1% λ-carrageenan solution into the left hind paw. The right hind paws received 50 µl of sterile saline intraplantarly, serving as the control.

Quantification of *in vivo* luciferase gene expression was monitored using an IVIS bioluminescent imaging system, which measured real-time bioluminescence in photons, following an intraperitoneal injection (200 µl) of the luciferin substrate. Imaging was performed before inflammation (0 hours) and 3, 24 and 72 hours post-carrageenan injection at a binning of 8 for 5 minutes. The light emission in a defined region of interest (ROI), around each hind paw, allowed the luciferase expression to be monitored for consecutive images at each time point (Fig 4.8 A-G). The footpad thickness was measured before inflammation (0 hours) and 3, 24, 48, 72 and 96 hours post-carrageenan injection, using a dial caliper (Fig 4.8 H).

[A]

LV-mCMV-Luc⁺

| 0 hours
pre-inflammation | 3 hours
post-inflammation | 24 hours
post-inflammation | 72 hours
post-inflammation |

[B]

LV-SFFV-Luc⁺

| 0 hours
pre-inflammation | 3 hours
post-inflammation | 24 hours
post-inflammation | 72 hours
post-inflammation |

[C] LV-4NFκB-Luc⁺

0 hours
pre-inflammation

3 hours
post-inflammation

24 hours
post-inflammation

72 hours
post-inflammation

[D] LV-2-Luc⁺

0 hours
pre-inflammation

3 hours
post-inflammation

24 hours
post-inflammation

72 hours
post-inflammation

[E] LV-9-Luc⁺

0 hours
pre-inflammation

3 hours
post-inflammation

24 hours
post-inflammation

72 hours
post-inflammation

[F] LV-12-Luc⁺

0 hours
pre-inflammation

3 hours
post-inflammation

24 hours
post-inflammation

72 hours
post-inflammation

In Vivo Synthetic Promoter Activity

Figure 4.8. Composite synthetic promoters exhibit inflammation-specific luciferase gene expression in carrageenan-induced inflamed mouse paws. Representative images are shown of CD1 mouse hind paws 7 days post-injection with lentiviral particles LV-mCMV-Luc[+] [A], LV-SFFV-Luc[+] [B], LV-4NFκB-Luc[+] [C], LV-2-Luc[+] [D], LV-9-Luc[+] [E], LV-12-Luc[+] [F]. The left hind paws received an intraplantar injection of 50 µl of 1% λ-carrageenan whereas the control right paws were injected with 50 µl of saline. The luciferase expression was monitored before inflammation (0 hours) and 3, 24 and 72 hours post-inflammation. Light emission (photons/second/cm^2/sr) was measured using an IVIS CCCD camera at a binning of 8 for 5 minutes. The quantitative bioluminescence data within the defined region of interest over individual paws represents the mean ± SD of n= 3 mice, per group [G]. The statistical significance compared to baseline gene expression was calculated using the Student t-test (ns = p>0.05, * = p≤0.05, ** = p≤0.01) and only statistically significant points are labelled. Paw thickness was measured before inflammation (0 hours) and 3, 24, 48, 72 and 96 hours post-inflammation, using a dial caliper where each time-point represents the mean ± SD of n= 3 mice, per group [H].

As anticipated, at 0 hours (before inflammation) there was no detectable luciferase expression induced by the negative control promoter LV-mCMV. In contrast, the constitutively active LV-SFFV positive control promoter induced constitutive luciferase gene expression whereas inflammation-inducible promoters LV-4NFκB, LV-2, LV-9 and LV-12 induced low basal luciferase expression (Figure 4.8 A-F).

Encouragingly, at 3 hours post-carrageenan injection, the inflammation-inducible LV-2, LV-9 and LV-4NFκB promoters induced robust inflammation-specific luciferase expression which was comparable to that of the strong viral LV-SFFV promoter (Fig 4.8 G), however only the induction by LV-2 was statistically significant (* p=0.03). In parallel to the *in vitro* expression profile of promoter LV-12, this promoter induced low, albeit inflammation-specific, luciferase gene expression at 3 hours post-inflammation (Fig 4.8 A-F).

The caliper measurements of the saline control right paws, which were also injected with lentivirus, provided consistently low baseline measurements of paw thickness throughout the 96 hour monitoring period, therefore, the changes in luciferase expression in the treated left paw can be attributed to the level of inflammation within the paw. The peak-intensity of carrageenan-induced paw inflammation was observed at 24-48 hours post-inflammation (Fig 4.8 H). However, at 24 hours post-inflammation, there was a slight decrease in luciferase expression from the LV-4NFκB, LV-2 and LV-9 treated left hind paws (Fig 4.8 C-E). It is feasible that the substantial paw thickness potentially impaired photon emission which consequently, resulted in decreased quantified luciferase gene expression. Nonetheless, the gene expression levels at 24 hours post-inflammation were still comparable to the luciferase expression induced by the constitutively active LV-SFFV promoter, at the same time-point (Fig 4.8 G). At 72 hours, paw thickness was generally lower than at the previous time points but had not returned to baseline (Fig 4.8 H) and the luciferase expression from the induced promoters had also declined (Fig 4.8 G).

4.4. Discussion

The requirements of successful gene therapy approaches include efficient gene delivery, sustained expression of the transgene, limited biodistribution and non-immunogenicity, which are characteristics governed by the mode of delivery and gene-delivery vector. Many gene therapy vectors are limited by their storage capacity, duration of transgene expression and immunogenic profile, which restrict their utility. Non-viral vectors i.e. plasmid DNA commonly yield low gene transfer efficiency because they persist episomally in the nucleus and lack the inherent ability to enter cells and localise to the nucleus. In contrast, lentiviral vectors (LV) offer the advantage of stable genomic integration of the transgene into dividing or non-dividing cells to provide long-term gene expression with low-immunogenicity (Gould and Favorov, 2003).

Earlier concerns regarding the safety profile of lentiviral-mediated gene delivery have been overcome by extensively modifying the structure of the lentiviral vector. A major step towards clinical acceptability has been the development of self-inactivating (SIN) lentiviral vectors with a 400-nucleotide deletion in the 3'-long terminal repeat (LTR). These deletions eradicate the TATA box and the LTR promoter to decrease the emergence of replication-competent viruses without affecting the viral titres or transgene expression, thereby improving the biosafety profile of lentiviral vectors (Zufferey *et al.,* 1998). Chapter 4 reports the *in vitro* and *in vivo* translational studies of luciferase gene expression induced by the candidate composite synthetic promoters within the SIN lentiviral vector LV-Luc$^+$.

Comparisons between the luciferase gene expression of the composite promoters in plasmid DNA and lentiviral vectors (section 4.2.1) demonstrated the dramatic improvements in the magnitude of luciferase expression following lentiviral transduction of 293T cells. The overall performance and fold inductions of the lentiviral integrated promoters greatly surpassed that displayed by analogous plasmid DNA promoters. Encouragingly, many of the composite promoters acquired improved synergistic expression profiles whilst other promoters retained

their synergistic gene expression and overall, the lentiviral promoters displayed better expression profiles than their plasmid DNA equivalents. This initial screening of lentiviral-mediated luciferase expression facilitated the selection of the composite promoters LV-2, LV-9 and LV-12 which generally exhibited low basal and high induced luciferase expression, multi-responsiveness and robust synergistic inducibility following combined stimulation.

In agreement with our concept of combining inflammation-inducible regulatory elements for enhanced gene expression, the IL-1/IL-6 hybrid synthetic promoter developed by van de Loo *et al.,* (2004) exhibited additive luciferase gene expression in response to combined inflammatory stimulation. However, the fold inductions observed following stimulation with combined LPS, PMA and *dbc*AMP in HIG82 cells (5.3 fold) and RAW 264.1 cells (18 fold) following transient transfection of plasmid DNA and combined TNFα, IL-1β and IL-18 in an adenoviral infected murine chondrocyte cell line (upto 4 fold) were substantially lower than the fold changes induced by the plasmid and lentiviral composite synthetic promoters described in this chapter, which strengthens the notion of constructing composite promoters for synergistic transcriptional activation.

The stable and long-term luciferase expression conferred by lentiviral genomic integration enabled the kinetics of luciferase expression induced by the composite promoters LV-2,   LV-9 and LV-12 to be monitored in stable 293T cells (section 4.2.2). These experiments provided evidence of dose- and time-dependent increases in luciferase gene expression which confirmed the high-sensitivity (Fig 4.4) and rapid activation (Fig 4.6) of the composite promoters in response to various concentrations and durations of exposure to inflammatory stimulation, which further supported the choice of the promoters LV-2, LV-9 and LV-12 as candidates for local RA gene therapy. Also, the high induced luciferase expression in response to low concentrations of inflammatory stimuli (Fig 4.4) suggest that high therapeutic gene expression can potentially be induced by the composite promoters in conditions of relatively low inflammation, which may be beneficial for the early stages of RA disease relapse. In these

situations, the composite promoter can function prophylactically to reduce the level of inflammation before the onset of chronic inflammation in the RA joint.

The limited gene expression variability between the plasmid DNA and lentiviral vector constructs and the retention of favourable gene expression profiles displayed by the lentiviral composite synthetic promoters can be attributed to the performance-enhancing modifications of the LV-Luc$^+$ vector. For example, the similar expression profiles of the lentiviral integrated synthetic promoters to their plasmid DNA equivalents may have been due to the presence of the chromatin insulator from the 5' end of the chicken β-globin locus (5'-cHS4) which flanks the composite promoter and luciferase gene within the LV-Luc$^+$ vector. The HIV-1 lentiviral vectors preferentially integrate into transcriptionally active areas in the host cell genome (Schroder *et al.,* 2002) and therefore are subject to influences by external factors. The presence of the 5'-cHS4 insulators shield the synthetic promoter from the action of external distal enhancers (enhancer blocking effect), which is likely to have prevented high basal gene expression from the lentivirally-integrated synthetic promoters. This enhancer-blocking activity is of great importance for gene therapy applications as it might prevent or reduce insertional gene activation of growth-regulatory genes or proto-oncogenes located near the vector insertion site. Also, the insulators protect the luciferase gene against the advancement of adjacent inactive condensed chromatin to confer protection against gene silencing (barrier effect) (Pikaart *et al.,* 1998).

Also, in contrast to the pCpG-Luc$^+$ plasmid vector, the LV-Luc$^+$ contains the woodchuck post-transcriptional regulatory element (WPRE) in the 3' untranslated region (UTR) of the luciferase gene, which has been demonstrated to substantially increase gene expression levels by functioning in the nucleus to stimulate gene expression post-transcriptionally and to increase the levels of nuclear transcripts which consequently increases protein expression (Zufferey *et al.,* 1999). Therefore, the greatly enhanced luciferase gene expression induced by lentiviral composite promoters can be partly attributed to increased luciferase protein levels facilitated by the incorporation of the WPRE in the lentiviral vector.

Intra-articular injections of lentiviral vectors typically result in the transduction of tendons, ligaments, cartilage, muscle, adipose/areolar synovium and synoviocytes within the joint (Gouze *et al.,* 2007). Lentiviral vectors pseudotyped with the VSV-G envelope have increased host-cell range due to the ability of the VSV-G envelope binding to phospholipid components of the cell membrane to mediate viral entry by membrane fusion (Burns *et al.,* 1993) and therefore enable lentiviral-mediated delivery of the composite synthetic promoters into various cell types. Consequently, the VSV-G pseudotyped lentiviral particles were used to transduce cells *in vivo*, to monitor promoter activity during paw inflammation.

Inflammation is the underlying pathology in RA which perpetuates the structural damage to the joint and many animal models of paw inflammation provide an accurate representation of the immunological and pathological events similar to those observed in human RA (Bendele, 2001). In this chapter, carrageenan-induced paw oedema was used as an experimental model of mouse paw inflammation to examine the *in vivo* profile of luciferase gene expression induced by the inflammation-inducible composite synthetic promoters; LV-2, LV-9 and LV-12 (section 4.2). The induction of paw inflammation by carrageenan in mouse paws was previously demonstrated by Levy *et al.,* (1969), and similarly shown by Winter *et al.,* (1962) in rats. The application of the carrageenan-induced paw oedema mouse model for the initial translation studies, described in section 4.3 offered several advantages over other experimental models of inflammation.

Firstly, in the routinely used CIA mouse model, paw inflammation can occur in any of the paws after several days post-immunisation. In contrast, a single local intraplantar injection of 1% λ-carrageenan into the mouse hind paw ensures the rapid induction of inflammation within hours which is confined to the paw injected with carrageenan, as described by Morris (2003). Therefore, this limits the inefficient use of valuable lentiviral composite promoter particles. Secondly, the carrageenan-induced paw oedema model has been shown to exhibit a biphasic profile where the inflammatory reaction first peaks at 4 hours and then at 72 hours following a local carrageenan injection and decreases thereafter (Henriques *et al.,* 1987). As

demonstrated in the preceding time-course kinetic experiments (section 4.2.2), the composite promoters induced time-dependent increases in luciferase expression (upto 48 hours and potentially longer) which coincides with the time-course of carrageenan-induced inflammation and represents a quick initial screen of promoter activity *in vivo.* Thirdly, the inflammatory process induced by carrageenan is characterised by the production of prostaglandins and leukotrienes in addition to the release of other mediators by infiltrated immune cells, particularly neutrophils (Cuzzocrea *et al.,* 1999). These mediators can directly or indirectly activate NFκB, AP-1 and possibly HIF-1α to induce additive/synergistic luciferase gene expression.

The *in vivo* luciferase imaging results (section 4.3) confirmed the disease-specific induction of luciferase gene expression from the lentiviral composite synthetic promoters LV-2, LV-9 and LV-12 in the carrageenan-induced paw inflammation model, although only LV-2 induction at 3 hours post-inflammation was statistically significant (* p=0.03). Nevertheless, mice injected with the inflammation-inducible LV-2, LV-9 and LV-4NFκB constructs displayed robust luciferase gene expression which was localised to the inflamed left hind paw and comparable to the luciferase expression induced by the strong viral promoter LV-SFFV, as confirmed by real-time bioluminescence imaging (Fig 4.8 A-G).

Interestingly, the luciferase gene expression induced by the LV-SFFV promoter in the inflamed left paw continuously and significantly declined after 3 hours post-inflammation, whilst the luciferase expression from the control saline paw remained unchanged throughout the 72 hour monitoring period (Fig 4.8 G). In contrast, Garaulet *et al.,* (2013) reported marginal fluctuations in the fold inductions of luciferase gene expression induced by the lentiviral-integrated SFFV promoter for 30 days in a zymosan-induced paw inflammation model. However, the authors did not disclose the specific luciferase gene expression values in the control and zymosan-induced paws, since the data was represented as fold induction over unstimulated conditions, making it difficult to identify variations in the SFFV promoter activity over the course of inflammation. The decrease in LV-SFFV promoter activity described in this chapter (Fig 4.8

G) might indicate SFFV promoter silencing, however, definite confirmation e.g. using DNA methylation analysis, highlights an option for further work.

The luciferase gene expression in the inflamed paws peaked at 3 hours post-inflammation (Fig 4.8 A-G) which corresponds to the reported trend of carrageenan induced-paw inflammation described by Henriques *et al.,* (1987) but was contradicted by the dial caliper measurements of paw swelling (Fig 4.8 H) which showed the peak inflammation at 24-48 hours post-carrageenan injection. It is feasible that the significant degree of paw swelling at 24 hours may have obstructed the detection of luciferase expression by the IVIS machine, which consequently resulted in lower quantified luciferase gene expression. Interestingly, the second peak of inflammation at 72 hours reported by Henriques *et al.,* (1987) was not observed in this experiment. Nonetheless, the lowest luciferase expression was observed at 72 hours post-inflammation in mice injected with the composite synthetic promoters (Fig 4.8 G), which was supported by the reduced paw swelling at this time point (Fig 4.8 H).

Interestingly, the functional analysis data obtained from plasmid DNA transient transfections (section 4.2.1), lentivirus integrated 293T cells (section 4.2.1) and lentivirus infected mouse paws (section 4.3) showed that promoter LV-2 consistently displayed a low/moderate basal and very high induced luciferase gene expressions whereas promoter LV-9 induced a lower basal and a high induced luciferase gene expression and promoter LV-12 generally exhibited low basal and low induced luciferase gene expression. Therefore, it is feasible to expect similar expression profiles to be displayed by the composite promoters when regulating therapeutic gene expression during paw inflammation.

Overall, the initial *in vivo* experiment described in this discussion confirmed the development of disease-regulated composite synthetic promoters which can be applicable for RA gene therapy. Using the luciferase reporter gene as a surrogate for a therapeutic gene, the lentiviral composite synthetic promoters have demonstrated versatility by displaying robust inflammation-inducible profiles in human kidney 293T cells (*in vitro*) and mouse synovial cells

(*in vivo*), which can be attributed to the ubiquitous activation of NFκB, AP-1 and HIF-1α in numerous cell types and broad cell tropism of the VSV-G envelope in the lentiviral vector. Collectively, the experiments described in Chapters 3 and 4 have enabled the selection of LV-2 and LV-9 as the most promising promoters to regulate local therapeutic gene expression to potentially reduce joint inflammation *in vivo*.

CHAPTER 5:

Evaluating the Therapeutic Efficacy of Soluble

mTNFRII-Fc and hIL-1Ra Expressed from

Inflammation-Inducible Composite Promoters

5.1. Introduction

There is substantial evidence implicating TNFα and IL-1β in the pathological processes in a wide range of chronic inflammatory diseases, particularly in RA (Feldmann and Maini, 2008). Selective inhibition of TNFα and IL-1β activity through the use of soluble TNF-receptors (TNFR-Fc) and IL-1 receptor antagonist (IL-1Ra), respectively, has demonstrated reduced inflammation and suppression of disease in experimental models of paw inflammation. In this chapter, the regulated expression of mouse TNFRII-Fc (mTNFRII-Fc) and human IL-1Ra (hIL-1Ra) therapeutic proteins using the inflammation-inducible composite promoters, LV-2 and LV-9 was quantified using ELISA and the therapeutic potential of the candidate promoter was assessed in the carrageenan-induced paw inflammation model.

5.1.1. Construction of lentiviral composite synthetic promoters regulating mTNFRII-Fc and hIL-1Ra therapeutic gene expression

The schematic diagram below illustrates the cloning strategy used to clone the therapeutic genes downstream of the composite synthetic promoters in the lentiviral plasmid DNA (Fig 5.1).

A. pLV.CMV.eGFP was digested with BamHI and SalI to release the eGFP gene. The resulting 8357 bp fragment served as the cloning vector pLV.CMV (Fig 5.1 A).
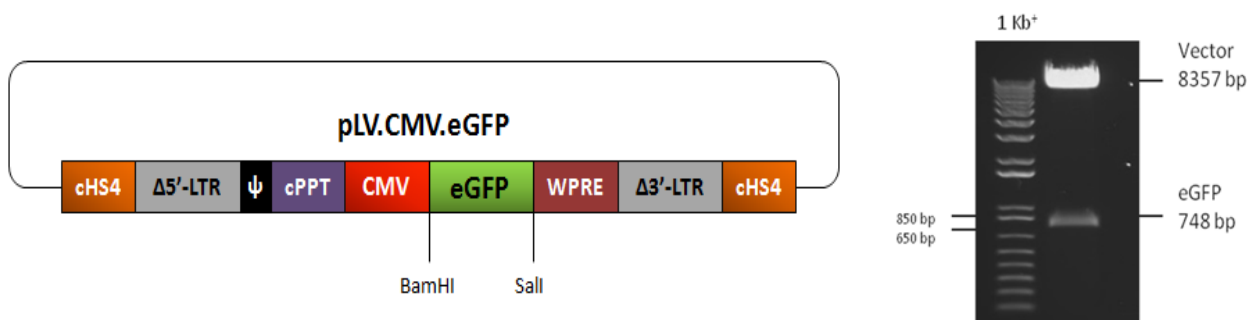
B. The hIL-1Ra gene (540 bp) was amplified from pcIL-1Ra using forward BamHI and reverse SalI primers (Fig 5.1 B).
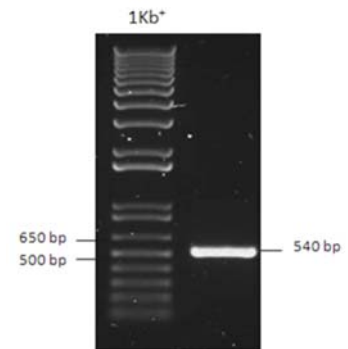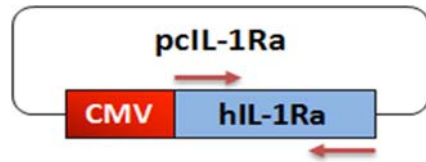
C. The hIL-1Ra PCR product was digested with BamHI and SalI and cloned into the equivalent site within the pLV.CMV vector to generate pLV.CMV.hIL-1Ra. The resulting construct was digested with PmeI and BamHI to release the CMV promoter. The 6931 bp fragment served as the cloning vector pLV.hIL-1Ra (Fig 5.1 C).
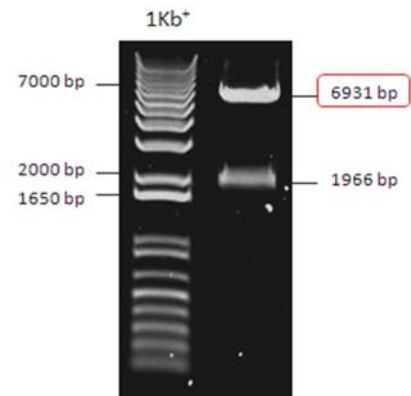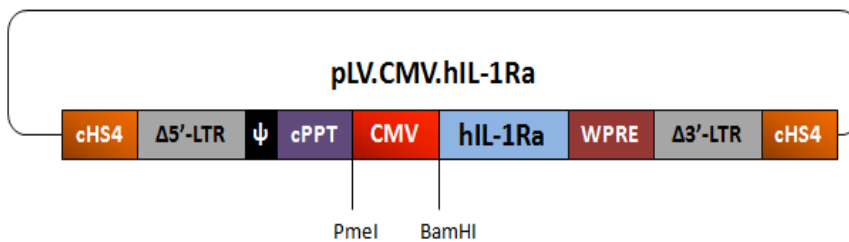
D. The mTNFRII-Fc gene (1356 bp) contains an internal SalI site therefore the reverse PCR primer contained an XhoI site. Restriction enzyme digestion with SalI and XhoI produce fragments with compatible 5'-TCGA-3' overhangs therefore, the mTNFRII-Fc gene was amplified from pFuse-mTNFRII-Fc using forward BamHI and reverse XhoI primers (Fig 5.1 D).
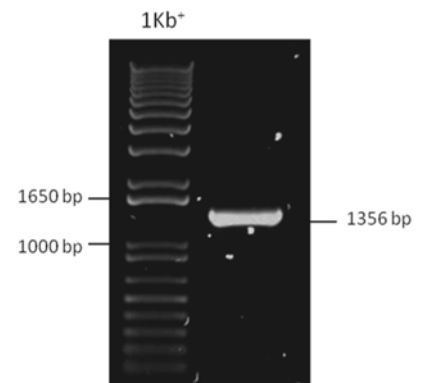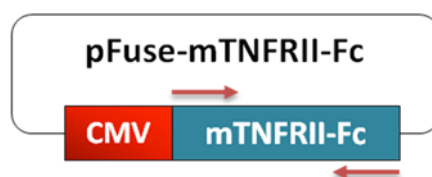
E. The mTNFRII-Fc PCR product was digested with BamHI and XhoI and cloned into the BamHI/SalI site within the pLV.CMV vector to generate pLV.CMV.mTNFRII-Fc The resulting construct was digested with PmeI and BamHI to release the CMV promoter. The 7756 bp fragment served as the cloning vector pLV.mTNFRII-Fc (Fig 5.1 E).
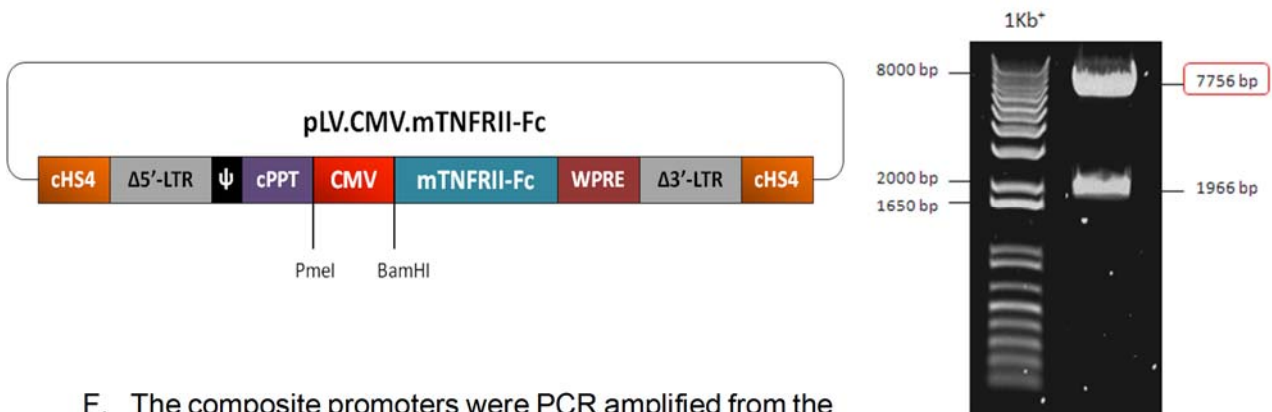
F. The composite promoters were PCR amplified from the original pCpG-20bp-composite promoter constructs, respectively (Chapter 3), using forward SnaBI and reverse BamHI primers (Fig 5.1 F).
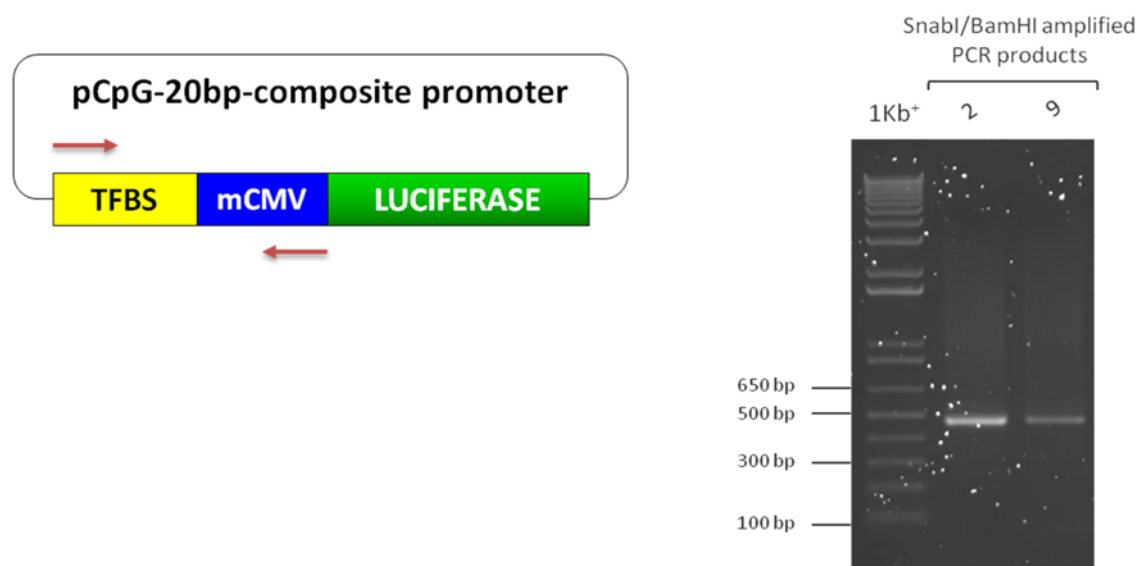
G. The PmeI and SnaBI enzymes produce blunt ended fragments which can be ligated together. Therefore, the synthetic promoter PCR products were digested with SnaBI and BamHI and cloned into the PmeI and BamHI site within the pLV.hIL-1Ra and

pLV.mTNFRII-Fc cloning vectors to generate lentiviral constructs containing the composite promoters driving expression of the therapeutic genes hIL-1Ra and mTNFRII-Fc, respectively (Fig 5.1 G).

pLV.Synthetic Promoter.hIL-1Ra

cHS4 | Δ5'-LTR | Ψ | cPPT | TFBS | mCMV | hIL-1Ra | WPRE | Δ3'-LTR | cHS4

Synthetic promoter

pLV.Synthetic Promoter.mTNFRII-Fc

cHS4 | Δ5'-LTR | Ψ | cPPT | TFBS | mCMV | mTNFRII-Fc | WPRE | Δ3'-LTR | cHS4

Synthetic promoter
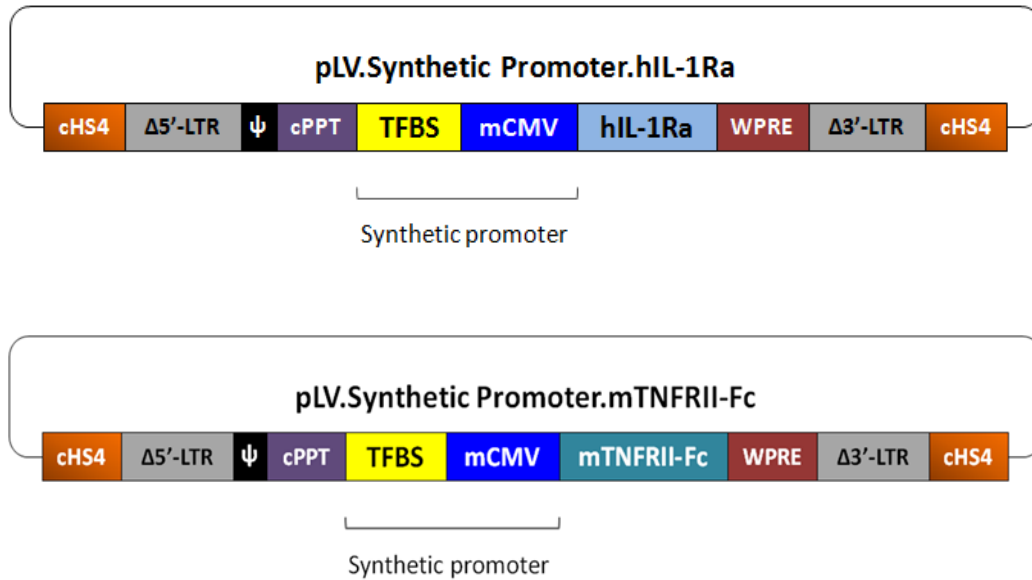
H. Forward and reverse oligonucleotides containing the mCMV promoter sequence were ordered from Sigma-Aldrich with 5'-SnabI and 3'-BamHI overhangs. The oligonucleotides were annealed using the boiling method and cloned into the PmeI and BamHI site within the pLV.hIL-1Ra and pLV.mTNFRII-Fc cloning vectors to generate the LV-mCMV-IL-1Ra and LV-mCMV-mTNFRII-Fc negative control constructs, respectively.

I. The SFFV promoter (560 bp) was PCR amplified from the pUCL-Luc⁺ construct using forward SnaBI and reverse BamHI primers (Fig 5.1 I).
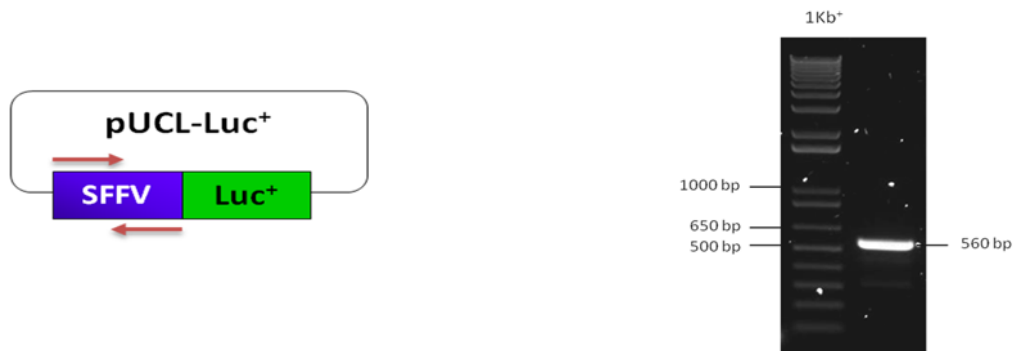
J. The PCR products were digested with SnaBI and BamHI and cloned into the PmeI/BamHI site within the pLV.hIL-1Ra and pLV.mTNFRII-Fc cloning vectors to generate lentiviral constructs containing the constitutively active SFFV promoter driving expression of the therapeutic genes hIL-1Ra and mTNFRII-Fc, respectively (Fig 5.1 J).
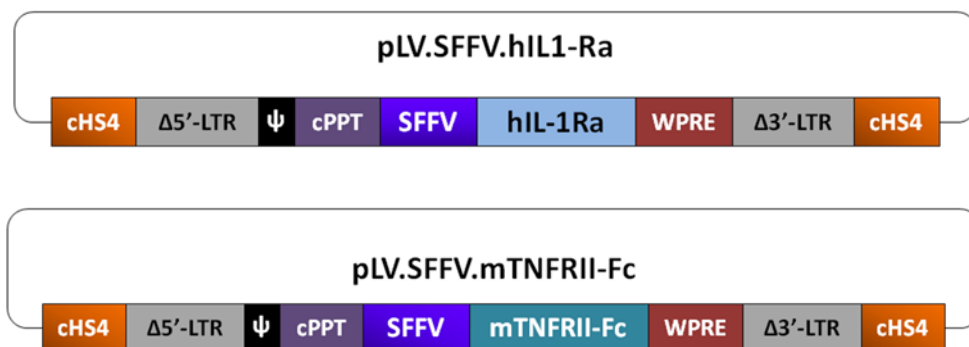
K. Following the successful construction of the lentiviral synthetic promoter constructs, the lentiviral particles were produced using a three-plasmid transient transfection into 293T cells with the transfer, packaging and envelope plasmids to generate lentiviral particles pseudotyped with the VSV-G envelope glycoprotein. The VSV-G pseudotyped lentiviral particles were transduced into 293T cells to generate a stable 293T cell line with the integrated synthetic promoters and the therapeutic genes.

Figure 5.1. Schematic diagram illustrating the cloning method used to generate the lentiviral composite synthetic promoters expressing soluble mTNFRII-Fc and hIL-1Ra. The pLV.CMV.eGFP lentiviral plasmid was systematically modified to generate the lentiviral constructs expressing the therapeutic genes

5.2. *In vitro* quantification of stable therapeutic protein expression regulated by inflammation-inducible composite synthetic promoters

Inflammation-inducible regulation of soluble mTNFRII-Fc and hIL-1Ra therapeutic protein expression by the composite promoters LV-2 and LV-9 and the control promoters LV-mCMV and LV-SFFV were evaluated using ELISA quantification. The stable 293T cells were unstimulated, incubated in hypoxia or stimulated with TNFα, PMA or their combination. The following day, the cell supernatant was collected and soluble mTNFRII-Fc and hIL-1Ra protein expression was quantified by ELISA. The therapeutic protein expression induced by promoters LV-2 and LV-9 were compared to the protein levels induced by the negative and positive control promoters, LV-mCMV and LV-SFFV, respectively (Fig 5.2).
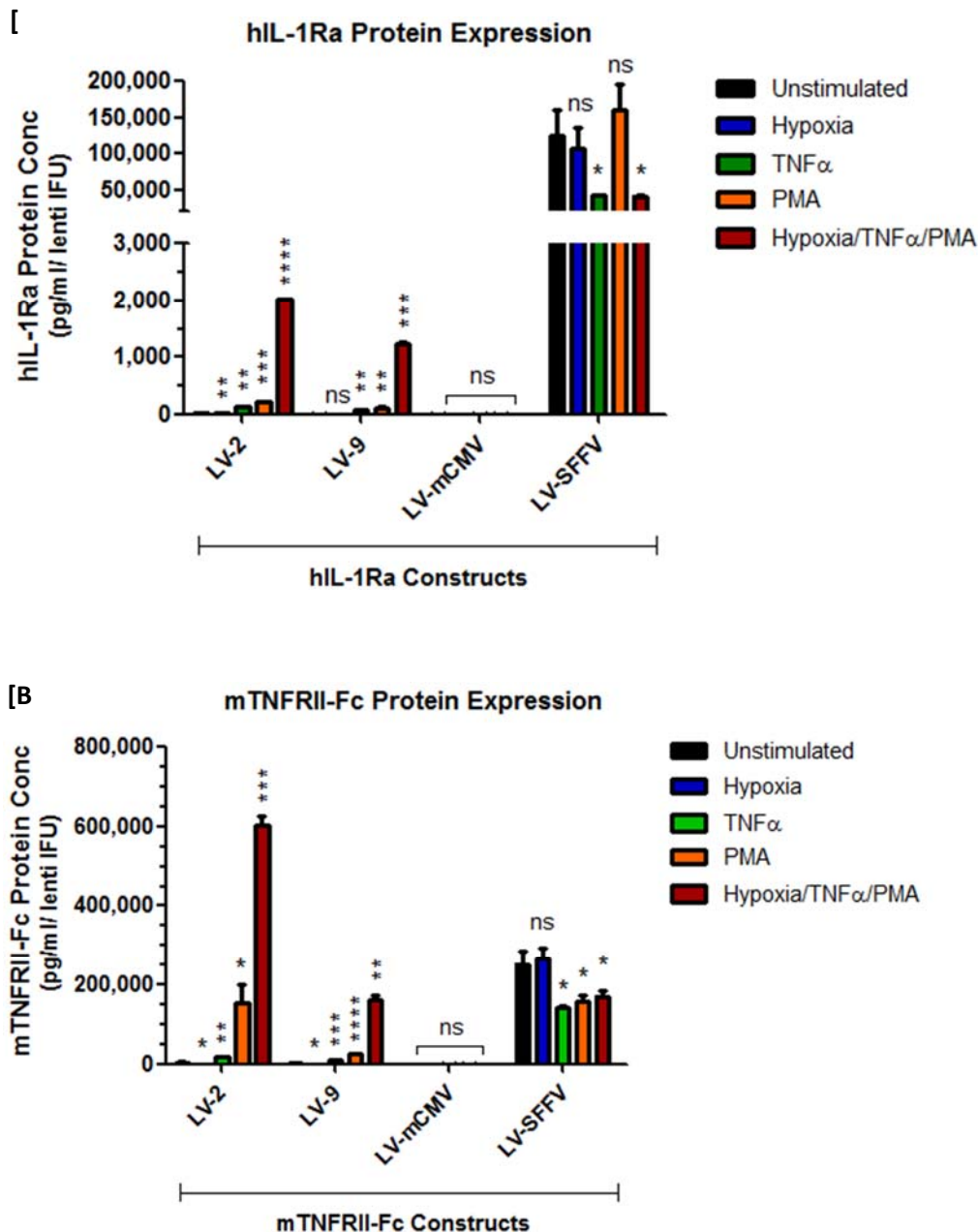
Figure 5.2. Promoters LV-2 and LV-9 demonstrate inflammation-inducible and synergistic expression of soluble mTNFRII-Fc and hIL-1Ra proteins. The stable 293T cells were seeded at a density of 50,000 cells per well in a 6-well plate and cultured at 37°C for 24 hours. After 24 hours, the cells were unstimulated, incubated in hypoxia (0.1% O_2) or stimulated with TNFα (10 ng/ml), PMA (10 ng/ml) or a combination of TNFα and PMA in hypoxia for 18 hours. The cell supernatant was collected 18 hours post-stimulation and the expression of hIL-1Ra [A] and soluble mTNFRII-Fc [B] protein was quantified by ELISA and normalised to the lentiviral titre (expressed as pg/ml/lenti IFU). The data represents the mean ± SD of triplicate values. The statistical significance between the unstimulated and stimulated luciferase gene expression was calculated using the Student's t-test (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, **** = p≤0.0001).

The inflammation-inducible regulation of therapeutic protein expression by the composite promoters LV-2 and LV-9 was confirmed by quantifying the secreted protein expression of mTNFRII-Fc and hIL-1Ra in the supernatant of unstimulated or stimulated LV-2, LV-9,   LV-mCMV and LV-SFFV-stably transduced 293T cells, using ELISA. The sensitivity of mTNFRII-Fc and hIL-1Ra detection was 37 pg/ml and 4 pg/ml, respectively.

As anticipated, promoter LV-mCMV did not induce mTNFRII-Fc and hIL-1Ra therapeutic protein expression irrespective of stimulation (ns $p > 0.05$). In contrast, constitutive therapeutic protein expression was induced by the constitutively active promoter LV-SFFV, although significant decreases in mTNFRII-Fc and hIL-1Ra expression can be observed following inflammatory stimulation (* $p \leq 0.05$). Generally, the composite promoters LV-2 and LV-9 retained their differential, multi-responsive and synergistically-inducible gene expression profiles. In the absence of stimulation, the LV-2 induced low/moderate basal protein expression of mTNFRII-Fc and hIL-1Ra whereas LV-9 induced very low basal expression of mTNFRII-Fc and barely detectable hIL-1Ra protein expression. Encouragingly, the low/moderate basal expression exhibited by promoter LV-2 in comparison to the very low basal expression induced by promoter LV-9 was consistent with all of the previous functional analysis data of these promoters in Chapters 3 and 4.

Following hypoxia incubation, there were marginal increases in mTNFRII-Fc and hIL-1Ra protein expression induced by promoters LV-2 and LV-9, which is in accordance with the previously demonstrated low hypoxia-inducibility of these promoters. However, there were significant but modest increases in hIL-1Ra protein expression induced by LV-2 and LV-9, in response to TNFα and PMA stimulation ($p \leq 0.01$ to $p \leq 0.001$). In contrast, robust mTNFRII-Fc protein expression was induced by LV-2 and LV-9 following TNFα and PMA stimulation ($p \leq 0.02$ to $p \leq 0.0001$). Noticeably, promoters LV-2 and LV-9 demonstrated highly significant synergistic induction of mTNFRII-Fc and hIL-1Ra protein expression following combined inflammatory and hypoxic stimulation ($p \leq 0.01$ to $p \leq 0.0001$). Therefore, on the basis of low/moderate basal and high inflammation-inducible gene expression levels, promoter LV-2

was selected as the most promising promoter to regulate mTNFRII-Fc and hIL-1Ra therapeutic protein expression to potentially reduce paw inflammation.

5.3.   Evaluating the therapeutic efficacy of soluble mTNFRII-Fc and hIL-1Ra protein expression regulated by the inflammation-inducible composite promoter LV-2 during paw inflammation

The therapeutic efficacy of soluble mTNFRII-Fc and hIL-1Ra therapeutic proteins regulated by the inflammation-inducible composite promoter LV-2, the negative control LV-mCMV promoter and the positive control LV-SFFV promoter was assessed in a carrageenan-induced paw inflammation model (Fig 5.3).
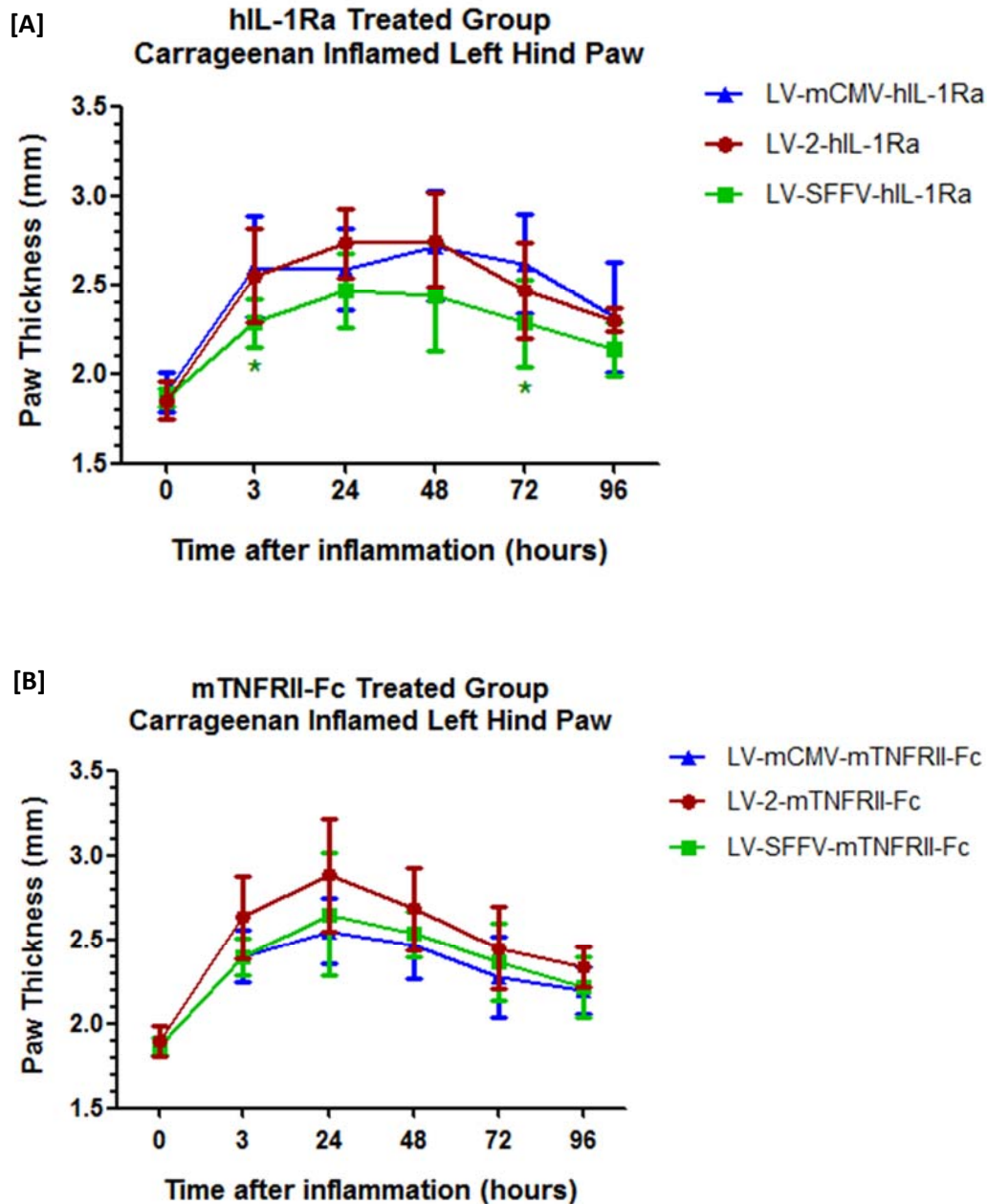
Figure 5.3. Comparative analysis of the therapeutic effect of mTNFRII-Fc and hIL-1Ra protein expression regulated by the inflammation-inducible composite promoter LV-2. Adult CD1 male mice (n=6 per group) were injected intraplantarly with 830,000 lentiviral IFU (25 μl) into the left hind paw with LV-mCMV-hIL-1Ra, LV-2-hIL-1Ra, LV-SFFV-hIL-1Ra [A] and LV-mCMV-mTNFRII-Fc, LV-2-mTNFRII-Fc, LV-SFFV-mTNFRII-Fc [B]. After 7 days, the left hind paws received an intraplantar injection of 50 μl 1% λ-carrageenan whereas the control right paws were injected intraplantarly with 50 μl of saline. Paw thickness was measured before inflammation (0 hours)

The lentiviral particles expressing the therapeutic proteins were administered 7 days prior to the induction of carrageenan-induced paw inflammation. During this time interval before inflammation, the synthetic promoters are expected to induce basal expression of the therapeutic gene; the SFFV promoter is expected to induce a high level of constitutive therapeutic gene expression, the LV-2 promoter should display low/moderate basal expression whereas the mCMV promoter should induce negligible therapeutic gene expression during this un-inflamed state, as confirmed by the ELISA experiment in Figure 5.2. It was proposed that mice injected with LV-mCMV-mTNFRII-Fc and LV-mCMV-hIL-1Ra would not have reduced inflammation due to the lack of therapeutic protein expression induced by the LV-mCMV promoter thus serving as the negative control group for comparisons to the other data groups.

As anticipated, Figure 5.3 A shows that the mice treated with LV-SFFV-hIL-1Ra had the lowest degree of paw swelling at all time-points, which was significantly lower than the paw thickness of LV-mCMV-hIL-1Ra treated mice at 3 hours (* p=0.04) and 72 hours (* p=0.05) post-inflammation only. The general reduction of paw thickness in the LV-SFFV-hIL-1Ra treated mice (compared to mice treated with LV-mCMV-hIL-1Ra) is consistent with the principle of IL-1Ra therapy: IL-1Ra blocks IL-1-mediated signalling by competitively binding to the IL-1 receptors and after the onset of carrageenan-induced inflammation, some of the IL-1 receptors were occupied by IL-1Ra induced by the constitutive SFFV promoter, thereby reducing the overall level of paw inflammation in the LV-SFFV-hIL-1Ra treated mice. High molar excess of IL-1Ra, preferably before the onset of disease is required for high therapeutic efficacy of IL-1Ra therapy and despite the lower magnitudes of paw inflammation compared to the LV-mCMV-hIL-1Ra treated mice, the level of paw inflammation continued to increase in the LV-SFFV-hIL-1Ra treated mice. These observations suggest that the   LV-SFFV promoter may not have produced sufficient concentrations of IL-1Ra to block IL-1 activity and prevent progressive paw inflammation, which would indicate that suboptimal lentiviral titres were used in this experiment. Similarly, the administration of LV-2-hIL-1Ra did not reduce paw

inflammation following carrageenan-induced paw inflammation which may be due to the low basal IL-1Ra expression from the LV-2 promoter and possibly, the delivery of suboptimal lentiviral titres.

In contrast, the principle of mTNFRII-Fc therapy is to bind free TNFα at a 1:1 ratio thereby serving as a decoy receptor to prevent TNFα-mediated cell signalling. It was expected that mTNFRII-Fc expressed from the constitutive LV-SFFV promoter would achieve the greatest reduction in paw inflammation due to high constitutive mTNFRII-Fc protein concentrations. It was also expected that the LV-2 promoter would synergistically-induce therapeutic protein expression during peak carrageenan-induced inflammation (~24 hours post-carrageenan injection) to subsequently reduce paw inflammation after this time point. These actions would require high therapeutic protein concentrations however, neither LV-SFFV nor LV-2 promoter induced sufficient mTNFRII-Fc protein expression to reduce paw inflammation, when compared to the LV-mCMV-mTNFRII-Fc treated mice (Fig 5.3 B). The general reduction of paw inflammation observed in Figure 5.3 (A and B) is consistent with the profile of carrageenan-induced paw inflammation observed in section 4.3 (Fig 4.8 H), which implies that although the carrageenan-induced paw inflammation was reliably reproduced in both experiments, the local delivery of therapeutic proteins did not reduce paw inflammation.

5.4. Discussion

It is widely recognised that the pro-inflammatory cytokines TNFα and IL-1β are abundant in the RA joint and play crucial roles in the pathogenesis of RA. TNFα and IL-1β-mediated cell signalling induces the expression of downstream pro-inflammatory mediators and destructive enzymes which exacerbate chronic inflammation and promote cartilage and bone resorption (Feldmann and Maini, 2008). The notion of neutralising the biological activities of TNFα and IL-1β led to the development of anti-cytokine therapy for the treatment of RA and significant experimental evidence has been demonstrated by local and systemic administration of soluble TNFRII-Fc and IL-1Ra, respectively.

The soluble TNFRII-Fc consists of the extracellular binding domain of the p75 TNFα receptor fused to the Fc region of mouse IgG1. This fusion protein is a competitive inhibitor of TNFα which binds to and neutralises TNFα preventing TNFα-mediated signal transduction (Moreland *et al.,* 1997). IL-1Ra inhibits the biological activity of IL-1 (α and β) by competitively binding to the type-1 IL-1 receptor and preventing IL-1 receptor mediated signalling (McIntyre *et al.,* 1991). The recombinant human therapeutic proteins, soluble TNFR-Fc and IL-1Ra, are commercially known as Etanercept (Enbrel®) and Anakinra (Kineret®), respectively and Etanercept is a routine 'biological' therapy for treating RA. Although significant therapeutic efficacy has been demonstrated with IL-1Ra therapy in experimental models of RA, Anakinra is not recommended for the treatment of RA, except in the context of a controlled, long-term clinical study due to the imbalance in clinical benefits and cost effectiveness (NICE clinical guideline 79: Rheumatoid Arthritis).

Despite clinical success of anti-TNFα protein therapy in RA patients, the systemic and repeated administrations of such biological therapies at high-doses continue to raise safety concerns such as immunosuppression which increases the risk of developing infections. An attractive alternative is the local delivery of therapeutic genes expressed from inflammation-inducible composite synthetic promoters which offer anti-inflammatory effects during disease

flares and reduced activity during disease remission (van de Loo *et al.,* 2004). Although many inflammation-inducible systems have demonstrated therapeutic efficacy in animal models of paw inflammation, the feasibility of multi-responsive and synergistically-inducible composite synthetic promoters for local RA gene therapy has yet to be demonstrated and this was a central component of my PhD.

Section 5.2 described the *in vitro* quantification of soluble mTNFRII-Fc and hIL-1Ra proteins in the supernatant of unstimulated or stimulated LV-mCMV, LV-SFFV, LV-2 and LV-9 stable 293T cells using ELISA. These results verified the high expression of the therapeutic proteins by LV-2 and LV-9 composite promoters in a synergistic and inflammation-inducible manner, compared to the constitutive and the undetected therapeutic protein expressions induced by the LV-SFFV and LV-mCMV promoters, respectively.

Based on the consistent and favourable expression profile of promoter LV-2, the therapeutic potential of LV-2-regulated therapeutic protein expression was evaluated in a carrageenan-induced paw inflammation model. Unfortunately, LV-2 promoter-driven hIL-1Ra and mTNFRII-Fc protein expression did not reduce paw inflammation. However, mice treated with LV-SFFV-hIL-1Ra displayed lower paw inflammation at all time-points throughout the 96 hour monitoring period, although only the changes at 3 hours and 72 hours post-inflammation were statistically significant (both * $p \leq 0.05$). This effect might be due to the high and constitutive hIL-1Ra gene expression and semi-efficient blockade of the IL-1 receptors with IL-1Ra during the 7 days prior to carrageenan-induced paw inflammation. This observation is in agreement with the suggestion that IL-1Ra therapy is more effective when administered before the onset of disease, which ensures that the IL-1 receptors are occupied with IL-1Ra to block IL-1 cell signalling (Bakker *et al.,* 1997). However, following carrageenan-injection, paw inflammation in the LV-SFFV-hIL-1Ra treated mice still continued to increase (albeit lower), which suggests that the lentiviral titres were not enough to achieve high hIL-1Ra protein concentrations to efficiently block the IL-1 receptors and elicit the desired therapeutic effects. This observation corroborates the finding that efficient biological inactivation of *in vitro* and *in vivo* IL-1-

meditated signalling requires a large molar excess (10-1000 fold) of IL-1Ra over IL-1 (Arend *et al.,* 1990; Fischer *et al.,* 1991) and more than 95% of IL-1 receptors need to be occupied by IL-1Ra to efficiently block IL-1 signalling (Arend, 2002) which may not have been achieved due to the low lentiviral titre. Also, the very low hIL-1Ra protein expression *in vitro* (Fig 5.2 A) is in accordance with the absence of LV-2-IL-1Ra mediated reduction in paw swelling observed in Figure 5.3 A.

The delivery of mTNFRII-Fc (and soluble TNFRII) has been demonstrated to successfully reduce paw inflammation in experimental mouse models of paw inflammation due to the inhibition of TNFα-mediated signalling (Khoury *et al.,* 2007; Hughes *et al.,* 2010). Using the carrageenan-induced paw inflammation mouse model, Mazzon *et al.,* (2008) demonstrated that the systemic administration of Etanercept (2 hours before inflammation) significantly attenuated the development of paw inflammation to a similar extent as that seen in TNFα receptor-1 knockout mice, which confirmed the importance of TNFα in this model, particularly during the early phase of paw inflammation. In contrast to the findings of Mazzon *et al.,* (2008), the results presented in Figure 5.3 B showed that the mice treated with mTNFRII-Fc induced by promoters LV-2 and LV-SFFV did not display reduced paw inflammation when compared to the level of paw inflammation in LV-mCMV-mTNFRII-Fc treated mice, which is inconsistent with published research and provides further evidence that the suboptimal lentiviral titre of particles expressing mTNFRII-Fc may have also been too low to efficiently block TNFα-mediated signalling.

In a recent study by Garaulet *et al.,* (2013), the authors developed an inflammation-inducible E-selectin (ESEL) synthetic promoter-based lentiviral-expression system. The intraplantar delivery of the lentiviral particles encoding the ESEL-promoter into the paws of mice resulted in inflammation-specific therapeutic IL-10 gene expression which reduced paw inflammation following the induction of zymosan-induced paw inflammation. Their experiments were of a similar format to the experiments discussed in this chapter, however, the titre of lentiviral particles encoding therapeutic IL-10 gene (2×10^7 transducing units) was substantially higher

than that used in the experiments in section 5.3 (830,000 lentiviral IFU). The authors also noted that the zymosan-induced mouse model was a suitable model for their experiment due to the high induction of inflammatory-interleukins following local zymosan injection, which could be reduced by therapeutic IL-10 expression (Garaulet *et al.,* 2013). This logically implied that the efficacy of the administered therapy is dependent on the presence of the target inflammatory mediators in the experimental model.

For example, in order to efficiently evaluate the anti-inflammatory effects of locally expressed mTNFRII-Fc and hIL-1Ra proteins, the target pro-inflammatory cytokines TNFα and IL-1β should play a major role in the inflammatory process in the carrageenan-induced inflammation mouse model. Carrageenan-induced paw inflammation in mice is invariably characterised by elevated production of histamine, serotonin, bradykinin, prostaglandins, nitric oxide and infiltrating neutrophils (Posadas *et al.,* 2004) and with the exception of Rocha *et al.,* (2006) and Mazzon *et al.,* (2008) who showed that TNFα plays a role in carrageenan-induced paw inflammation, there is limited research implicating TNFα (and IL-1β) as key inflammatory mediators in this model. With hindsight, the therapeutic effects of soluble TNFRII-Fc and IL-1Ra could be more effective in experimental models of paw inflammation mediated by TNFα and IL-1β such as the CIA or AIA models, which share more similarities to human RA than the carrageenan-induced inflammation model. Mouse CIA and AIA models also offer the advantage of longer durations of inflammation and the ability to reactivate inflammation to mimic the relapse/remission phases of the RA would enable longer monitoring periods to fully evaluate the therapeutic efficacy of therapeutic genes expressed from inflammation-inducible composite synthetic promoters. Due to time constraints, this experiment could not be repeated in a more suitable mouse model of arthritis and/or paw inflammation, and represents an important option for further investigation.

In the preceding Chapter 4, the luciferase gene was used as a surrogate for the therapeutic genes and promoter LV-2 successfully demonstrated inflammation-specific luciferase expression in the carrageenan-induced inflamed paw as confirmed by real-time

bioluminescent imaging. In contrast to the highly sensitive detection of luciferase gene expression using adjustable bioluminescent imaging, the functional readout of the therapeutic efficacy of soluble mTNFRII-Fc and hIL-1Ra proteins was their ability to significant reduce the level of paw inflammation, which was a more challenging task that will require optimisation.

Overall, the reduction of inflammation in mice treated with LV-2-regulated mTNFRII-Fc and hIL-1Ra proteins was anticipated but not observed. It is likely that the lack of significant reductions in paw inflammation may have been due to 1) suboptimal lentiviral titres which impaired the abundant expression of the therapeutic proteins prior to and during inflammation, 2) low basal expression of IL-1Ra which failed to efficiently occupy the IL-1 receptors to prevent IL-1-mediated signalling or 3) the use of the carrageenan-induced inflammation model where the roles of TNFα and IL-1β have not been fully defined. To conclude, these limitations can potentially be rectified by optimising the lentiviral titre, reconstructing the composite promoters to achieve higher basal and induced gene expression levels and utilising other mouse models of paw inflammation, all of which are interesting future work options.

CHAPTER 6:

Post-Transcriptional Regulation of Luciferase

Gene Expression by Inflammation-responsive

miR-23b

6.1. Introduction

The development of inflammation-responsive gene expression systems, which harness the endogenous transcriptional and post-transcriptional cellular machinery, is the central concept of the experiments described in this thesis. In this chapter, I investigate the possibility of post-transcriptional regulation of constitutive luciferase gene expression using microRNAs (miRNAs), whose expression is regulated by inflammatory signals.

Mature miRNAs are small (~21 nt), noncoding RNA species which bind to their target sites within the 3' untranslated region (3'UTR) of mRNAs to negatively regulate gene expression by promoting the degradation or inhibiting the translation of target mRNAs (Bartel, 2004). Many recent publications have demonstrated the feasibility of exploiting the endogenous activity of miRNAs by incorporating the target sequences of miRNAs into the 3'UTR of the transgene. In this way, the transgene mRNA is subjected to post-transcriptional regulation by the candidate miRNA, resulting in differential gene expression (Brown *et al.,* 2007b; Brown and Naldini, 2009). Analogous to this concept, the aim of the experiments in Chapter 6 is to provide proof-of-principle of inflammation-specific gene expression mediated by an inflammation-repressed miRNA.

A promising candidate to demonstrate this concept was miR-23b (miR-23b-3p), which was identified by Zhu *et al.,* (2012) to be:

- downregulated in inflammatory lesions of patients with autoimmune RA and systemic lupus erythematosus (SLE)
- downregulated in tissue samples from mice with collagen-induced arthritis (CIA) and experimental autoimmune encephalomyelitis (EAE)
- downregulated in human fibroblast-like synoviocytes, mouse primary kidney cells and mouse embryonic fibroblasts treated with inflammatory stimuli
- highly expressed in normal tissues

Exploiting the differential expression of miR-23b between inflammation and the normal state holds great promise for the development of inflammation-regulated gene expression. Figure 6.1 depicts the concept of miR-23b-mediated regulation of therapeutic gene expression.
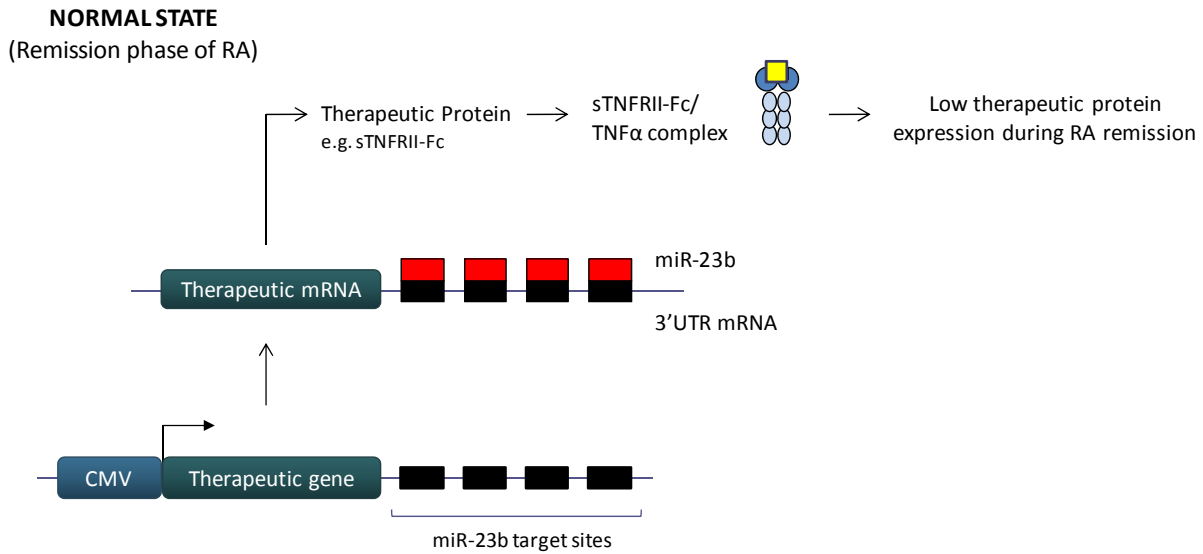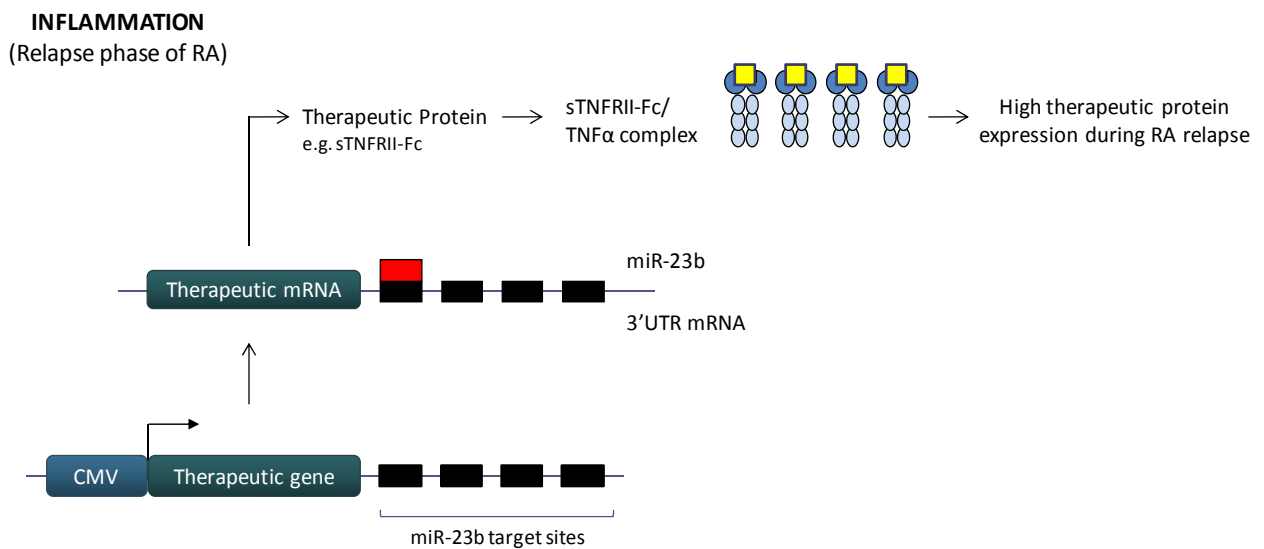
Figure 6.1. Schematic diagram illustrating the concept of post-transcriptional regulation of therapeutic gene expression by inflammation-repressed miR-23b. Perfectly complementary miR-23b target sites can be cloned into 3'UTR of the therapeutic gene, driven by a constitutive promoter. Following transcription, the therapeutic mRNA containing miR-23b target sites will be subjected to regulation by inflammation-repressed miR-23b. During RA remission, miR-23b is highly expressed and can bind to their cognate sites in the therapeutic mRNA 3'UTR and target the therapeutic mRNAs for degradation, resulting in low therapeutic protein concentrations [A]. During inflammation, the miR-23b expression is downregulated, therefore few or no miR-23b molecules will bind to their target sites in the therapeutic mRNA 3'UTR and the untargeted mRNAs can be translated into the therapeutic protein to reduce inflammation [B].

6.1.1. Construction of miR-23b-regulated luciferase expression cassettes

The luciferase gene was used as a surrogate marker for the expression levels of a therapeutic gene. The schematic diagrams below (Fig 6.2) depict the cloning strategy used to construct the miR-23b-regulated luciferase expression cassettes.

A. The plasmid DNA construct pcLuc$^+$ (modified from pcDNA3.1$^+$ by Gould *et al.,* 2004) contains the constitutive CMV promoter driving expression of the luciferase gene (Luc$^+$). The construct was digested with XbaI and ApaI, within the 3'UTR of the luciferase gene and upstream of the poly A signal, to create the pcLuc$^+$ cloning vector (Fig 6.2 A).
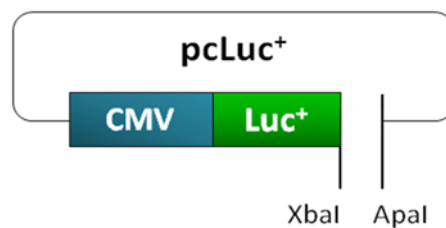
B. The miR-23b-target oligonucleotide contained two perfect complementary miR-23b target sites, which are the reverse complement sequences of human miR-23b-3p. The oligonucleotides were ordered with 5'-XbaI and 3'-ApaI overhangs and cloned into the XbaI/ApaI site within the pcLuc$^+$ vector to generate the pcLuc$^+$-miR-23b-2T construct (Fig 6.2 B).
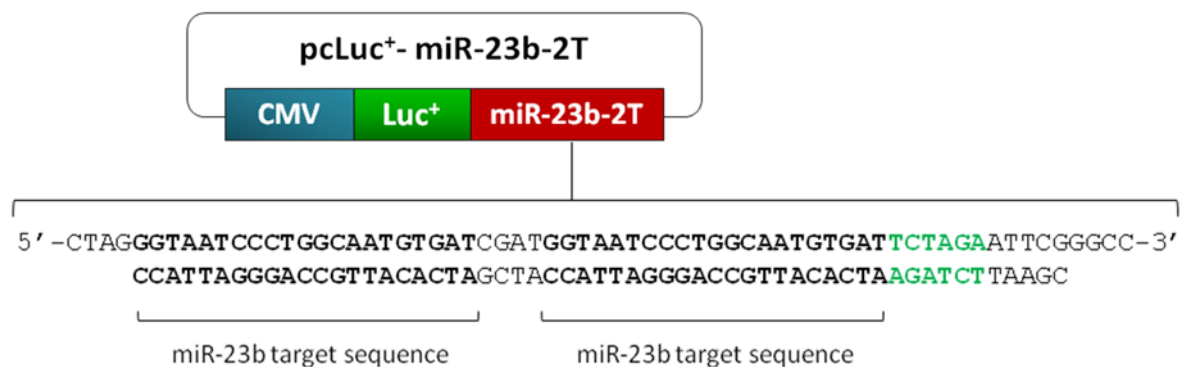
C. Cloning of the miR-23b-target oligonucleotide into the pcLuc⁺ vector destroys the XbaI site and restores the ApaI restriction site. The miR-23b-target oligonucleotide contains an internal XbaI restriction site (5'-TCTAGA-3', shown in green above). Therefore, the pcLuc⁺-miR-23b-2T construct was digested with XbaI and ApaI to clone an additional miR-23b-target oligonucleotide to generate the expression construct pcLuc⁺-miR-23b-4T, containing four miR-23b target sites.

D. Successful cloning of pcLuc⁺-miR-23b-2T and pcLuc⁺-miR-23b-4T constructs were authenticated by an EcoRI restriction enzyme digest, which released 1800 bp and 1860 bp fragments, respectively (Fig 6.2 D). Positive constructs were verified by DNA sequencing (Appendix 12.1 and 12.2).
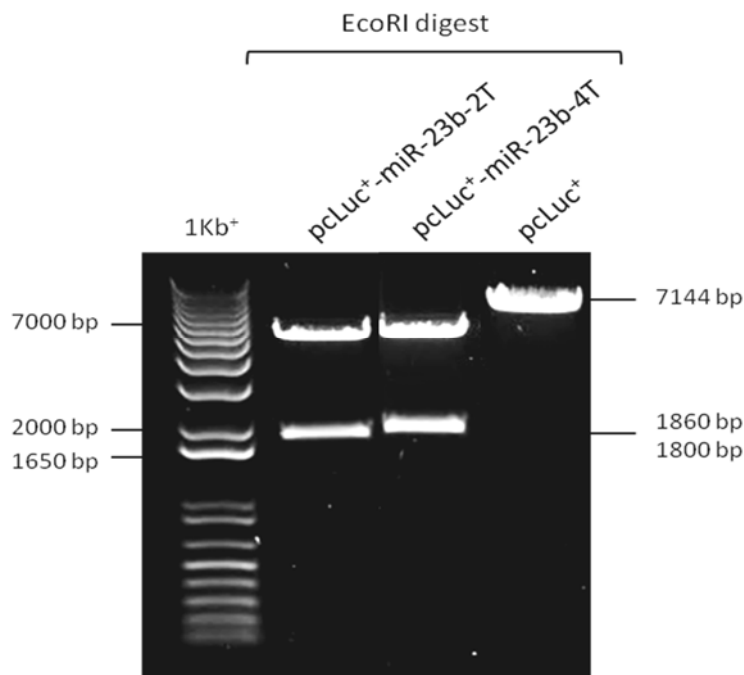
Figure 6.2. Construction of pcLuc⁺-miR23b-target expression vectors. The pcLuc⁺ plasmid was systematically modified to generate constructs with two or four miR-23b target sites in the 3'UTR of the luciferase gene.

6.2. Successful knockdown of luciferase gene expression using synthetic miR-23b mimics

Synthetic miRNA mimics are double stranded RNA molecules which 'mimic' the activity of endogenous miRNAs by post-transcriptionally downregulating target mRNAs. To investigate whether miR-23b can target and repress luciferase mRNAs containing miR-23b target sites, 1 μM miR-23b mimic or 1 μM control miR-648 mimic was co-transfected with pcLuc$^+$, pcLuc$^+$-miR-23b-2T or pcLuc$^+$-miR-23b-4T and also pRL-CMV in 293T cells. The luciferase protein expression was quantified and presented in Figure 6.3.
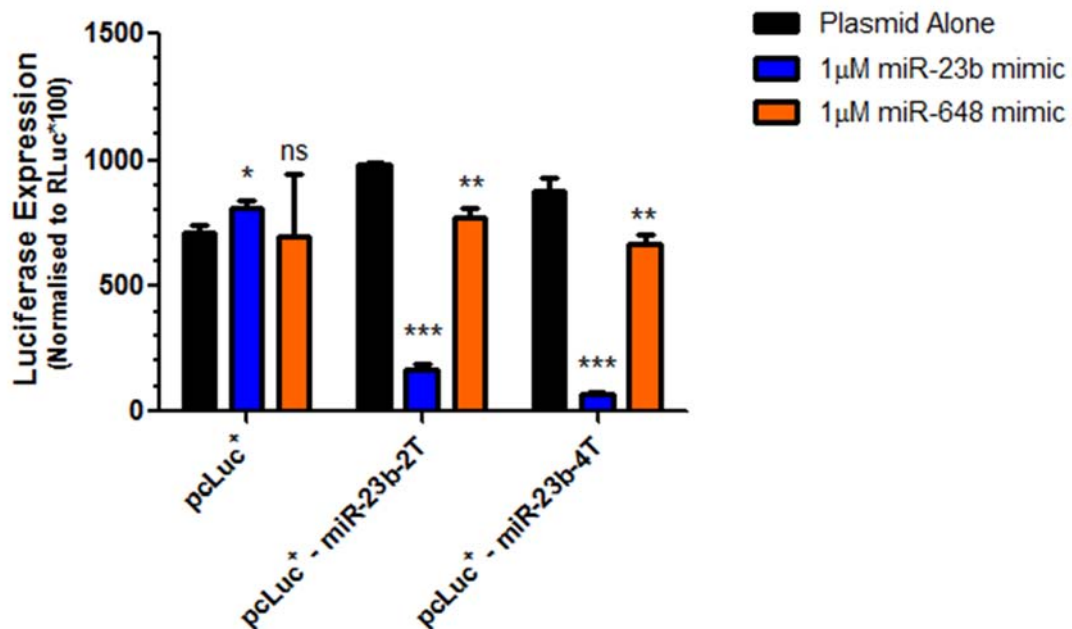
Figure 6.3. Successful knockdown of luciferase gene expression using a synthetic miR-23b mimic. 293T cells (32,000 cells) were seeded in a 48-well plate. After 24 hours, the pcLuc$^+$, pcLuc$^+$-miR-23b-2T and pcLuc$^+$-miR-23b-4T constructs (90 ng) were co-transfected with the renilla expressing plasmid pRL-CMV (10 ng) and 1 μM miR-23b mimic and 1 μM miR-648 mimic (control) using Lipofectamine 2000 in 293T cells. The data represents the mean ± SD of triplicate values normalised to renilla luciferase. The statistical significance compared to the corresponding 'plasmid alone' datasets for each group was calculated using the Student's t-test (ns= p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001).

As anticipated, miR-23b or miR-648 mimic had marginal effects on the miRNA target-less luciferase mRNA expressed from pcLuc$^+$ control vector (* p=0.02 and ns =0.94, respectively). Impressively, miR-23b mimic significantly reduced luciferase gene expression from the pcLuc$^+$-miR-23b-2T (*** p=0.0002; 83% reduction) and pcLuc$^+$-miR-23b-4T   (*** p=0.001; 92% reduction), which contain two and four miR-23b target sites, respectively. In contrast, the non-specific miR-648 control mimic had a less significant effect on luciferase mRNAs with miR-23b target sites, expressed from pcLuc$^+$-miR-23b-2T (** p=0.01; 22% reduction) and pcLuc$^+$-miR-23b-4T (** p=0.01; 24% reduction).

Luciferase mRNAs with four miR-23b target sites (expressed from pcLuc$^+$-miR-23b-4T) were suppressed to a greater extent than luciferase mRNAs with two target sites which indicates that downregulation of gene expression by miRNAs is influenced by the number of miRNA target sites within the mRNA.

This experiment strengthened the established concept of cloning miRNA target sites into the transgene 3'UTR to post-transcriptionally regulate transgene expression. The use of synthetic miRNA mimics confirmed that the miR-23b target sites were correctly cloned into the expression cassette and could potentially be targeted by endogenous miR-23b.

6.3. Inflammation-induced downregulation of endogenous miR-23b expression in NIH3T3 mouse embryonic fibroblasts

Endogenous miR-23b is downregulated in autoimmune pathology and also downregulated in primary human FLS, mouse primary kidney cells and mouse embryonic fibroblasts treated with inflammatory stimuli (Zhu *et al.,* 2012). Therefore, the differential expression of miR-23b in response to the presence or absence of inflammatory stimulation in NIH3T3 mouse embryonic fibroblasts was confirmed by Absolute Real-time qPCR. The miR-23b copy number (per cell) was quantified in unstimulated, mIL-17A, TNFα and TNFα+mIL-17A stimulated NIH3T3 cells using the standard curve method (Fig 6.4). The standard curve plots are presented in Appendix 15.
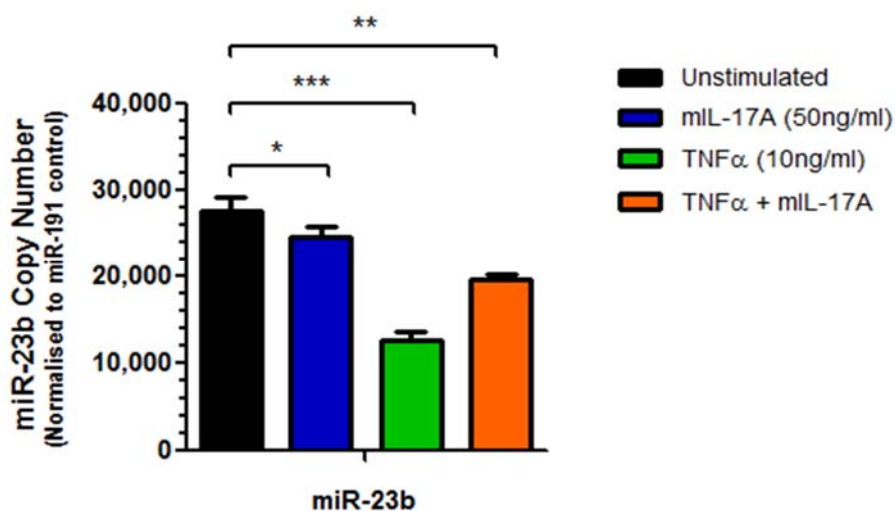
Figure 6.4. miR-23b is significantly downregulated in response to inflammatory stimuli in NIH3T3 cells. NIH3T3 cells (1.5×10^6) were unstimulated or stimulated with mIL-17A (50 ng/ml), TNFα (10 ng/ml) or mIL-17A with TNFα for 18 hours. The small RNA fraction containing miRNAs was isolated and 100 ng RNA was reverse transcribed to generate cDNA. Each cDNA sample was amplified using miR-23b, U6 control, miR-16 control, miR-17-5p control, miR-103 control, miR-191 control forward primers and a universal reverse primer in an Absolute Real-time qPCR reaction. Absolute quantification of the miRNAs was determined using the

This experiment shows that miR-23b is highly expressed in unstimulated NIH3T3 cells and is significantly downregulated in response to stimulation with mIL-17A (* p=0.048; 11% reduction), TNFα (*** p=0.0002; 55% reduction) and a combination of TNFα and mIL-17A (** p=0.01; 29% reduction) in the same cell type. These results confirmed the previously reported inflammation-repressed expression profile of endogenous miR-23b, which can potentially be harnessed for inflammation-responsive gene regulation.

6.4 Efficient regulation of luciferase gene expression by inflammation-responsive endogenous miR-23b in NIH3T3 mouse embryonic fibroblasts

The ability of endogenous miR-23b to achieve inflammation-specific luciferase gene expression was investigated in unstimulated and stimulated NIH3T3 and 293T cells. The cells were co-transfected with pcLuc$^+$, pcLuc$^+$-miR23b-2T, pcLuc$^+$-miR23b-4T and also pRL-CMV. Human and mouse miR-23b sequences are identical, therefore the constructs can be active in both human and mouse cells (Fig 6.5).

miR-23b-mediated regulation in NIH3T3 cells

miR-23b-mediated regulation in 293T cells
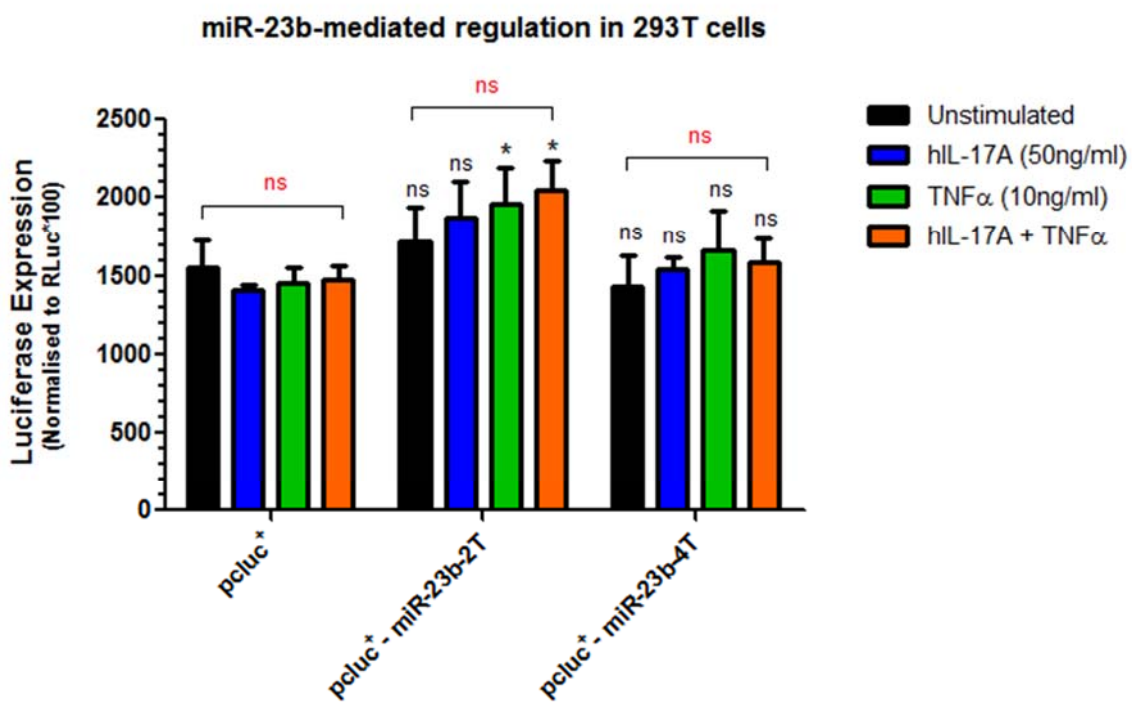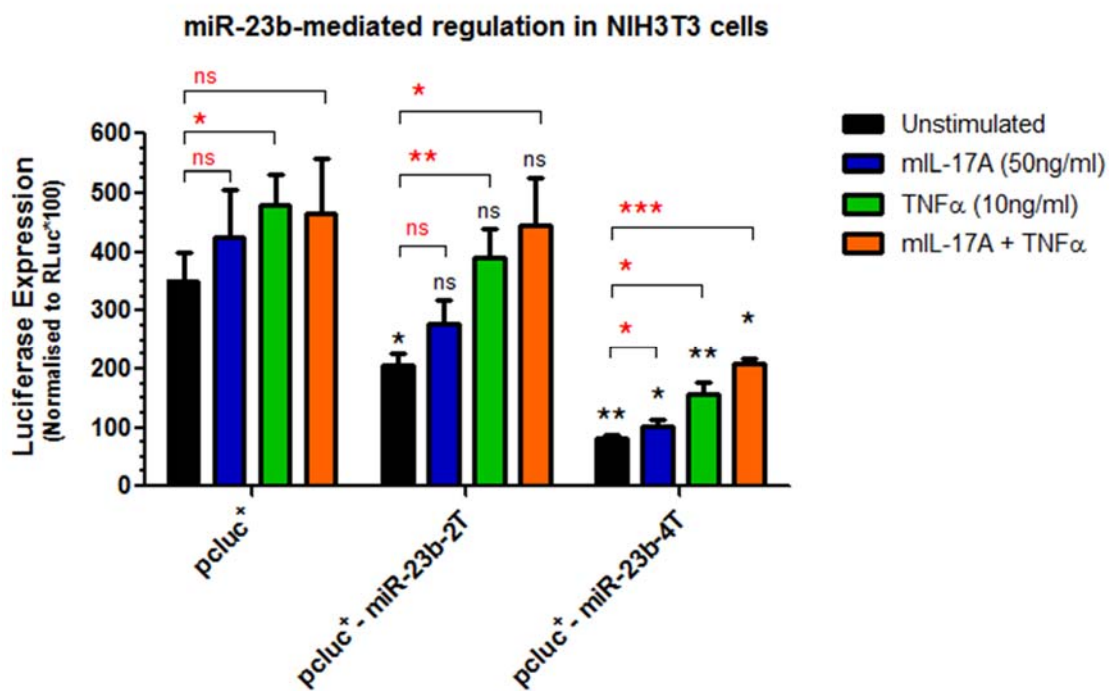
Figure 6.5. Cell-specific and inflammation-responsive luciferase gene regulation mediated by endogenous miR-23b. NIH3T3 cells [A] and 293T cells [B] were seeded at 20,000 cells/well in a 96 well plate. After 24 hours, pcLuc⁺, pcLuc⁺-miR-23b-2T or pcLuc⁺-miR-23b-4T (180 ng) were co-transfected with pRL-CMV (20 ng) in NIH3T3 cells [A] and 293T cells [B]. Transfected NIH3T3 and 293T cells were unstimulated or stimulated with IL-17A (50 ng/ml), TNFα (10 ng/ml) or TNFα + IL-17A for 18 hours. Mouse IL-17A (mIL-17A) and human IL-17A (hIL-17A) was used to stimulate the mouse NIH3T3 cells and human 293T cells, respectively. The data represents the mean ± SD of triplicate values normalised to renilla luciferase. The statistical significance was calculated using the Student's t-test (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001). The statistical significance shown in black compares the luciferase expression from each miR-23b target-bearing constructs (miR-23b-2T and -4T) to the target-less pcLuc⁺ construct during the same conditions. The statistical significance shown in red compares the luciferase expression during inflammatory stimulation compared to the basal expression induced by the same construct.

In NIH3T3 cells, the presence of four miR-23b target sites in the luciferase mRNA resulted in a greater reduction in basal luciferase gene expression than the presence of two miR-23b target sites. The highly significant reduction in basal luciferase gene expression from pcLuc⁺-miR-23b-4T (** p=0.01, black) and pcLuc⁺-miR-23b-2T (* p=0.02, black) compared to pcLuc⁺, confirmed that suppression of gene expression by miR-23b was greater with increasing numbers of miR-23b target sites within the luciferase mRNA.

The Real-time qPCR data (section 6.3) showed that miR-23b expression was significantly downregulated in NIH3T3 cells stimulated with inflammatory cytokines. However, in Figure 6.5A, transfection of NIH3T3 cells with pcLuc⁺-miR-23b-4T resulted in the significant reduction in luciferase gene expression, irrespective of inflammatory stimulation when compared to gene expression induced from the target-less pcLuc⁺. Interestingly, the changes in miR-23b concentrations during inflammatory stimulation had a significant effect on luciferase gene expression from pcLuc⁺-miR-23b-4T when compared to its basal expression (* p=0.05 mIL-17A red, * p=0.02 TNFα red and *** p=0.0003 mIL-17A+TNFα red), demonstrating that

inflammation-responsive gene regulation was still observed in pcLuc$^+$-miR-23b-4T. However, inflammatory stimulation did not decrease miR-23b expression to a level where luciferase mRNA remained untargeted and consequently, the luciferase gene expression levels in stimulated cells remained low (Fig 6.5 A).

In contrast, when NIH3T3 cells were transfected with pcLuc$^+$-miR-23b-2T and stimulated with inflammatory cytokines, the changes in luciferase expression were not significantly different from that induced by the target-less pcLuc$^+$ construct (ns = $p>0.05$, black), indicating that the miR-23b concentration was significantly downregulated to a level where miR-23b had no significant effect on luciferase gene expression, which also corroborated the Real-time qPCR data (section 6.3). Furthermore, the significant changes in luciferase gene expression from pcLuc$^+$-miR-23b-2T during the basal state compared to that in the presence of TNFα (** $p=0.01$, red) and mIL-17+TNFα stimulation (* $p=0.03$, red) show that the inflammation-repressed expression of miR-23b in NIH3T3 cells permits high gene expression during inflammation and low gene expression during the uninflamed state, which is the desirable expression profile for miR-23b-mediated gene therapy (Fig 6.5 A).

In contrast, the absolute copy number of miR-23b in 293T cells is low (Brown *et al.,* 2007b), which corroborates the data in Figure 6.5B. In 293T cells, miR-23b generally had no significant effect on the luciferase mRNA expression, irrespective of the number of miR-23b target sites or inflammatory stimulation. These observations confirmed the cell-specific expression profile of miR-23b and also suggest that a threshold miR-23b concentration is required to achieve target mRNA downregulation in specific cells.

6.5.   Evaluating the inflammation-responsive miR-23b regulation of luciferase mRNA expressed from lentiviral constructs

To assist the potential *in vivo* application of miR-23b-mediated expression cassette, two or four miR-23b target sites were cloned into the 3'UTR of the luciferase gene within the lentiviral vector. The luciferase gene expression was driven by the constitutive SFFV promoter as this promoter is less susceptible to changes in activity due to inflammatory stimulation than the constitutive CMV promoter.

6.5.1.   Construction of lentiviral miR-23b-target constructs for regulation of stable luciferase gene expression

The schematic diagram below (Fig 6.6) illustrates the cloning strategy used to incorporate miR-23b-target sites within the lentiviral vector.

A.   The pLV.SFFV.Luc⁺ construct (from Chapter 4) was digested with EcoRI and KpnI, immediately downstream of the luciferase gene and the woodchuck post-transcriptional regulatory element (WPRE), to serve as the pLV.SFFV.Luc⁺ cloning vector for the miR-23b-2T and miR-23b-4T PCR products (Fig 6.6 A).
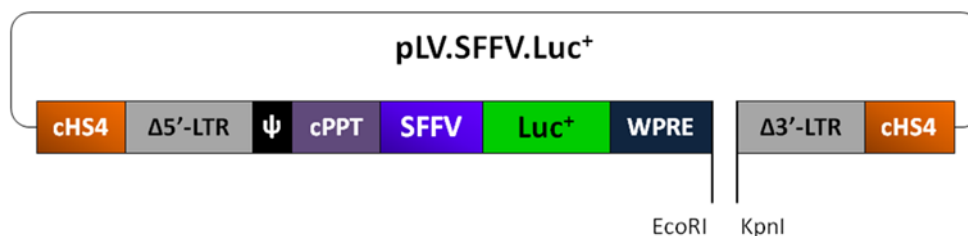
B. The miR-23b-target sequences were PCR amplified from pcLuc⁺-miR-23b-2T and pcLuc⁺-miR-23b-4T using forward and reverse PCR primers with MfeI and KpnI restriction sites, respectively (Fig 6.6 B).
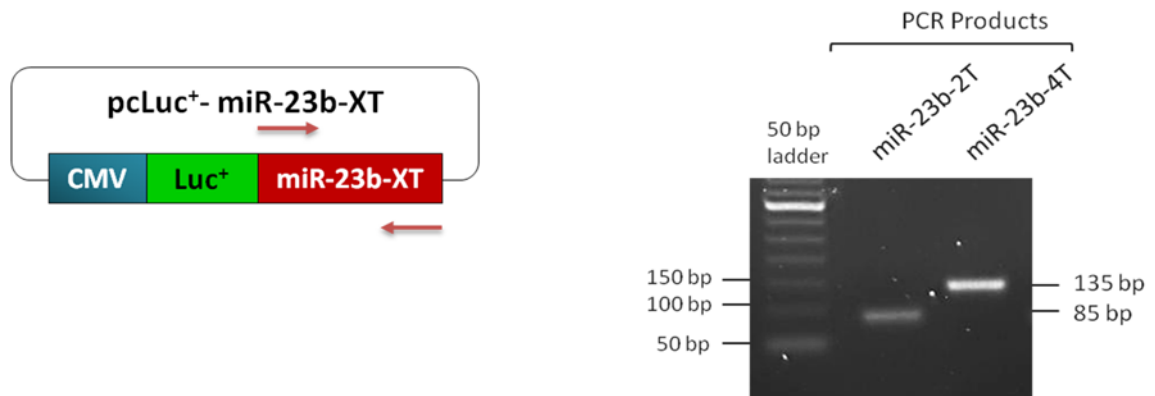
C. EcoRI and MfeI digested DNA have compatible ends therefore the miR-23b-target PCR products were digested with MfeI and KpnI and cloned into the EcoRI/KpnI site within the pLV.SFFV.Luc⁺ vector to generate pLV.SFFV.miR-23b-XT, where X is 2 or 4 miR-23b target sites (Fig 6.6 C). The resulting constructs were verified by DNA sequencing (Appendix 12.3 and 12.4).

D. Following the construction of the lentiviral miR-23b constructs, the lentiviral particles were produced by a three-plasmid transient transfection into 293T cells with the transfer, packaging and envelope plasmids to generate lentiviral particles pseudotyped with the VSV-G envelope glycoprotein. The VSV-G pseudotyped lentiviral particles were transduced into 293T and NIH3T3 cells to generate stable cell lines with the integrated luciferase gene containing miR-23b target sites.

Figure 6.6. Schematic diagram illustrating the cloning method used to generate the lentiviral miR-23b-target constructs. The lentiviral constructs express the luciferase gene under the control of

6.5.2. Synthetic miR-23b mimics significantly downregulate luciferase expression from lentivirally transduced 293T cells

Similar to the experiments in section 6.2, the synthetic miR-23b mimics were used to confirm whether the luciferase mRNA with miR-23b target sites, expressed from lentivirally transduced 293T cells, can be efficiently downregulated by a miR-23b mimic. The stable 293T cells were transfected with 1 µM miR-23b mimic or 1 µM miR-648 mimic (control) and the luciferase protein expression was quantified and presented in Figure 6.7.
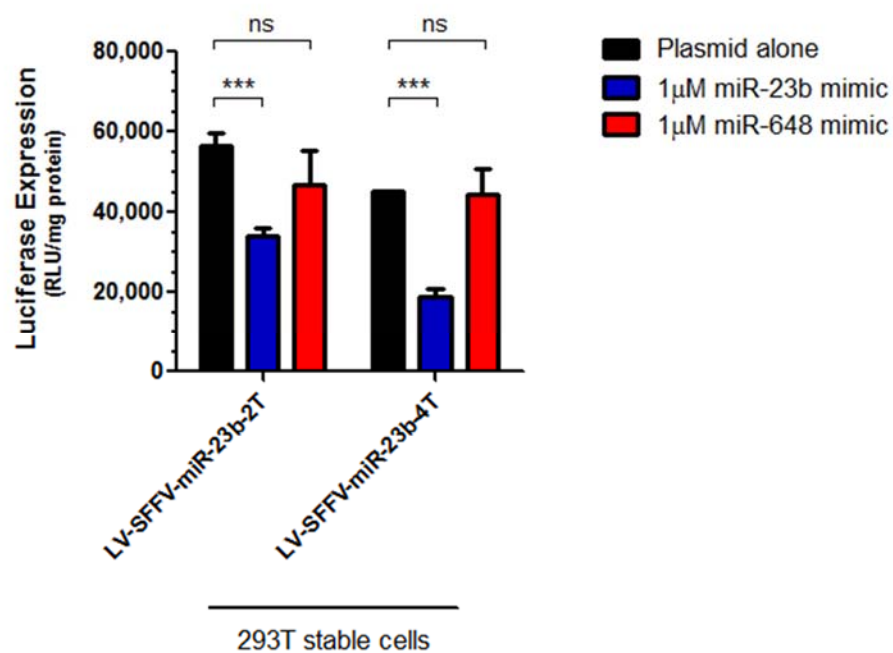
Figure 6.7. Significant downregulation of luciferase gene expression from stable 293T cells, using synthetic miR-23b mimics. 293T cells (50,000) were transduced with LV-SFFV-miR-23b-2T and LV-SFFV-miR-23b-4T lentiviral particles to generate stable 293T cells. The stable 293T cells (32,000 cells) were seeded in a 48-well plate. After 24 hours, 1 µM miR-23b mimic or 1 µM miR-648 mimic (control) were transfected into the stable 293T cells using Lipofectamine 2000 transfection reagent. Firefly luciferase was normalised to the protein content in the cell lysate (expressed as RLU/mg protein). The data represents the mean ± SD of triplicate values. The statistical significance compared to the corresponding 'plasmid alone' datasets for each group was calculated using the Student's t-test (ns = p>0.05, *** = p≤0.001).

Analogous to the results in section 6.2, the transfection of the miR-23b mimic resulted in a significant downregulation of luciferase gene expression from LV-SFFV-miR-23b-2T and LV-SFFV-miR-23b-4T stable 293T cells (both *** p=0.001). In contrast, the non-specific miR-648 control mimic had no-significant effect on the luciferase gene expression.

This experiment confirmed the ability of miR-23b to target luciferase mRNAs expressed from the lentiviral SFFV promoter in stable 293T cells. Therefore, the ability of endogenous inflammation-responsive miR-23b to regulate luciferase gene expression in lentivirally transduced stable NIH3T3 cells was evaluated in section 6.5.3.

6.5.3. Endogenous miR-23b activity exhibits miR-23b-target number dependent and miR-23b concentration dependent regulation of luciferase gene expression in stable NIH3T3 cells

To determine whether endogenous miR-23b can regulate luciferase mRNA expressed from lentivirally transduced stable NIH3T3 cells in an inflammation-responsive manner, the stable NIH3T3 cells were unstimulated, or stimulated with mIL-17A, TNFα or their combination. The luciferase protein was quantified and presented in Figure 6.8.
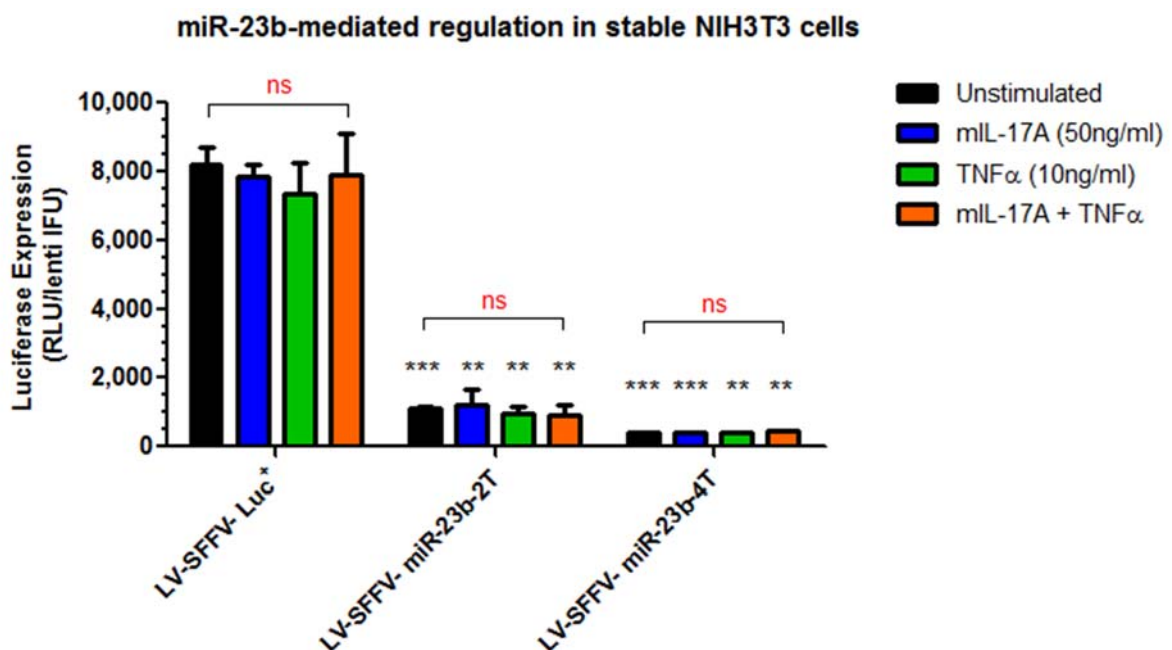
Figure 6.8. miR-23b target-number and concentration-dependent regulation of luciferase gene expression. NIH3T3 cells (50,000) were transduced with LV-SFFV-Luc+, LV-SFFV-miR-23b-2T or LV-SFFV-miR-23b-4T lentiviral particles to generate stable NIH3T3 cells. The stable NIH3T3 cells (20,000 cells) were seeded in a 96-well plate and after 24 hours, the cells were unstimulated or stimulated with mIL-17A, TNFα or their combination for 18 hours. The luciferase protein was quantified and normalised to the lentiviral titre of the respective constructs (RLU/lenti IFU). The data represents the mean ± SD of triplicate values and the statistical significance was calculated using the Student's t-test (ns = p>0.05, * = p≤0.05, ** = p≤0.01, *** = p≤0.001). The

Interestingly, basal luciferase expressions from LV-SFFV-miR-23b-2T and -4T cassettes were significantly lower than expression from LV-SFFV-Luc⁺ in lentivirally transduced NIH-3T3 cells (both *** p=0.001, black, respectively). This indicates that basal levels of miR-23b were sufficient to effectively target luciferase mRNA.

Importantly, the downregulation of miR-23b by inflammatory stimulation was not sufficient to reverse miR-23b-mediated suppression of luciferase expression from either cassette, as indicated by the absence of inflammation-regulated luciferase gene expression by miR-23b (ns= p>0.05, red). These results further corroborate that efficient regulation by miR-23b is dependent on the number of miR-23b target sites and also dependent on the concentration of miR-23b in the cell.

6.6 Inflammation-specific dual regulation of gene expression by an NFκB synthetic promoter and endogenous miR-23b in NIH3T3 cells

The combination of inflammation-responsive synthetic promoters and miRNA target sites within the same vector offers the prospect of stringently regulating gene expression induced by synthetic promoters with high basal gene expression.

The construct pCpG-6NFκB-Luc$^+$, which contains 6NFκB motifs, previously demonstrated high inducibility to TNFα in 293T cells (Chapter 3). Inflammation-regulated gene expression in NIH3T3 cells was only observed with mRNAs containing two miR-23b target sites (section 6.4) therefore, two miR-23b target sites were cloned into the 3'UTR of the luciferase gene, downstream of the synthetic promoter, in order to investigate whether miR-23b can reduce the basal gene expression without impairing the induced luciferase gene expression. Figure 6.9 (below) schematically depicts the cloning strategy used to generate the dual-regulated construct 6NFκB-miR-23b-2T.

A. pCpG-6NFκB-Luc$^+$ was digested with PpumI and FseI to isolate the NFκB-responsive synthetic promoter and the 5'- portion of the luciferase gene. The resulting construct served as the cloning vector 6NFκB-Luc$^+$ (Fig 6.9 A).
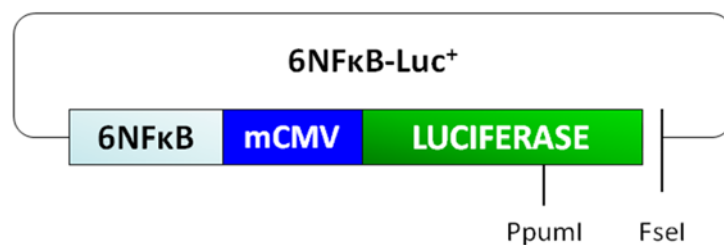
B. The 3'-portion of the luciferase gene and two miR-23b target sites were PCR amplified from the pcLuc⁺-miR-23b-2T construct using a forward PCR primer and a reverse PCR primer that contained an FseI restriction enzyme site (Fig 6.9 B).
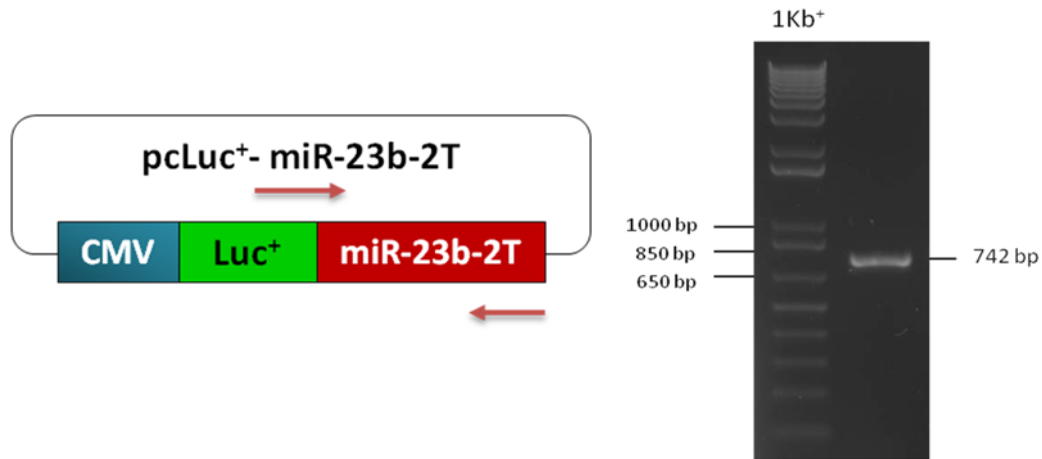
C. The PCR product was digested with PpumI and FseI and cloned into the equivalent site within the 6NFκB-Luc⁺ vector to restore the luciferase gene and generate 6NFκB-miR-23b-2T (Fig 6.9 C). The construct was confirmed by DNA sequencing (Appendix 12.5).
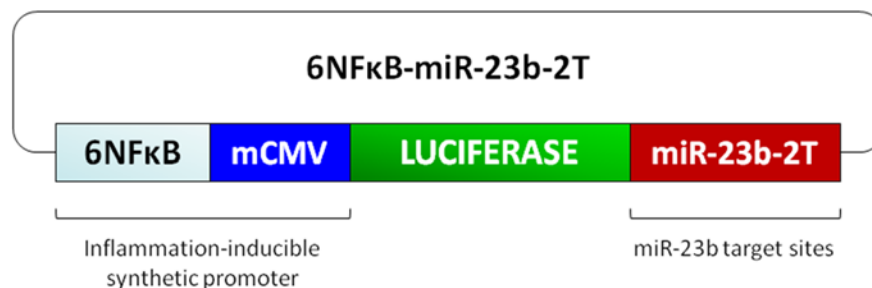
Figure 6.9. Schematic diagram depicting a dual-regulated construct with an inflammation-inducible synthetic promoter and two miR-23b target sites. The pCpG-6NFκB-Luc⁺ construct was modified to incorporate two miR-23b target sites within the 3'UTR of the luciferase gene, downstream of an inflammation-inducible promoter.

The 6NFκB-Luc⁺ and 6NFκB-miR-23b-2T constructs were co-transfected with pRL-CMV into NIH3T3 cells, which were then unstimulated or stimulated with TNFα. The luciferase gene expression is presented in Figure 6.10.
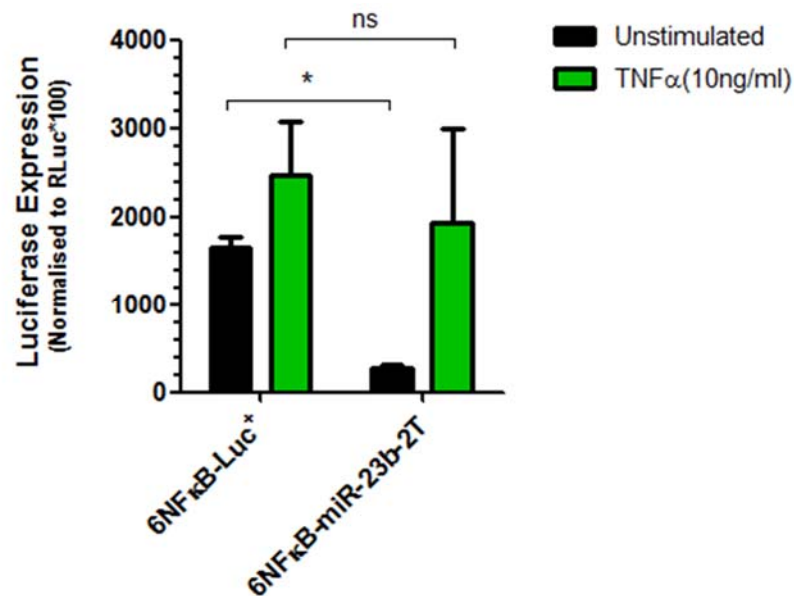
Figure 6.10. Regulated luciferase expression from a transcriptionally and post-transcriptionally responsive vector in NIH3T3 cells. The 6NFκB-Luc⁺ and 6NFκB-miR-23b-2T constructs (180 ng) were co-transfected with pRL-CMV (20 ng) into NIH3T3 cells (20,000 cells in a 96-well plate) and were either unstimulated or stimulated with TNFα (10 ng/ml) for 18 hours. Firefly luciferase was normalised to renilla luciferase and the data represents the mean ± SD of triplicate values. The statistical significance between the corresponding datasets was calculated using the Student's t-test (ns= p>0.05, * = p≤0.05).

Endogenous miR-23b significantly downregulated the basal luciferase gene expression from 6NFκB-miR-23b-2T (* p=0.03, 84% reduction) when compared to luciferase expression from 6NFκB-Luc⁺. Importantly, luciferase expression was induced by TNFα stimulation from both vectors. However, due to the reduced basal expression from 6NFκB-miR-23b-2T, the fold change following TNFα-stimulation was 7.1 fold compared to 1.5 fold for the unmodified vector, which represents a dramatic difference in the regulation of gene expression by the two constructs. Overall, Figure 6.10 confirmed the concept of exploiting the endogenous

transcriptional and post-transcriptional components to achieve dual-regulation of gene expression in an inflammation-responsive manner.

6.7. Discussion

The intra-articular synthesis of the therapeutic proteins expressed from constitutive promoters can potentially achieve sustained, local suppression of inflammation (Adriaansen *et al.,* 2006). However, the constitutive promoters i.e. CMV and SFFV used in these systems induce continuous and unregulated gene expression which can lead to adverse effects in non-target tissues and/or disease states. The clinical course of RA is characterised by episodes of spontaneous remission and relapsing chronic inflammation, therefore, the prospect of inflammation-regulated gene expression systems offers the potential of permitting high therapeutic gene expression only during inflammation thus improving the safety and efficacy of local RA gene therapy. In this chapter, regulated gene expression induced by inflammatory stimuli was achieved by miRNA-mediated post-transcriptional regulation.

As reviewed by Gentner and Naldini, (2012), the distinct expression profiles of numerous miRNAs have been exploited to achieve cell-, tissue- and disease-specific gene expression. However, exploiting the endogenous activities of inflammation-repressed miRNAs to achieve disease-specific gene expression is an unexplored concept and in this chapter, I explored the use of miR-23b to develop an inflammation-responsive gene expression system.

A crucial element of miRNA-mediated, inflammation-responsive gene regulation was the choice of miRNA. The candidate miRNA was required to be highly expressed during the uninflamed state and significantly downregulated in inflammation and following extensive literature searches, miR-23b was identified to display the desired expression profile by Zhu *et al.,* (2012), who used miRNA microarray and relative Real-time qPCR validation. This observation was confirmed by absolute Real-time qPCR in section 6.3 which verified that miR-23b was highly expressed in the absence of inflammatory stimulation and significantly downregulated in NIH3T3 mouse embryonic fibroblasts stimulated with inflammatory

cytokines. Unlike the relative quantification method used by Zhu *et al.,* (2012), the quantification of absolute miR-23b concentration in NIH3T3 cells facilitated logical interpretations of the miR-23b-mediated luciferase gene regulation data, which highlighted the important relationship between miR-23b concentration and efficient gene regulation.

Using the basic principle of miRNA-mediated regulation, two and four perfectly complementary miR-23b target sites were cloned into the 3'UTR of the luciferase gene to subject the luciferase mRNA to inflammation-responsive regulation by miR-23b. The binding of miR-23b to the luciferase 3'UTR was investigated by target validation experiments (section 6.2) which confirmed the successful downregulation of the luciferase gene using synthetic miR-23b mimics.

The importance of quantifying the absolute expression of the candidate miRNAs was highlighted by Brown *et al.,* (2007b) who demonstrated that the target mRNA suppression is dependent on a threshold miRNA concentration. This group quantified the absolute expression of several miRNAs in 293T (kidney) and U937 (monocyte) cells using absolute Real-time qPCR and showed that miRNAs expressed at low levels (<100 copies/pg small RNA) do not significantly downregulate target mRNAs whereas highly expressed miRNAs (>1000 copies/pg small RNA) significantly repressed target mRNAs. Interestingly, they quantified miR-23b to have a relatively low concentration in unstimulated 293T cells (~400 copies/ pg small RNA) which provided a logical explanation for the lack of miR-23b-mediated regulation that was observed in 293T cells (Fig 6.5 B). In addition to the relatively low miR-23b concentrations in 293T cells, it is also possible that the number of miR-23b target sites may have exceeded the miR-23b concentration which resulted in poor miR-23b-mediated regulation of gene expression. Overexpression of miR-23b-target bearing luciferase mRNAs may have resulted from the amplification of the pcLuc$^+$-miR-23b expression constructs (derived from pcDNA3) which contain an SV40 origin of replication (ori). Transient transfection of these constructs into 293T cells results in the episomal replication of the vector due to the presence of the SV40 Large T-antigen in 293T cells (Prelich *et al.,* 1987). Therefore, the

combined effect of high constitutive luciferase expression driven by the CMV promoter, amplification of the expression vector in 293T cells and the low miR-23b concentration in 293T cells resulted in the ratio of low miR-23b concentration relative to the high expression of target luciferase mRNAs which may have impaired miR-23b-mediated regulation in 293T cells (Fig 6.5 B).

In contrast, the luciferase gene expression from lentivirally transduced stable NIH3T3 cells was significantly downregulated, irrespective of inflammatory stimulation (Fig 6.8). It is possible that the low transduction efficiency of NIH3T3 cells resulted in low luciferase mRNA levels which promoted a ratio of high miR-23b concentration relative to the low expression of target luciferase mRNA in stable NIH3T3 cells, ultimately resulting in strong suppression of luciferase gene expression. This data also suggested that the highly abundant miR-23b concentration in NIH3T3 cells was not downregulated by inflammatory-cytokine stimulation to an extent which prevented luciferase mRNA targeting, thereby further strengthening the concept that target mRNA suppression is dependent on a threshold miRNA concentration (Brown *et al.,* 2007b).

In the case of optimal miR-23b concentrations relative to the target mRNA levels, efficient regulation of gene expression can be achieved, as shown in Figure 6.5 A. In this experiment, the luciferase mRNAs containing two miR-23b target sites expressed in transiently transfected NIH3T3 cells were efficiently regulated in an inflammation-responsive manner. However, although inflammatory-regulation was observed with luciferase mRNAs containing four miR-23b target sites, the overall luciferase expression was significantly lower than that observed with luciferase mRNAs containing two miR-23b target sites, which supported the finding that target mRNA suppression is also dependent on the number of miRNA target sites (Doench *et al.,* 2003). This observation was also confirmed by the miR-23b mimic target validation experiments (Fig 6.3 and 6.7) which clearly demonstrated that luciferase mRNA suppression was greatest with four miR-23b target sites compared to two miR-23b target sites within the luciferase mRNA. Similarly, numerous publications have also implemented the findings of

Doench *et al.,* (2003) to achieve greater target suppression of gene expression by incorporating four miRNA target sites.

Encouraged by the inflammation-responsive regulation of luciferase gene expression in NIH3T3 cells by miR-23b, a dual-regulated expression system was generated by cloning two perfectly complementary miR-23b target sites into the 3'UTR of the luciferase gene, positioned downstream of an NFκB-responsive synthetic promoter (section 6.6). In this experiment, the ratio of miR-23b concentration relative to the luciferase mRNA levels were optimal which resulted in the significant downregulation of basal luciferase expression in unstimulated NIH3T3 cells by miR-23b, without significantly impairing the TNFα-induced luciferase gene expression. These results provided proof-of-concept of combining transcriptional and post-transcriptional regulatory components to achieve stringent inflammation-responsive gene expression. This approach also provided a novel strategy to improve the safety profile of synthetic promoters with high basal gene expression in cells with inflammation-repressed expression of miR-23b, at the threshold concentration.

Evidently, a fine balance between miR-23b concentrations and target mRNA levels is required for inflammation-regulated gene expression by miR-23b and profiling the absolute expression of miR-23b concentration within the cells served as a reliable predictor of efficient miR-23b-mediated gene regulation. Table 6.1 summarises the luciferase gene regulation by endogenous miR-23b in transiently transfected 293T and NIH3T3 cells and lentivirally transduced stable NIH3T3 cells.

[A]

Cell type	miR-23b concentration (Unstimulated state)	Luciferase target mRNA expression	Luciferase gene regulation
Transient transfected 293T cells	Low	High (miR-23b-2T)	No significant effect
Transient transfected 293T cells	Low	Very high (miR-23b-4T)	No significant effect
Transient transfected NIH3T3 cells	**High**	**Moderate (miR-23b-2T)**	**Significant downregulation**
Transient transfected NIH3T3 cells	High	Moderate/High (miR-23b-4T)	Significant downregulation
Stable NIH3T3 cells	High	Low (miR-23b-2T and 4T)	Significant downregulation
Transient transfection NIH3T3 cells Dual-regulated promoter	**High**	**Moderate (miR-23b-2T)**	**Significant downregulation**

[B]

Cell type	miR-23b concentration (Inflammatory stimulation)	Luciferase target mRNA expression	Luciferase gene regulation
Transient transfected 293T cells	Low	High (miR-23b-2T)	No significant effect
Transient transfected 293T cells	Low	Very high (miR-23b-4T)	No significant effect
Transient transfected NIH3T3 cells	**Significantly reduced**	**Moderate (miR-23b-2T)**	**No significant effect**
Transient transfected NIH3T3 cells	Significantly reduced	Moderate/High (miR-23b-4T)	Significant downregulation
Stable NIH3T3 cells	Significantly reduced	Low (miR-23b-2T and 4T)	Significant downregulation
Transient transfection NIH3T3 cells Dual-regulated promoter	**Significantly reduced**	**Moderate (miR-23b-2T)**	**No significant effect**

Table 6.1. Efficient luciferase gene regulation is dependent on the ratio between the level of luciferase target mRNA and the threshold miR-23b concentration within the cell. The differences in luciferase gene regulation by miR-23b during the unstimulated [A] and inflammatory-stimulated [B] states are shown.

A potential concern regarding the high expression of mRNAs with four miR-23b target sites is the inadvertent competition introduced between synthetic miR-23b (luciferase) target mRNAs and the endogenous miR-23b target mRNAs. Endogenous miR-23b suppresses IL-17, TNFα and IL-1β-induced NFκB activation by targeting the 3'UTR of mRNA encoding TAK1-binding protein 2 (TAB2), TAB3 and the IκB-kinase α (IKKα), resulting in repressed inflammation in autoimmune diseases such as RA and SLE (Zhu *et al.,* 2012). Theoretically, the overexpression of synthetic miR-23b target mRNA (luciferase mRNA) can potentially disrupt the miR-23b-mediated targeting of endogenous mRNAs by acting as a 'decoy' target, which may result in the activation of the NFκB pathway and perpetuate inflammation by untargeted TAB2, TAB3 and IKKα.

However, it is unlikely that the recruitment of miR-23b to the target sites within the luciferase mRNA would result in loss of regulation of endogenous miR-23b targets because firstly, the concentration of miR-23b is highly abundant in NIH3T3 cells and greatly exceeds the luciferase target mRNA expression. Secondly, endogenous mRNA targets often have several targeting miRNAs due to the similarities with their 5'-miRNA seed sequences (Brennecke *et al.,* 2005) and therefore the endogenous miR-23b targets, i.e. TAB2, TAB3 and IKKα can be regulated by other targeting miRNAs. Thirdly, the incorporation of perfectly complementary miR-23b target sites increases the catalytic rate of the RISC complex which decreases the likelihood that the endogenous miRNAs become saturated. In agreement with these assumptions, the results of Brown *et al.,* (2007b) experimentally demonstrated that the high expression of perfectly complementary miRNA target sites does not disturb the regulation of endogenous target mRNAs.

Logically, the cell-state specific expression profile of miR-23b limits the broad applicability of miR-23b-mediated regulation systems to specific cell types and tissues. Encouragingly, miR-23b expression is significantly downregulated in primary RA fibroblast-like synoviocytes (Zhu *et al.,* 2012), which are the main target cells for the disease-regulated local gene therapy described in this thesis. Therefore, the miR-23b-mediated expression systems described in this chapter can potentially promote inflammation-regulated therapeutic gene expression in RA synovial fibroblasts, provided that these cells express sufficient miR-23b concentrations relative to the target therapeutic mRNA expression. Consequently, the unwanted high therapeutic gene expression driven by a constitutive promoter or synthetic promoter can potentially be downregulated by highly expressed miR-23b during the uninflamed state, whilst permitting high therapeutic gene expression during inflammation in RA joints.

Chapter 7: General Discussion and Conclusion

Rheumatoid arthritis (RA) is characterised by elevated pro-inflammatory cytokines in the synovial joint, i.e. TNFα and IL-1β, which orchestrate many of the inflammatory and destructive processes in the disease and are therefore attractive therapeutic targets (McInnes and Schett, 2011). At the molecular level, the inflammatory and hypoxic pathological environment in the RA joint activates responsive transcription factors which control the gene expression of pro-inflammatory cytokines, chemokines, growth factors, cartilage degrading enzymes and other mediators integral to the pathogenesis of RA (Okamoto *et al.,* 2008).

Such advances in the understanding of the pathophysiology of RA have enabled researchers to exploit the endogenous activity of activated transcription factors using inflammation-inducible synthetic promoters to regulate therapeutic gene expression in response to inflammation (Khoury *et al.,* 2007; Geurts *et al.,* 2007; Henningson *et al.,* 2012 and Garaulet *et al.,* 2013). These self-regulating promoters can potentially induce high levels of anti-inflammatory agents during disease flare and lower levels during disease remission, to mirror the clinical course and severity of inflammation within the joint (Adriaansen *et al.,* 2006). Gene therapy strategies utilising inflammation-inducible synthetic promoters for the treatment of arthritis in experimental animal models of RA have generally been successful (e.g. Geurts *et al.,* 2007 and Henningson *et al.,* 2012). However, the development and application of multi-responsive and synergistically-inducible composite synthetic promoters for RA gene therapy has yet to be demonstrated by others and this was the primary objective of my PhD.

This thesis details the rationale design, construction and functional characterisation of inflammation-inducible composite synthetic promoters for local RA gene therapy (Chapters 3-5). Engineering of composite synthetic promoters to contain randomly arranged core binding sites of AP-1, HIF-1α, NFκB, and also C/EBPβ, Egr-1 and Ets-1 represented a means to exploit the simultaneous binding of the candidate TFs to potentially provide multi-responsive and additive/synergistic induction of gene expression in response to the multiple pathological

stimuli in the RA joint, in contrast to the single-responsive synthetic promoters such as those used by Shibata *et al.,* (2000) and Khoury *et al.,* (2007), which are restricted to hypoxia and inflammatory stimulation, respectively.

Collectively, the *in vitro* functional characterisation of the composite promoters from different libraries revealed that changes in promoter architecture can have profound effects on gene expression. Interestingly, the impaired induction of transcriptional synergism by the composite promoters with a 4bp space between the TFBSs, generated by the random ligation cloning method, was suggestive of potential steric hindrance of TFs during combined inflammatory and hypoxic stimulation in transiently transfected 293T cells. Logical interpretations of these results were provided by the established principles of TF-DNA interactions, indicating that sliding TFs in search of their respective binding sites can dislodge DNA-bound TFs (Berg *et al.,* 1981; Halford and Marko, 2004; Hu *et al.,* 2008), which potentially impairs synergistic induction. Also, it is possible that many TFBSs were rendered inaccessible by adjacently bound TFs, due to the close proximity of the TFBSs, thereby preventing simultaneous TF binding which hindered the potential for synergistic gene expression (section 3.2.2 and 3.4).

Subsequently, the spatial arrangements of TFBSs were reorganised using an optimised Assembly PCR cloning method which provided a comprehensive tool to engineer various composite promoter libraries containing randomly arranged yet evenly spaced TFBSs in high-throughput PCR reactions. Functional characterisation of these synthetic promoters revealed that increasing the spacing between the TFBSs and also between the proximal TFBS and the TATA box resulted in decreased basal and induced luciferase gene expression levels. Supporting bioinformatic evidence by Gotea *et al.,* (2010) reported that active endogenous promoters typically possess clusters of the same type of TFBS. These homotypic clusters are likely to enhance TF recruitment and the efficiency of the local search process (Brackley *et al.,* 2012) which can increase gene expression. Also, many studies have shown that expanding the distance between upstream elements i.e. TFBSs and the TATA box can reduce

transcription (Guarente and Hoar, 1984; McKnight, 1982; Takahashi *et al.,* 1986, Wu and Berk, 1988; Smith *et al.,* 1995; Dobi and Winston, 2007; Sharon *et al.,* 2012).

Without prior knowledge on the significance of computational biology in the design and functionality of synthetic promoters, the results discussed in Chapter 3 have demonstrated that studies from the field of synthetic biology are useful in interpreting the data. The strongest evidence supporting the majority of the results from Chapter 3 comes from the studies of Sharon *et al.,* (2012). This group integrated computational biology to design a library of 6500 promoters and also devised a high-throughput microarray-based method to measure the effect of systematic changes to TFBS location, number, affinity and organisation, when positioned upstream of the TATA box and the yellow fluorescence protein (YFP) reporter gene. Strikingly, their data showed that gene expression was reduced when TFBSs (i.e. Gal4 and Gcn4 sites) were positioned further away from the core promoter, gene expression increased with increasing number of TFBSs (upto 3-4 sites) and then saturated and the close proximity of adjacent Gal4 binding sites (1bp space) resulted in steric occlusion of Gal4 molecules, which was also weakly demonstrated by Gcn4 molecules binding to Gcn4 sites separated by 5bp space. Taken together, the data of Sharon *et al.,* (2012) strongly supports many of the observed trends discussed in Chapter 3 and also highlights the advantages of integrating computational biology to rationally design synthetic promoters, as also demonstrated by Geurts *et al.,* (2009).

As anticipated, the vast majority of the composite synthetic promoters with the optimised 20bp space between the TFBSs and 66bp space between the proximal TFBS and the TATA box, exhibited low basal, multi-responsive and additively or synergistically-inducible gene expression (section 3.5), which was suggestive of alleviated steric hindrance of TFs. Although the sought-after expression profiles was observed in most of the optimised composite promoters, an inherent limitation to the screening method described in this thesis is that the identity and number of TFs binding to sites within the synthetic promoter has not been confirmed. An experimental technique such as chromatin immunoprecipitation (ChIP) would

have been useful in identifying the specific binding locations of the TFs, which might have provided more insight into the number of TFBSs needed for the desired level and pattern of gene expression. Although the in-depth investigation of TF binding was not a proposed research question, it is an interesting and important option for further research, where analysis of TF-TFBS interactions within the composite promoters could be useful in the rationale design of the synthetic promoters for optimal activity.

Promoters 2, 9 and 12, displaying the best expression profiles, were cloned into a SIN lentiviral vector to confer long-term and stable gene expression *in vitro* and *in vivo* and their kinetics of gene expression were assessed in lentiviral-transduced stable 293T cells. The lentiviral promoters LV-2, LV-9 and LV-12 were highly sensitive and rapidly activated and generally exhibited dose- and time-dependent increases in luciferase gene expression in response to increasing concentrations and durations of exposure to TNFα and PMA stimulation, respectively. Importantly, the lentiviral promoters had dramatically improved induction levels and retained their favourable expression profiles (section 4.2.2). This may have been due to the presence of performance enhancing elements i.e. WPRE within the lentiviral vector or the presence of high copy numbers of integrated proviral DNA in the stable 293T cells, where the latter could have been confirmed by qPCR, had time permitted.

Encouragingly, the local delivery of LV-2, LV-9 and LV-12 in the mouse paw resulted in disease-specific induction of luciferase gene expression following carrageenan-induced paw inflammation, as determined by real-time bioluminescence imaging. Although only LV-2 induction was statistically significant (at 3 hours post-inflammation), both LV-2 and LV-9 induced robust gene expression levels which were localised to the inflamed paw and comparable to gene expression induced by the constitutive LV-SFFV promoter (section 4.3).

Due to the consistently favourable performance of promoter LV-2, this promoter was used to regulate the expression of mTNFR-Fc and IL-1Ra therapeutic genes in the carrageenan-induced paw inflammation mouse model (section 5.3). In accordance with the principle of IL-

1Ra therapy, the constitutive SFFV promoter expressed IL-1Ra before the onset of inflammation which was able to occupy the IL-1 receptors (McIntyre *et al.,* 1991; Bakker *et al.,* 1997) and resulted in significantly inhibited paw inflammation. However, mice treated with LV-2-driven therapeutic genes did not exhibit reduced paw inflammation. Furthermore, none of the vectors encoding mTNFRII-Fc inhibited inflammation in the model which may have been due to the delivery of suboptimal lentiviral titres resulting in insufficient expression of therapeutic proteins prior to and during inflammation. The lentiviral titre used in this experiment (830,000 IFU) was substantially lower than the lentiviral titre used by similar studies such as Garaulet *et al.,* (2013) who observed reduced paw inflammation following the local delivery of IL-10 encoding lentivectors (2×10^7 transducing units) expressed from an inflammation-inducible synthetic promoter. Therefore, the optimisation of lentiviral titres used in the experiments described in section 5.3 represents an important direction for future work.

A further objective of my PhD was to exploit the differential expression of miRNAs during inflammation to post-transcriptionally regulate gene expression induced by a constitutive promoter (Chapter 6). Absolute real-time qPCR was used to confirm the inflammation-repressed expression of miR-23b in mouse embryonic fibroblasts, as previously identified by Zhu *et al.,* (2012). The fundamental principles of microRNA-mRNA targeting were implemented to construct expression cassettes with two or four miR-23b target sites in the 3'UTR of luciferase gene to subject the CMV-driven luciferase gene expression to regulation by miR-23b during inflammation. In agreement with the hypothesis, endogenous miR-23b significantly downregulated luciferase gene expression in unstimulated NIH3T3 cells (where miR-23b is highly expressed) and did not impair the high gene expression in NIH3T3 cells treated with inflammatory stimuli (where miR-23b is downregulated). Importantly, the efficiency of miR-23b-mediated gene regulation was highly dependent on the threshold   miR-23b concentration in the cells, the level of target mRNA expression and also the number of miR-23b target sites in the mRNA 3'UTR, as confirmed by the pioneering studies of Brown *et al.* (2007b). Furthermore, the incorporation of two miR-23b-target sites in the 3'UTR of the

luciferase gene controlled by an NFκB-responsive promoter resulted in significantly reduced basal gene expression in unstimulated NIH3T3 cells and consistently high gene expression following TNFα stimulation. The coupling of an inflammation-inducible synthetic promoter with the target sites of an inflammation-repressed miRNA in a dual-regulated cassette represented a novel approach to improving the efficiency and safety of 'leaky' synthetic promoters by reducing unnecessarily high basal gene expression without impairing the induced gene expression.

Encouragingly, Zhu *et al.,* (2012) reported the downregulation of miR-23b expression in the synovial tissues from RA patients and CIA mice, in the kidney tissues from SLE (lupus) patients and the related MRL/*lpr* mice and in the spinal cords from EAE mice (multiple sclerosis mouse model) compared to healthy controls. Therefore, provided that miR-23b is expressed at the threshold concentrations in the target tissues, these miR-23b-regulated expression cassettes can be applicable for gene therapy to treat RA, lupus and multiple sclerosis.

To conclude, the results in Chapters 3-5 have confirmed the development of a novel approach to construct transcriptionally-regulated inflammation-responsive composite synthetic promoters. The promoters exhibited differential expression, multi-responsiveness and synergistic induction *in vitro* and disease-specific induction of luciferase gene expression *in vivo*. However, the general lack of therapeutic efficacy by the candidate promoter LV-2 (and also LV-SFFV) may have been a result of suboptimal lentiviral titres. Nevertheless, the ubiquitous activation of TFs during inflammatory conditions offers the prospect of utilising these versatile composite promoters for gene therapy applications in other diseases characterised by local changes in transcription factor activation, which is currently underway (Professor Adrian Hobbs, William Harvey Heart Centre, UK). Furthermore, the results in Chapter 6 have proven the concept of post-transcriptional regulation of constitutive (CMV-driven) and inducible (NFκB-driven) gene expression by the inflammation-repressed miR-23b,

which is a promising approach for a combined transcriptionally and post-transcriptionally responsive expression system for future RA gene therapy.

Chapter 8: Further Work

The results in this thesis reveal the potential for further work in a few important areas. Further work would entail:

- Investigating the binding interactions of the candidate TFs to their cognate binding sites within the composite promoters, using chromatin immunoprecipitation

The level of gene expression from synthetic promoters is the functional readout of TF binding, however, these simple reporter gene assays provide no information regarding the specific TF binding location and how many TFs bind to the TFBSs within the synthetic promoter. A chromatin immunoprecipitation assay will be used to investigate TF-TFBS interactions to provide more insight into the pattern of TF binding during inflammatory and/or hypoxic conditions.

- Optimising the lentiviral titre to achieve therapeutic effects *in vivo*

Lentiviral titres will be optimised in an experimental model of RA i.e. CIA or AIA model, which would facilitate the monitoring of promoter activity for a longer duration, with the option of re-inducing inflammation to fully characterise the promoter activity *in vivo.* When repeating this experiment, an additional control group treated with either TNFRII or IL-1Ra proteins will be included for comparison. Also, biodistribution analyses using qPCR will be performed to determine the distribution of the transgene to the target site (within the paw) and to non-target sites (e.g. liver), which should provide insight into the safety and efficacy of the delivered therapy.

- Construct composite promoters with higher basal and induced gene expression than promoter LV-2

A higher basal and induced expression of the composite promoters might be a feature that could overcome the challenges in achieving therapeutic effects *in vivo*. This is particularly relevant to the principles of targeted IL-1 and TNFα therapies, which require high IL-1Ra and mTNFRII-Fc expression before and during inflammation. Using the knowledge obtained from the experiments in Chapter 3 (sections 3.3.5 and 3.3.7), the TFBSs can be positioned closer together or positioned closer to the TATA box within the composite promoters to achieve higher basal and higher induced gene expression levels than demonstrated by promoter LV-2. Functional characterisation of the promoters will be required to confirm multi-responsiveness and synergistic induction of gene expression.

- Identify the methylation status of the promoters

Promoters are often silenced by DNA methylation, therefore it would be interesting to perform DNA methylation analysis studies to detect the methylation status of the promoters used *in vivo* to determine whether they are prone to silencing.

- Quantify the changes in luciferase mRNA expression in response to the changes in miR-23b expression, during inflammation

Perfect complementarity between miRNAs and the target sites in the mRNA 3'UTR results in mRNA cleavage (Bartel, 2004). Therefore, it is feasible to assume that the luciferase mRNA (containing perfectly complementary miR-23b target sites) was degraded following miR-23b binding, as indicated by the downregulation in luciferase protein expression. However, it would be useful to quantify the changes in luciferase mRNA in response to the changes in miR-23b expression during inflammation using Real-time qPCR, to determine the threshold luciferase mRNA levels for efficient regulation by miR-23b.

Potential future work of interest would be to re-design the composite promoters to incorporate additional regulatory components which could improve the efficacy and safety of the gene therapy. For example;

- Incorporate drug-inducible components into the composite promoter to achieve pharmacological- and disease-regulated gene expression

A general limitation of inflammation-inducible synthetic promoters is that they respond indiscriminately to any inflammatory environment and can therefore induce gene expression in response to inflammation which is not associated with the disease. By combining drug-inducible and inflammation-inducible regulatory elements in the same construct, gene expression can be magnified and more importantly, completely switched off in the case of RA-unrelated inflammation. This strategy was demonstrated by Gould's lab who constructed a novel hybrid hypoxia-responsive and Dox-regulated transcriptional system (Subang *et al.,* (2012) which can be applied to the composite promoters and miRNA-regulated constructs described in this thesis to generate stringently regulated gene therapy cassettes.

Chapter 9: References

1. Adriaansen, J., Vervoordeldonk, M. J., & Tak, P. P. (2006). Gene therapy as a therapeutic approach for the treatment of rheumatoid arthritis: innovative vectors and therapeutic genes. *Rheumatology (Oxford), 45*(6), 656-668.

2. Ahlqvist, J. (1984). A hypothesis on the pathogenesis of rheumatoid and other non-specific synovitides. IV A. The possible intermediate role of local hypoxia and metabolic alterations. *Med Hypotheses, 13*(3), 257-302.

3. Aicher, W. K., Dinkel, A., Grimbacher, B., Haas, C., Seydlitz-Kurzbach, E. V., Peter, H. H., et al. (1999). Serum response elements activate and cAMP responsive elements inhibit expression of transcription factor Egr-1 in synovial fibroblasts of rheumatoid arthritis patients. *Int Immunol, 11*(1), 47-61.

4. Aicher, W. K., Heer, A. H., Trabandt, A., Bridges, S. L., Jr., Schroeder, H. W., Jr., Stransky, G., et al. (1994). Overexpression of zinc-finger transcription factor Z-225/Egr-1 in synoviocytes from rheumatoid arthritis patients. *J Immunol, 152*(12), 5940-5948.

5. Aicher, W. K., Sakamoto, K. M., Hack, A., & Eibel, H. (1999a). Analysis of functional elements in the human Egr-1 gene promoter. *Rheumatol Int, 18*(5-6), 207-214.

6. Aikawa, Y., Morimoto, K., Yamamoto, T., Chaki, H., Hashiramoto, A., Narita, H., et al. (2008). Treatment of arthritis with a selective inhibitor of c-Fos/activator protein-1. *Nat Biotechnol, 26*(7), 817-823.

7. Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., et al. (1990). A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *EMBO J, 9*(6), 1897-1906.

8. Alamanos, Y., Voulgari, P. V., & Drosos, A. A. (2006). Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review. *Semin Arthritis Rheum, 36*(3), 182-188.

9. Alsaleh, G., Suffert, G., Semaan, N., Juncker, T., Frenzel, L., Gottenberg, J. E., et al. (2009). Bruton's tyrosine kinase is involved in miR-346-related regulation of IL-18 release by lipopolysaccharide-activated rheumatoid fibroblast-like synoviocytes. *J Immunol, 182*(8), 5088-5097.

10. Amarzguioui, M., Rossi, J. J., & Kim, D. (2005). Approaches for chemically synthesized siRNA and vector-mediated RNAi. *FEBS Lett, 579*(26), 5974-5981.

11. Amaya-Amaya, J., Sarmiento-Monroy, J. C., Mantilla, R. D., Pineda-Tamayo, R., Rojas-Villarraga, A., & Anaya, J. M. (2013). Novel risk factors for cardiovascular

disease in rheumatoid arthritis. *Immunol Res, 56*(2-3), 267-286.

12. Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T., & Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. *Curr Biol, 13*(10), 807-818.

13. Andersen, C. L., Jensen, J. L., & Orntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res, 64*(15), 5245-5250.

14. Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., et al. (1987). Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. *Cell, 49*(6), 729-739.

15. Apparailly, F., Millet, V., Noel, D., Jacquet, C., Sany, J., & Jorgensen, C. (2002). Tetracycline-inducible interleukin-10 gene transfer mediated by an adeno-associated virus: application to experimental arthritis. *Hum Gene Ther, 13*(10), 1179-1188.

16. Aravin, A. A., Hannon, G. J., & Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science, 318*(5851), 761-764.

17. Arend, W. P. (2002). The balance between IL-1 and IL-1Ra in disease. *Cytokine Growth Factor Rev, 13*(4-5), 323-340.

18. Arend, W. P., Welgus, H. G., Thompson, R. C., & Eisenberg, S. P. (1990). Biological properties of recombinant human monocyte-derived interleukin 1 receptor antagonist. *J Clin Invest, 85*(5), 1694-1697.

19. Arnett, F. C., Edworthy, S. M., Bloch, D. A., McShane, D. J., Fries, J. F., Cooper, N. S., et al. (1988). The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum, 31*(3), 315-324.

20. Asahara, H., Fujisawa, K., Kobata, T., Hasunuma, T., Maeda, T., Asanuma, M., et al. (1997). Direct evidence of high DNA binding activity of transcription factor AP-1 in rheumatoid arthritis synovium. *Arthritis Rheum, 40*(5), 912-918.

21. Asquith, D. L., Miller, A. M., McInnes, I. B., & Liew, F. Y. (2009). Animal models of rheumatoid arthritis. *Eur J Immunol, 39*(8), 2040-2044.

22. Aupperle, K. R., Yamanishi, Y., Bennett, B. L., Mercurio, F., Boyle, D. L., & Firestein, G. S. (2001). Expression and regulation of inducible IkappaB kinase (IKK-i) in human fibroblast-like synoviocytes. *Cell Immunol, 214*(1), 54-59.

23. Bakker, A. C., Joosten, L. A., Arntz, O. J., Helsen, M. M., Bendele, A. M., van de Loo, F. A., et al. (1997). Prevention of murine collagen-induced arthritis in the knee and ipsilateral paw by local expression of human interleukin-1 receptor antagonist

protein in the knee. *Arthritis Rheum, 40*(5), 893-900.

24. Bakker, A. C., van de Loo, F. A., Joosten, L. A., Arntz, O. J., Varley, A. W., Munford, R. S., et al. (2002). C3-Tat/HIV-regulated intraarticular human interleukin-1 receptor antagonist gene therapy results in efficient inhibition of collagen-induced arthritis superior to cytomegalovirus-regulated expression of the same transgene. *Arthritis Rheum, 46*(6), 1661-1670.

25. Barde, I., Laurenti, E., Verp, S., Wiznerowicz, M., Offner, S., Viornery, A., et al. (2011). Lineage- and stage-restricted lentiviral vectors for the gene therapy of chronic granulomatous disease. *Gene Ther, 18*(11), 1087-1097.

26. Barkett, M., & Gilmore, T. D. (1999). Control of apoptosis by Rel/NF-kappaB transcription factors. *Oncogene, 18*(49), 6910-6924.

27. Barnes, D., Kunitomi, M., Vignuzzi, M., Saksela, K., & Andino, R. (2008). Harnessing endogenous miRNAs to control virus tissue tropism as a strategy for developing attenuated virus vaccines. *Cell Host Microbe, 4*(3), 239-248.

28. Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell, 116*(2), 281-297.

29. Bendele, A. (2001). Animal models of rheumatoid arthritis. *J Musculoskelet Neuronal Interact, 1*(4), 377-385.

30. Benham, C. J. (1979). Torsional stress and local denaturation in supercoiled DNA. *Proc Natl Acad Sci U S A, 76*(8), 3870-3874.

31. Berg, O. G., Winter, R. B., & von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry, 20*(24), 6929-6948.

32. Berse, B., Hunt, J. A., Diegel, R. J., Morganelli, P., Yeo, K., Brown, F., et al. (1999). Hypoxia augments cytokine (transforming growth factor-beta (TGF-beta) and IL-1)-induced vascular endothelial growth factor secretion by human synovial fibroblasts. *Clin Exp Immunol, 115*(1), 176-182.

33. Bevaart, L., Vervoordeldonk, M. J., & Tak, P. P. (2010). Evaluation of therapeutic targets in animal models of arthritis: how does it relate to rheumatoid arthritis? *Arthritis Rheum, 62*(8), 2192-2205.

34. Bloquel, C., Denys, A., Boissier, M. C., Apparailly, F., Bigey, P., Scherman, D., et al. (2007). Intra-articular electrotransfer of plasmid encoding soluble TNF receptor variants in normal and arthritic mice. *J Gene Med, 9*(11), 986-993.

35. Bluml, S., Bonelli, M., Niederreiter, B., Puchner, A., Mayr, G., Hayer, S., et al. (2011). Essential role of microRNA-155 in the pathogenesis of autoimmune arthritis in mice. *Arthritis Rheum, 63*(5), 1281-1288.

36. Bokhoven, M., Stephen, S. L., Knight, S., Gevers, E. F., Robinson, I. C., Takeuchi,

Y., et al. (2009). Insertional gene activation by lentiviral and gammaretroviral vectors. *J Virol, 83*(1), 283-294.

37. Bolinger, C., & Boris-Lawrie, K. (2009). Mechanisms employed by retroviruses to exploit host factors for translational control of a complicated proteome. *Retrovirology, 6*, 8.

38. Borchert, G. M., Lanier, W., & Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol, 13*(12), 1097-1101.

39. Boshart, M., Weber, F., Jahn, G., Dorsch-Hasler, K., Fleckenstein, B., & Schaffner, W. (1985). A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell, 41*(2), 521-530.

40. Bouard, D., Alazard-Dany, D., & Cosset, F. L. (2009). Viral vectors: from virology to transgene expression. *Br J Pharmacol, 157*(2), 153-165.

41. Brackertz, D., Mitchell, G. F., & Mackay, I. R. (1977). Antigen-induced arthritis in mice. I. Induction of arthritis in various strains of mice. *Arthritis Rheum, 20*(3), 841-850.

42. Brackley, C. A., Cates, M. E., & Marenduzzo, D. (2012). Facilitated diffusion on mobile DNA: configurational traps and sequence heterogeneity. *Phys Rev Lett, 109*(16), 168103.

43. Bradley, R. K., Li, X. Y., Trapnell, C., Davidson, S., Pachter, L., Chu, H. C., et al. (2010). Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related Drosophila species. *PLoS Biol, 8*(3), e1000343.

44. Brand, D. D., Latham, K. A., & Rosloniec, E. F. (2007). Collagen-induced arthritis. *Nat Protoc, 2*(5), 1269-1275.

45. Breathnach, R., & Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem, 50*, 349-383.

46. Brennan, P., Bankhead, C., Silman, A., & Symmons, D. (1997). Oral contraceptives and rheumatoid arthritis: results from a primary care-based incident case-control study. *Semin Arthritis Rheum, 26*(6), 817-823.

47. Brennecke, J., Stark, A., Russell, R. B., & Cohen, S. M. (2005). Principles of microRNA-target recognition. *PLoS Biol, 3*(3), e85.

48. Broderick, J. A., & Zamore, P. D. (2011). MicroRNA therapeutics. *Gene Ther, 18*(12), 1104-1110.

49. Brown, B. D., Cantore, A., Annoni, A., Sergi, L. S., Lombardo, A., Della Valle, P., et al. (2007c). A microRNA-regulated lentiviral vector mediates stable correction of hemophilia B mice. *Blood, 110*(13), 4144-4152.

50. Brown, B. D., Gentner, B., Cantore, A., Colleoni, S., Amendola, M., Zingale, A., et

al. (2007b). Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state. *Nat Biotechnol, 25*(12), 1457-1467.

51. Brown, B. D., & Naldini, L. (2009). Exploiting and antagonizing microRNA regulation for therapeutic and experimental applications. *Nat Rev Genet, 10*(8), 578-585.

52. Brown, B. D., Sitia, G., Annoni, A., Hauben, E., Sergi, L. S., Zingale, A., et al. (2007a). In vivo administration of lentiviral vectors triggers a type I interferon response that restricts hepatocyte gene transfer and promotes vector clearance. *Blood, 109*(7), 2797-2805.

53. Brown, B. D., Venneri, M. A., Zingale, A., Sergi Sergi, L., & Naldini, L. (2006). Endogenous microRNA regulation suppresses transgene expression in hematopoietic lineages and enables stable gene transfer. *Nat Med, 12*(5), 585-591.

54. Buch M.H., & Emery P. (2002) The aetiology and pathogenesis of rheumatoid arthritis. *Hospital Pharmacist*, **9** (1), pp. 5-10.

55. Buchschacher, G. L., Jr., & Wong-Staal, F. (2000). Development of lentiviral vectors for gene therapy for human diseases. *Blood, 95*(8), 2499-2504.

56. Burns, J. C., Friedmann, T., Driever, W., Burrascano, M., & Yee, J. K. (1993). Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc Natl Acad Sci U S A, 90*(17), 8033-8037.

57. Butler, D. M., Malfait, A. M., Mason, L. J., Warden, P. J., Kollias, G., Maini, R. N., et al. (1997). DBA/1 mice expressing the human TNF-alpha transgene develop a severe, erosive arthritis: characterization of the cytokine cascade and cellular composition. *J Immunol, 159*(6), 2867-2876.

58. Butler, J. E., & Kadonaga, J. T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev, 16*(20), 2583-2592.

59. Callen, J. P. (2007). Complications and adverse reactions in the use of newer biologic agents. *Semin Cutan Med Surg, 26*(1), 6-14.

60. Castoldi, M., Schmidt, S., Benes, V., Noerholm, M., Kulozik, A. E., Hentze, M. W., et al. (2006). A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA). *RNA, 12*(5), 913-920.

61. Chatzikyriakidou, A., Voulgari, P. V., Georgiou, I., & Drosos, A. A. (2010). A polymorphism in the 3'-UTR of interleukin-1 receptor-associated kinase (IRAK1), a target gene of miR-146a, is associated with rheumatoid arthritis susceptibility.

Joint Bone Spine, 77(5), 411-413.

62. Chavrier, P., Janssen-Timmen, U., Mattei, M. G., Zerial, M., Bravo, R., & Charnay, P. (1989). Structure, chromosome location, and expression of the mouse zinc finger gene Krox-20: multiple gene products and coregulation with the proto-oncogene c-fos. *Mol Cell Biol, 9*(2), 787-797.

63. Chernajovsky, Y., Gould, D. J., & Podhajcer, O. L. (2004). Gene therapy for autoimmune diseases: quo vadis? *Nat Rev Immunol, 4*(10), 800-811.

64. Chou, C. T., Pei, L., Chang, D. M., Lee, C. F., Schumacher, H. R., & Liang, M. H. (1994). Prevalence of rheumatic diseases in Taiwan: a population study of urban, suburban, rural differences. *J Rheumatol, 21*(2), 302-306.

65. Choy, E. (2012). Understanding the dynamics: pathways involved in the pathogenesis of rheumatoid arthritis. *Rheumatology (Oxford), 51 Suppl 5*, v3-11.

66. Christy, B., & Nathans, D. (1989). Functional serum response elements upstream of the growth factor-inducible gene zif268. *Mol Cell Biol, 9*(11), 4889-4895.

67. Cogswell, P. C., Mayo, M. W., & Baldwin, A. S., Jr. (1997). Involvement of Egr-1/RelA synergy in distinguishing T cell activation from tumor necrosis factor-alpha-induced NF-kappa B1 transcription. *J Exp Med, 185*(3), 491-497.

68. Courtenay, J. S., Dallman, M. J., Dayan, A. D., Martin, A., & Mosedale, B. (1980). Immunisation against heterologous type II collagen induces arthritis in mice. *Nature, 283*(5748), 666-668.

69. Crofford, L. J. (2013). Use of NSAIDs in treating patients with arthritis. *Arthritis Res Ther, 15 Suppl 3*, S2.

70. Cutolo, M., & Lahita, R. G. (2005). Estrogens and arthritis. *Rheum Dis Clin North Am, 31*(1), 19-27, vii.

71. Cuzzocrea, S., Costantino, G., Mazzon, E., & Caputi, A. P. (1999). Regulation of prostaglandin production in carrageenan-induced pleurisy by melatonin. *J Pineal Res, 27*(1), 9-14.

72. Czauderna, F., Fechtner, M., Dames, S., Aygun, H., Klippel, A., Pronk, G. J., et al. (2003). Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells. *Nucleic Acids Res, 31*(11), 2705-2716.

73. Decker, E. L., Nehmann, N., Kampen, E., Eibel, H., Zipfel, P. F., & Skerka, C. (2003). Early growth response proteins (EGR) and nuclear factors of activated T cells (NFAT) form heterodimers and regulate proinflammatory cytokine gene expression. *Nucleic Acids Res, 31*(3), 911-921.

74. Decker, E. L., Skerka, C., & Zipfel, P. F. (1998). The early growth response protein (EGR-1) regulates interleukin-2 transcription by synergistic interaction with the nuclear factor of activated T cells. *J Biol Chem, 273*(41), 26923-26930.

75. Demaison, C., Parsley, K., Brouns, G., Scherr, M., Battmer, K., Kinnon, C., et al. (2002). High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. *Hum Gene Ther, 13*(7), 803-813.

76. Descombes, P., Chojkier, M., Lichtsteiner, S., Falvey, E., & Schibler, U. (1990). LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. *Genes Dev, 4*(9), 1541-1551.

77. Di Domizio, J., Dorta-Estremera, S., & Cao, W. (2013). Methylated BSA mimics amyloid-related proteins and triggers inflammation. *PLoS One, 8*(5), e63214.

78. Dittmer, J. (2003). The biology of the Ets1 proto-oncogene. *Mol Cancer, 2*, 29.

79. Doan, Q. V., Chiou, C. F., & Dubois, R. W. (2006). Review of eight pharmacoeconomic studies of the value of biologic DMARDs (adalimumab, etanercept, and infliximab) in the management of rheumatoid arthritis. *J Manag Care Pharm, 12*(7), 555-569.

80. Dobi, K. C., & Winston, F. (2007). Analysis of transcriptional activation at a distance in Saccharomyces cerevisiae. *Mol Cell Biol, 27*(15), 5575-5586.

81. Doench, J. G., Petersen, C. P., & Sharp, P. A. (2003). siRNAs can function as miRNAs. *Genes Dev, 17*(4), 438-442.

82. Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D., et al. (1998). A third-generation lentivirus vector with a conditional packaging system. *J Virol, 72*(11), 8463-8471.

83. Durand, S., & Cimarelli, A. (2011). The inside out of lentiviral vectors. *Viruses, 3*(2), 132-159.

84. Ebert, M. S., Neilson, J. R., & Sharp, P. A. (2007). MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods, 4*(9), 721-726.

85. Ebert, M. S., & Sharp, P. A. (2010). MicroRNA sponges: progress and possibilities. *RNA, 16*(11), 2043-2050.

86. Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C., & Marks, D. S. (2003). MicroRNA targets in Drosophila. *Genome Biol, 5*(1), R1.

87. Evans, C. H., Ghivizzani, S. C., & Robbins, P. D. (2013). Arthritis gene therapy and its tortuous path into the clinic. *Transl Res, 161*(4), 205-216.

88. Feldmann, M., & Maini, S. R. (2008). Role of cytokines in rheumatoid arthritis: an education in pathophysiology and therapeutics. *Immunol Rev, 223*, 7-19.

89. Finnegan, E. F., & Pasquinelli, A. E. (2013). MicroRNA biogenesis: regulating the regulators. *Crit Rev Biochem Mol Biol, 48*(1), 51-68.

90. Fischer, E., Marano, M. A., Barber, A. E., Hudson, A., Lee, K., Rock, C. S., et al.

(1991). Comparison between effects of interleukin-1 alpha administration and sublethal endotoxemia in primates. *Am J Physiol, 261*(2 Pt 2), R442-452.

91. Forman, J. J., Legesse-Miller, A., & Coller, H. A. (2008). A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. *Proc Natl Acad Sci U S A, 105*(39), 14879-14884.

92. Forster, K., Helbl, V., Lederer, T., Urlinger, S., Wittenburg, N., & Hillen, W. (1999). Tetracycline-inducible expression systems with reduced basal activity in mammalian cells. *Nucleic Acids Res, 27*(2), 708-710.

93. Forsythe, J. A., Jiang, B. H., Iyer, N. V., Agani, F., Leung, S. W., Koos, R. D., et al. (1996). Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. *Mol Cell Biol, 16*(9), 4604-4613.

94. Fulci, V., Scappucci, G., Sebastiani, G. D., Giannitti, C., Franceschini, D., Meloni, F., et al. (2010). miR-223 is overexpressed in T-lymphocytes of patients affected by rheumatoid arthritis. *Hum Immunol, 71*(2), 206-211.

95. Gaber, T., Dziurla, R., Tripmacher, R., Burmester, G. R., & Buttgereit, F. (2005). Hypoxia inducible factor (HIF) in rheumatology: low O2! See what HIF can do! *Ann Rheum Dis, 64*(7), 971-980.

96. Gabriel, S. E., & Michaud, K. (2009). Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. *Arthritis Res Ther, 11*(3), 229.

97. Garaulet, G., Alfranca, A., Torrente, M., Escolano, A., Lopez-Fontal, R., Hortelano, S., et al. (2013). IL10 released by a new inflammation-regulated lentiviral system efficiently attenuates zymosan-induced arthritis. *Mol Ther, 21*(1), 119-130.

98. Gentner, B., & Naldini, L. (2012). Exploiting microRNA regulation for genetic engineering. *Tissue Antigens, 80*(5), 393-403.

99. Geurts, J., Arntz, O. J., Bennink, M. B., Joosten, L. A., van den Berg, W. B., & van de Loo, F. A. (2007). Application of a disease-regulated promoter is a safer mode of local IL-4 gene therapy for arthritis. *Gene Ther, 14*(23), 1632-1638.

100. Geurts, J., Joosten, L. A., Takahashi, N., Arntz, O. J., Gluck, A., Bennink, M. B., et al. (2009). Computational design and application of endogenous promoters for transcriptionally targeted gene therapy for rheumatoid arthritis. *Mol Ther, 17*(11), 1877-1887.

101. Ghosh, Z., Chakrabarti, J., & Mallick, B. (2007). miRNomics-The bioinformatics of microRNA genes. *Biochem Biophys Res Commun, 363*(1), 6-11.

102. Giatromanolaki, A., Sivridis, E., Maltezos, E., Athanassou, N., Papazoglou, D., Gatter, K. C., et al. (2003). Upregulated hypoxia inducible factor-1alpha and -2alpha pathway in rheumatoid arthritis and osteoarthritis. *Arthritis Res Ther, 5*(4),

R193-201.

103. Goetze, S., Kintscher, U., Kaneshiro, K., Meehan, W. P., Collins, A., Fleck, E., et al. (2001). TNFalpha induces expression of transcription factors c-fos, Egr-1, and Ets-1 in vascular lesions through extracellular signal-regulated kinases 1/2. *Atherosclerosis, 159*(1), 93-101.

104. Gonzalez, A., Maradit Kremers, H., Crowson, C. S., Nicola, P. J., Davis, J. M., 3rd, Therneau, T. M., et al. (2007). The widening mortality gap between rheumatoid arthritis patients and the general population. *Arthritis Rheum, 56*(11), 3583-3587.

105. Goossens, P. H., Schouten, G. J., t Hart, B. A., Bout, A., Brok, H. P., Kluin, P. M., et al. (1999). Feasibility of adenovirus-mediated nonsurgical synovectomy in collagen-induced arthritis-affected rhesus monkeys. *Hum Gene Ther, 10*(7), 1139-1149.

106. Gossen, M., & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A, 89*(12), 5547-5551.

107. Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., & Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science, 268*(5218), 1766-1769.

108. Gotea, V., Visel, A., Westlund, J. M., Nobrega, M. A., Pennacchio, L. A., & Ovcharenko, I. (2010). Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. *Genome Res, 20*(5), 565-577.

109. Gottschalk, L. R., Giannola, D. M., & Emerson, S. G. (1993). Molecular regulation of the human IL-3 gene: inducible T cell-restricted expression requires intact AP-1 and Elf-1 nuclear protein binding sites. *J Exp Med, 178*(5), 1681-1692.

110. Gould, D. J., & Chernajovsky, Y. (2007). Novel delivery methods to achieve immunomodulation. *Curr Opin Pharmacol, 7*(4), 445-450.

111. Gould, D., Yousaf, N., Fatah, R., Subang, M. C., & Chernajovsky, Y. (2007). Gene therapy with an improved doxycycline-regulated plasmid encoding a tumour necrosis factor-alpha inhibitor in experimental arthritis. *Arthritis Res Ther, 9*(1), R7.

112. Gould, D. J., Bright, C., & Chernajovsky, Y. (2004). Inhibition of established collagen-induced arthritis with a tumour necrosis factor-alpha inhibitor expressed from a self-contained doxycycline regulated plasmid. *Arthritis Res Ther, 6*(2), R103-113.

113. Gould, D. J., & Favorov, P. (2003). Vectors for the treatment of autoimmune disease. *Gene Ther, 10*(10), 912-927.

114. Gouze, E., Gouze, J. N., Palmer, G. D., Pilapil, C., Evans, C. H., & Ghivizzani, S.

C. (2007). Transgene persistence and cell turnover in the diarthrodial joint: implications for gene therapy of chronic joint diseases. *Mol Ther, 15*(6), 1114-1120.

115. Gravallese, E. M. (2002). Bone destruction in arthritis. *Ann Rheum Dis, 61 Suppl 2*, ii84-86.

116. Gregersen, P. K., Silver, J., & Winchester, R. J. (1987). The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum, 30*(11), 1205-1213.

117. Griffiths-Jones, S., Saini, H. K., van Dongen, S., & Enright, A. J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res, 36*(Database issue), D154-158.

118. Grimbacher, B., Aicher, W. K., Peter, H. H., & Eibel, H. (1997). Measurement of transcription factor c-fos and EGR-1 mRNA transcription levels in synovial tissue by quantitative RT-PCR. *Rheumatol Int, 17*(3), 109-112.

119. Grimbacher, B., Aicher, W. K., Peter, H. H., & Eibel, H. (1998). TNF-alpha induces the transcription factor Egr-1, pro-inflammatory cytokines and cell proliferation in human skin fibroblasts and synovial lining cells. *Rheumatol Int, 17*(5), 185-192.

120. Guarente, L., & Hoar, E. (1984). Upstream activation sites of the CYC1 gene of Saccharomyces cerevisiae are active when inverted but not when placed downstream of the "TATA box". *Proc Natl Acad Sci U S A, 81*(24), 7860-7864.

121. Halford, S. E., & Marko, J. F. (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res, 32*(10), 3040-3052.

122. Handel, M. L., McMorrow, L. B., & Gravallese, E. M. (1995). Nuclear factor-kappa B in rheumatoid synovium. Localization of p50 and p65. *Arthritis Rheum, 38*(12), 1762-1770.

123. Harada, M., Mitsuyama, K., Yoshida, H., Sakisaka, S., Taniguchi, E., Kawaguchi, T., et al. (1998). Vascular endothelial growth factor in patients with rheumatoid arthritis. *Scand J Rheumatol, 27*(5), 377-380.

124. Hashiya, N., Jo, N., Aoki, M., Matsumoto, K., Nakamura, T., Sato, Y., et al. (2004). In vivo evidence of angiogenesis induced by transcription factor Ets-1: Ets-1 is located upstream of angiogenesis cascade. *Circulation, 109*(24), 3035-3041.

125. Hayashida, M., Okazaki, K., Fukushi, J., Sakamoto, A., & Iwamoto, Y. (2009). CCAAT/enhancer binding protein beta mediates expression of matrix metalloproteinase 13 in human articular chondrocytes in inflammatory arthritis. *Arthritis Rheum, 60*(3), 708-716.

126. Hayden, M. S., & Ghosh, S. (2012). NF-kappaB, the first quarter-century: remarkable progress and outstanding questions. *Genes Dev, 26*(3), 203-234.

127. Haywood, L., & Walsh, D. A. (2001). Vasculature of the normal and arthritic

synovial joint. *Histol Histopathol, 16*(1), 277-284.

128. Henningsson, L., Eneljung, T., Jirholt, P., Tengvall, S., Lidberg, U., van den Berg, W. B., et al. (2012). Disease-dependent local IL-10 production ameliorates collagen induced arthritis in mice. *PLoS One, 7*(11), e49731.

129. Henriques, M. G., Silva, P. M., Martins, M. A., Flores, C. A., Cunha, F. Q., Assreuy-Filho, J., et al. (1987). Mouse paw edema. A new model for inflammation? *Braz J Med Biol Res, 20*(2), 243-249.

130. Hirata, M., Kugimiya, F., Fukai, A., Ohba, S., Kawamura, N., Ogasawara, T., et al. (2009). C/EBPbeta Promotes transition from proliferation to hypertrophic differentiation of chondrocytes through transactivation of p57. *PLoS One, 4*(2), e4543.

131. Hitchon, C. A., & El-Gabalawy, H. S. (2011). The synovium in rheumatoid arthritis. *Open Rheumatol J, 5*, 107-114.

132. Hochberg, M. C., Johnston, S. S., & John, A. K. (2008). The incidence and prevalence of extra-articular and systemic manifestations in a cohort of newly-diagnosed patients with rheumatoid arthritis between 1999 and 2006. *Curr Med Res Opin, 24*(2), 469-480.

133. Hodges, B. L., Taylor, K. M., Joseph, M. F., Bourgeois, S. A., & Scheule, R. K. (2004). Long-term transgene expression from plasmid DNA gene therapy vectors is negatively affected by CpG dinucleotides. *Mol Ther, 10*(2), 269-278.

134. Hoesel, B., & Schmid, J. A. (2013). The complexity of NF-kappaB signaling in inflammation and cancer. *Mol Cancer, 12*, 86.

135. Hollander, A. P., Corke, K. P., Freemont, A. J., & Lewis, C. E. (2001). Expression of hypoxia-inducible factor 1alpha by macrophages in the rheumatoid synovium: implications for targeting of therapeutic genes to the inflamed joint. *Arthritis Rheum, 44*(7), 1540-1544.

136. Hollenhorst, P. C., Jones, D. A., & Graves, B. J. (2004). Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic Acids Res, 32*(18), 5693-5702.

137. Hu, L., Grosberg, A. Y., & Bruinsma, R. (2008). Are DNA transcription factor proteins maxwellian demons? *Biophys J, 95*(3), 1151-1156.

138. Hu, W. S., & Hughes, S. H. (2012). HIV-1 reverse transcription. *Cold Spring Harb Perspect Med, 2*(10).

139. Huber, L. C., Distler, O., Tarner, I., Gay, R. E., Gay, S., & Pap, T. (2006). Synovial fibroblasts: key players in rheumatoid arthritis. *Rheumatology (Oxford), 45*(6), 669-675.

140. Hughes, C., Faurholm, B., Dell'Accio, F., Manzo, A., Seed, M., Eltawil, N., et al.

(2010). Human single-chain variable fragment that specifically targets arthritic cartilage. *Arthritis Rheum, 62*(4), 1007-1016.

141. Hydbring, P., & Badalian-Very, G. (2013). Clinical applications of microRNAs. *F1000Res, 2*, 136.

142. Iotsova, V., Caamano, J., Loy, J., Yang, Y., Lewin, A., & Bravo, R. (1997). Osteopetrosis in mice lacking NF-kappaB1 and NF-kappaB2. *Nat Med, 3*(11), 1285-1289.

143. Iwamoto, T., Okamoto, H., Toyama, Y., & Momohara, S. (2008). Molecular aspects of rheumatoid arthritis: chemokines in the joints of patients. *FEBS J, 275*(18), 4448-4455.

144. Iwasaka, C., Tanaka, K., Abe, M., & Sato, Y. (1996). Ets-1 regulates angiogenesis by inducing the expression of urokinase-type plasminogen activator and matrix metalloproteinase-1 and the migration of vascular endothelial cells. *J Cell Physiol, 169*(3), 522-531.

145. Jackson, L. M., Wu, K. C., Mahida, Y. R., Jenkins, D., & Hawkey, C. J. (2000). Cyclooxygenase (COX) 1 and 2 in normal, inflamed, and ulcerated human gastric mucosa. *Gut, 47*(6), 762-770.

146. Jawed, S., Gaffney, K., & Blake, D. R. (1997). Intra-articular pressure profile of the knee joint in a spectrum of inflammatory arthropathies. *Ann Rheum Dis, 56*(11), 686-689.

147. Jiang, B. H., Semenza, G. L., Bauer, C., & Marti, H. H. (1996). Hypoxia-inducible factor 1 levels vary exponentially over a physiologically relevant range of O2 tension. *Am J Physiol, 271*(4 Pt 1), C1172-1180.

148. Jimi, E., Aoki, K., Saito, H., D'Acquisto, F., May, M. J., Nakamura, I., et al. (2004). Selective inhibition of NF-kappa B blocks osteoclastogenesis and prevents inflammatory bone destruction in vivo. *Nat Med, 10*(6), 617-624.

149. Jochum, W., Passegue, E., & Wagner, E. F. (2001). AP-1 in mouse development and tumorigenesis. *Oncogene, 20*(19), 2401-2412.

150. Jones, E. Y., Fugger, L., Strominger, J. L., & Siebold, C. (2006). MHC class II proteins and disease: a structural perspective. *Nat Rev Immunol, 6*(4), 271-282.

151. Kai, Z. S., & Pasquinelli, A. E. (2010). MicroRNA assassins: factors that regulate the disappearance of miRNAs. *Nat Struct Mol Biol, 17*(1), 5-10.

152. Kannan, K., Ortmann, R. A., & Kimpel, D. (2005). Animal models of rheumatoid arthritis and their relevance to human disease. *Pathophysiology, 12*(3), 167-181.

153. Karn, J., & Stoltzfus, C. M. (2012). Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med, 2*(2), a006916.

154. Kawasaki, H., Komai, K., Nakamura, M., Yamamoto, E., Ouyang, Z., Nakashima, T., et al. (2003). Human wee1 kinase is directly transactivated by and increased in association with c-Fos/AP-1: rheumatoid synovial cells overexpressing these genes go into aberrant mitosis. *Oncogene, 22*(44), 6839-6844.

155. Kay, M. A., Glorioso, J. C., & Naldini, L. (2001). Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nat Med, 7*(1), 33-40.

156. Kelly, E. J., Hadac, E. M., Greiner, S., & Russell, S. J. (2008). Engineering microRNA responsiveness to decrease virus pathogenicity. *Nat Med, 14*(11), 1278-1283.

157. Kenneth, N. S., & Rocha, S. (2008). Regulation of gene expression by hypoxia. *Biochem J, 414*(1), 19-29.

158. Khoury, M., Adriaansen, J., Vervoordeldonk, M. J., Gould, D., Chernajovsky, Y., Bigey, P., et al. (2007). Inflammation-inducible anti-TNF gene expression mediated by intra-articular injection of serotype 5 adeno-associated virus reduces arthritis. *J Gene Med, 9*(7), 596-604.

159. Khvorova, A., Reynolds, A., & Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell, 115*(2), 209-216.

160. Kiely, P. D., Deighton, C., Dixey, J., & Ostor, A. J. (2012). Biologic agents for rheumatoid arthritis--negotiating the NICE technology appraisals. *Rheumatology (Oxford), 51*(1), 24-31.

161. Kloosterman, W. P., Wienholds, E., de Bruijn, E., Kauppinen, S., & Plasterk, R. H. (2006). In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat Methods, 3*(1), 27-29.

162. Knedla, A., Neumann, E., & Muller-Ladner, U. (2007). Developments in the synovial biology field 2006. *Arthritis Res Ther, 9*(2), 209.

163. Kobayashi, T., Okamoto, K., Kobata, T., Hasunuma, T., Kato, T., Hamada, H., et al. (2000). Differential regulation of Fas-mediated apoptosis of rheumatoid synoviocytes by tumor necrosis factor alpha and basic fibroblast growth factor is associated with the expression of apoptosis-related molecules. *Arthritis Rheum, 43*(5), 1106-1114.

164. Koch, A. E. (2003). Angiogenesis as a target in rheumatoid arthritis. *Ann Rheum Dis, 62 Suppl 2*, ii60-67.

165. Kollias, G., Papadaki, P., Apparailly, F., Vervoordeldonk, M. J., Holmdahl, R., Baumans, V., et al. (2011). Animal models for arthritis: innovative tools for prevention and treatment. *Ann Rheum Dis, 70*(8), 1357-1362.

166. Kouskoff, V., Korganow, A. S., Duchatelle, V., Degott, C., Benoist, C., & Mathis, D. (1996). Organ-specific disease provoked by systemic autoimmunity. *Cell, 87*(5),

811-822.

167. Kramer, B., Meichle, A., Hensel, G., Charnay, P., & Kronke, M. (1994). Characterization of an Krox-24/Egr-1-responsive element in the human tumor necrosis factor promoter. *Biochim Biophys Acta, 1219*(2), 413-421.

168. Kramer, B., Wiegmann, K., & Kronke, M. (1995). Regulation of the human TNF promoter by the transcription factor Ets. *J Biol Chem, 270*(12), 6577-6583.

169. Krieg, A. M., Yi, A. K., Matson, S., Waldschmidt, T. J., Bishop, G. A., Teasdale, R., et al. (1995). CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature, 374*(6522), 546-549.

170. Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., et al. (2005). Silencing of microRNAs in vivo with 'antagomirs'. *Nature, 438*(7068), 685-689.

171. Kujubu, D. A., Lim, R. W., Varnum, B. C., & Herschman, H. R. (1987). Induction of transiently expressed genes in PC-12 pheochromocytoma cells. *Oncogene, 1*(3), 257-262.

172. Kumar, V., Abbas, A. K., Fausto, N., & Mitchell, R. N. (2007). Chapter 5 - Diseases of the Immune System. In *Robbins Basic Pathology, 8th Edition* (pp. 107 - 172). Philadelphia: Saunders, an imprint of Elsevier Inc.

173. Kuroki, Y., Shiozawa, S., Yoshihara, R., & Hotta, H. (1993). The contribution of human c-fos DNA to cultured synovial cells: a transfection study. *J Rheumatol, 20*(3), 422-428.

174. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., & Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science, 294*(5543), 853-858.

175. Lau, N. C., Lim, L. P., Weinstein, E. G., & Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science, 294*(5543), 858-862.

176. Lee, R. C., & Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. *Science, 294*(5543), 862-864.

177. Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell, 75*(5), 843-854.

178. Lee, S. S., Joo, Y. S., Kim, W. U., Min, D. J., Min, J. K., Park, S. H., et al. (2001). Vascular endothelial growth factor levels in the serum and synovial fluid of patients with rheumatoid arthritis. *Clin Exp Rheumatol, 19*(3), 321-324.

179. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature, 425*(6956), 415-419.

180. Lee, Y., Jeon, K., Lee, J. T., Kim, S., & Kim, V. N. (2002). MicroRNA maturation:

stepwise processing and subcellular localization. *EMBO J, 21*(17), 4663-4670.

181. Lekstrom-Himes, J., & Xanthopoulos, K. G. (1998). Biological role of the CCAAT/enhancer-binding protein family of transcription factors. *J Biol Chem, 273*(44), 28545-28548.

182. Levy, L. (1969). Carrageenan paw edema in the mouse. *Life Sci, 8*(11), 601-606.

183. Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., & Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell, 115*(7), 787-798.

184. Li, J., Wan, Y., Guo, Q., Zou, L., Zhang, J., Fang, Y., et al. (2010). Altered microRNA expression profile with miR-146a upregulation in CD4+ T cells from patients with rheumatoid arthritis. *Arthritis Res Ther, 12*(3), R81.

185. Logan, S. K., Garabedian, M. J., Campbell, C. E., & Werb, Z. (1996). Synergistic transcriptional activation of the tissue inhibitor of metalloproteinases-1 promoter via functional interaction of AP-1 and Ets-1 transcription factors. *J Biol Chem, 271*(2), 774-782.

186. Lund-Olesen, K. (1970). Oxygen tension in synovial fluids. *Arthritis Rheum, 13*(6), 769-776.

187. Lytle, J. R., Yario, T. A., & Steitz, J. A. (2007). Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. *Proc Natl Acad Sci U S A, 104*(23), 9667-9672.

188. MacGregor, A. J., Snieder, H., Rigby, A. S., Koskenvuo, M., Kaprio, J., Aho, K., et al. (2000). Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum, 43*(1), 30-37.

189. Makarov, S. S. (2001). NF-kappa B in rheumatoid arthritis: a pivotal regulator of inflammation, hyperplasia, and tissue destruction. *Arthritis Res, 3*(4), 200-206.

190. Matsusaka, T., Fujikawa, K., Nishio, Y., Mukaida, N., Matsushima, K., Kishimoto, T., et al. (1993). Transcription factors NF-IL6 and NF-kappa B synergistically activate transcription of the inflammatory cytokines, interleukin 6 and interleukin 8. *Proc Natl Acad Sci U S A, 90*(21), 10193-10197.

191. Mazzon, E., Esposito, E., Di Paola, R., Muia, C., Crisafulli, C., Genovese, T., et al. (2008). Effect of tumour necrosis factor-alpha receptor 1 genetic deletion on carrageenan-induced acute inflammation: a comparison with etanercept. *Clin Exp Immunol, 153*(1), 136-149.

192. McInnes, I. B., & Schett, G. (2007). Cytokines in the pathogenesis of rheumatoid arthritis. *Nat Rev Immunol, 7*(6), 429-442.

193. McInnes, I. B., & Schett, G. (2011). The pathogenesis of rheumatoid arthritis. *N Engl J Med, 365*(23), 2205-2219.

194. McIntyre, K. W., Stepan, G. J., Kolinsky, K. D., Benjamin, W. R., Plocinski, J. M.,

Kaffka, K. L., et al. (1991). Inhibition of interleukin 1 (IL-1) binding and bioactivity in vitro and modulation of acute inflammation in vivo by IL-1 receptor antagonist and anti-IL-1 receptor monoclonal antibody. *J Exp Med, 173*(4), 931-939.

195. McKnight, S. L. (1982). Functional relationships between transcriptional control signals of the thymidine kinase gene of herpes simplex virus. *Cell, 31*(2 Pt 1), 355-365.

196. Meier, F. M., Frerix, M., Hermann, W., & Muller-Ladner, U. (2013). Current immunotherapy in rheumatoid arthritis. *Immunotherapy, 5*(9), 955-974.

197. Melikyan, G. B. (2008). Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology, 5*, 111.

198. Mi, Z., Ghivizzani, S. C., Lechman, E. R., Jaffurs, D., Glorioso, J. C., Evans, C. H., et al. (2000). Adenovirus-mediated gene transfer of insulin-like growth factor 1 stimulates proteoglycan synthesis in rabbit joints. *Arthritis Rheum, 43*(11), 2563-2570.

199. Miagkov, A. V., Kovalenko, D. V., Brown, C. E., Didsbury, J. R., Cogswell, J. P., Stimpson, S. A., et al. (1998). NF-kappaB activation provides the potential link between inflammation and hyperplasia in the arthritic joint. *Proc Natl Acad Sci U S A, 95*(23), 13859-13864.

200. Miagkov, A. V., Varley, A. W., Munford, R. S., & Makarov, S. S. (2002). Endogenous regulation of a therapeutic transgene restores homeostasis in arthritic joints. *J Clin Invest, 109*(9), 1223-1229.

201. Milbrandt, J. (1987). A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor. *Science, 238*(4828), 797-799.

202. Mirkin SM. (2001). DNA Topology: Fundamentals. Encyclopedia of Life Sciences. Nature Publishing Group.

203. Mitchell, J. A., & Warner, T. D. (1999). Cyclo-oxygenase-2: pharmacology, physiology, biochemistry and relevance to NSAID therapy. *Br J Pharmacol, 128*(6), 1121-1132.

204. Moodley, I. (2008). Review of the cardiovascular safety of COXIBs compared to NSAIDS. *Cardiovasc J Afr, 19*(2), 102-107.

205. Moreland, L. W., Baumgartner, S. W., Schiff, M. H., Tindall, E. A., Fleischmann, R. M., Weaver, A. L., et al. (1997). Treatment of rheumatoid arthritis with a recombinant human tumor necrosis factor receptor (p75)-Fc fusion protein. *N Engl J Med, 337*(3), 141-147.

206. Morris, C. J. (2003). Carrageenan-induced paw edema in the rat and mouse. *Methods Mol Biol, 225*, 115-121.

207. Mukherjee, D., Nissen, S. E., & Topol, E. J. (2001). Risk of cardiovascular events

associated with selective COX-2 inhibitors. *JAMA, 286*(8), 954-959.

208. Muller-Ladner, U., Ospelt, C., Gay, S., Distler, O., & Pap, T. (2007). Cells of the synovium in rheumatoid arthritis. Synovial fibroblasts. *Arthritis Res Ther, 9*(6), 223.

209. Murata, K., Yoshitomi, H., Tanida, S., Ishikawa, M., Nishitani, K., Ito, H., et al. (2010). Plasma and synovial fluid microRNAs as potential biomarkers of rheumatoid arthritis and osteoarthritis. *Arthritis Res Ther, 12*(3), R86.

210. Nakamachi, Y., Kawano, S., Takenokuchi, M., Nishimura, K., Sakai, Y., Chin, T., et al. (2009). MicroRNA-124a is a key regulator of proliferation and monocyte chemoattractant protein 1 secretion in fibroblast-like synoviocytes from patients with rheumatoid arthritis. *Arthritis Rheum, 60*(5), 1294-1304.

211. Nakasa, T., Miyaki, S., Okubo, A., Hashimoto, M., Nishida, K., Ochi, M., et al. (2008). Expression of microRNA-146 in rheumatoid arthritis synovial tissue. *Arthritis Rheum, 58*(5), 1284-1292.

212. Nakasa, T., Shibuya, H., Nagata, Y., Niimoto, T., & Ochi, M. (2011). The inhibitory effect of microRNA-146a expression on bone destruction in collagen-induced arthritis. *Arthritis Rheum, 63*(6), 1582-1590.

213. Naldini, L., Blomer, U., Gage, F. H., Trono, D., & Verma, I. M. (1996b). Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc Natl Acad Sci U S A, 93*(21), 11382-11388.

214. Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., et al. (1996a). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science, 272*(5259), 263-267.

215. National Collaborating Centre for Chronic Conditions. *Rheumatoid arthritis: national clinical guideline for management and treatment in adults*. London: Royal College of Physicians, February 2009.

216. Naughton, D., Whelan, M., Smith, E. C., Williams, R., Blake, D. R., & Grootveld, M. (1993). An investigation of the abnormal metabolic status of synovial fluid from patients with rheumatoid arthritis by high field proton nuclear magnetic resonance spectroscopy. *FEBS Lett, 317*(1-2), 135-138.

217. Necas, J., and Bartosikova, L. (2013). Carrageenan: a review. *Veterinarni Medicina, 58 (4):* 187–205.

218. Nielen, M. M., van Schaardenburg, D., Reesink, H. W., van de Stadt, R. J., van der Horst-Bruinsma, I. E., de Koning, M. H., et al. (2004). Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum, 50*(2), 380-386.

219. Niimoto, T., Nakasa, T., Ishikawa, M., Okuhara, A., Izumi, B., Deie, M., et al.

(2010). MicroRNA-146a expresses in interleukin-17 producing T cells in rheumatoid arthritis patients. *BMC Musculoskelet Disord, 11*, 209.

220. Nishioka, K., Ohshima, S., Umeshita-Sasai, M., Yamaguchi, N., Mima, T., Nomura, S., et al. (2000). Enhanced expression and DNA binding activity of two CCAAT/enhancer-binding protein isoforms, C/EBPbeta and C/EBPdelta, in rheumatoid synovium. *Arthritis Rheum, 43*(7), 1591-1596.

221. Nunn, M. F., Seeburg, P. H., Moscovici, C., & Duesberg, P. H. (1983). Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature, 306*(5941), 391-395.

222. Nye, J. A., Petersen, J. M., Gunther, C. V., Jonsen, M. D., & Graves, B. J. (1992). Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev, 6*(6), 975-990.

223. Okamoto, H., Cujec, T. P., Yamanaka, H., & Kamatani, N. (2008). Molecular aspects of rheumatoid arthritis: role of transcription factors. *FEBS J, 275*(18), 4463-4470.

224. Okamoto, K., Asahara, H., Kobayashi, T., Matsuno, H., Hasunuma, T., Kobata, T., et al. (1998). Induction of apoptosis in the rheumatoid synovium by Fas ligand gene transfer. *Gene Ther, 5*(3), 331-338.

225. Osada, S., Yamamoto, H., Nishihara, T., & Imagawa, M. (1996). DNA binding specificity of the CCAAT/enhancer-binding protein transcription factor family. *J Biol Chem, 271*(7), 3891-3896.

226. Ospelt, C., Neidhart, M., Gay, R. E., & Gay, S. (2004). Synovial activation in rheumatoid arthritis. *Front Biosci, 9*, 2323-2334.

227. Ossipow, V., Descombes, P., & Schibler, U. (1993). CCAAT/enhancer-binding protein mRNA is translated into multiple proteins with different transcription activation potentials. *Proc Natl Acad Sci U S A, 90*(17), 8219-8223.

228. Otero, M., & Goldring, M. B. (2007). Cells of the synovium in rheumatoid arthritis. Chondrocytes. *Arthritis Res Ther, 9*(5), 220.

229. Paleolog, E. M. (2002). Angiogenesis in rheumatoid arthritis. *Arthritis Res, 4 Suppl 3*, S81-90.

230. Paleolog, E. M. (2009). The vasculature in rheumatoid arthritis: cause or consequence? *Int J Exp Pathol, 90*(3), 249-261.

231. Papadakis, E. D., Nicklin, S. A., Baker, A. H., & White, S. J. (2004). Promoters and control elements: designing expression cassettes for gene therapy. *Curr Gene Ther, 4*(1), 89-113.

232. Papapetrou, E. P., Kovalovsky, D., Beloeil, L., Sant'angelo, D., & Sadelain, M. (2009). Harnessing endogenous miR-181a to segregate transgenic antigen

receptor expression in developing versus post-thymic T cells in murine hematopoietic chimeras. *J Clin Invest, 119*(1), 157-168.

233. Parker, C. S., & Topol, J. (1984). A Drosophila RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. *Cell, 36*(2), 357-369.

234. Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., et al. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature, 408*(6808), 86-89.

235. Pauley, K. M., Satoh, M., Chan, A. L., Bubb, M. R., Reeves, W. H., & Chan, E. K. (2008). Upregulated miR-146a expression in peripheral blood mononuclear cells from rheumatoid arthritis patients. *Arthritis Res Ther, 10*(4), R101.

236. Peters, C. L., Morris, C. J., Mapp, P. I., Blake, D. R., Lewis, C. E., & Winrow, V. R. (2004). The transcription factors hypoxia-inducible factor 1alpha and Ets-1 colocalize in the hypoxic synovium of inflamed joints in adjuvant-induced arthritis. *Arthritis Rheum, 50*(1), 291-296.

237. Peterson, S. M., Thompson, J. A., Ufkin, M. L., Sathyanarayana, P., Liaw, L., & Congdon, C. B. (2014). Common features of microRNA target prediction tools. *Front Genet, 5*, 23.

238. Philippe, L., Alsaleh, G., Suffert, G., Meyer, A., Georgel, P., Sibilia, J., et al. (2012). TLR2 expression is regulated by microRNA miR-19 in rheumatoid fibroblast-like synoviocytes. *J Immunol, 188*(1), 454-461.

239. Pikaart, M. J., Recillas-Targa, F., & Felsenfeld, G. (1998). Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev, 12*(18), 2852-2862.

240. Pluta, K., & Kacprzak, M. M. (2009). Use of HIV as a gene transfer vector. *Acta Biochim Pol, 56*(4), 531-595.

241. Poeschla, E. M. (2008). Integrase, LEDGF/p75 and HIV replication. *Cell Mol Life Sci, 65*(9), 1403-1424.

242. Ponder K.P. (2001). Chapter 4: Vectors of gene therapy. In: Kresina, T.F. ed. An *Introduction to Molecular Medicine and Gene Therapy.* John Wiley & Sons, Inc., pp. 77-112.

243. Pope, R. M., Lovis, R., Mungre, S., Perlman, H., Koch, A. E., & Haines, G. K., 3rd. (1999). C/EBP beta in rheumatoid arthritis: correlation with inflammation, not disease specificity. *Clin Immunol, 91*(3), 271-282.

244. Posadas, I., Bucci, M., Roviezzo, F., Rossi, A., Parente, L., Sautebin, L., et al. (2004). Carrageenan-induced mouse paw oedema is biphasic, age-weight dependent and displays differential nitric oxide cyclooxygenase-2 expression. *Br*

J Pharmacol, 142(2), 331-338.

245. Prelich, G., Kostura, M., Marshak, D. R., Mathews, M. B., & Stillman, B. (1987). The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication in vitro. *Nature, 326*(6112), 471-475.

246. Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. *Nature, 322*(6081), 697-701.

247. Quattrocchi, E., Dallman, M. J., & Feldmann, M. (2000). Adenovirus-mediated gene transfer of CTLA-4Ig fusion protein in the suppression of experimental autoimmune arthritis. *Arthritis Rheum, 43*(8), 1688-1697.

248. Ramji, D. P., & Foka, P. (2002). CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem J, 365*(Pt 3), 561-575.

249. Recillas-Targa, F., Pikaart, M. J., Burgess-Beusse, B., Bell, A. C., Litt, M. D., West, A. G., et al. (2002). Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci U S A, 99*(10), 6883-6888.

250. Redlich, K., Kiener, H. P., Schett, G., Tohidast-Akrad, M., Selzer, E., Radda, I., et al. (2001). Overexpression of transcription factor Ets-1 in rheumatoid arthritis synovial membrane: regulation of expression and activation by interleukin-1 and tumor necrosis factor alpha. *Arthritis Rheum, 44*(2), 266-274.

251. Reinhart, B. J., & Bartel, D. P. (2002). Small RNAs correspond to centromere heterochromatic repeats. *Science, 297*(5588), 1831.

252. Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature, 403*(6772), 901-906.

253. Robbins, P. D., Evans, C. H., & Chernajovsky, Y. (2003). Gene therapy for arthritis. *Gene Ther, 10*(10), 902-911.

254. Rocha, A. C., Fernandes, E. S., Quintao, N. L., Campos, M. M., & Calixto, J. B. (2006). Relevance of tumour necrosis factor-alpha for the inflammatory and nociceptive responses evoked by carrageenan in the mouse paw. *Br J Pharmacol, 148*(5), 688-695.

255. Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L., & Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res, 14*(10A), 1902-1910.

256. Roman-Blas, J. A., & Jimenez, S. A. (2006). NF-kappaB as a potential therapeutic target in osteoarthritis and rheumatoid arthritis. *Osteoarthritis Cartilage, 14*(9), 839-848.

257. Russo, M. W., Sevetson, B. R., & Milbrandt, J. (1995). Identification of NAB1, a

repressor of NGFI-A- and Krox20-mediated transcription. *Proc Natl Acad Sci U S A, 92*(15), 6873-6877.

258. Rutkauskaite, E., Zacharias, W., Schedel, J., Muller-Ladner, U., Mawrin, C., Seemayer, C. A., et al. (2004). Ribozymes that inhibit the production of matrix metalloproteinase 1 reduce the invasiveness of rheumatoid arthritis synovial fibroblasts. *Arthritis Rheum, 50*(5), 1448-1456.

259. Ruyssen-Witrand, A., Constantin, A., Cambon-Thomsen, A., & Thomsen, M. (2012). New insights into the genetics of immune responses in rheumatoid arthritis. *Tissue Antigens, 80*(2), 105-118.

260. Sarafianos, S. G., Marchand, B., Das, K., Himmel, D. M., Parniak, M. A., Hughes, S. H., et al. (2009). Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J Mol Biol, 385*(3), 693-713.

261. Scherr, M., Venturini, L., Battmer, K., Schaller-Schoenitz, M., Schaefer, D., Dallmann, I., et al. (2007). Lentivirus-mediated antagomir expression for specific inhibition of miRNA function. *Nucleic Acids Res, 35*(22), e149.

262. Schleif, R. (1992). DNA looping. *Annu Rev Biochem, 61*, 199-223.

263. Schmittgen, T. D., Jiang, J., Liu, Q., & Yang, L. (2004). A high-throughput method to monitor the expression of microRNA precursors. *Nucleic Acids Res, 32*(4), e43.

264. Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., & Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell, 110*(4), 521-529.

265. Seibl, R., Birchler, T., Loeliger, S., Hossle, J. P., Gay, R. E., Saurenmann, T., et al. (2003). Expression and regulation of Toll-like receptor 2 in rheumatoid arthritis synovium. *Am J Pathol, 162*(4), 1221-1227.

266. Semenza, G. L. (2000). HIF-1: mediator of physiological and pathophysiological responses to hypoxia. *J Appl Physiol (1985), 88*(4), 1474-1480.

267. Shaikhibrahim, Z., & Wernert, N. (2012). ETS transcription factors and prostate cancer: the role of the family prototype ETS-1 (review). *Int J Oncol, 40*(6), 1748-1754.

268. Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., et al. (2012). Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat Biotechnol, 30*(6), 521-530.

269. Shaulian, E., & Karin, M. (2001). AP-1 in cell proliferation and survival. *Oncogene, 20*(19), 2390-2400.

270. Shibata, T., Giaccia, A. J., & Brown, J. M. (2000). Development of a hypoxia-responsive vector for tumor-specific gene therapy. *Gene Ther, 7*(6), 493-498.

271. Shin, S. Y., Kim, J. H., Baker, A., Lim, Y., & Lee, Y. H. (2010). Transcription factor Egr-1 is essential for maximal matrix metalloproteinase-9 transcription by tumor necrosis factor alpha. *Mol Cancer Res, 8*(4), 507-519.

272. Shiozawa, S., Shimizu, K., Tanaka, K., & Hino, K. (1997). Studies on the contribution of c-fos/AP-1 to arthritic joint destruction. *J Clin Invest, 99*(6), 1210-1216.

273. Shiozawa, S., & Tsumiyama, K. (2009). Pathogenesis of rheumatoid arthritis and c-Fos/AP-1. *Cell Cycle, 8*(10), 1539-1543.

274. Shiozawa, S., Tsumiyama, K., Yoshida, K., & Hashiramoto, A. (2011). Pathogenesis of joint destruction in rheumatoid arthritis. *Arch Immunol Ther Exp (Warsz), 59*(2), 89-95.

275. Silman, A. J., & Pearson, J. E. (2002). Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res, 4 Suppl 3*, S265-272.

276. Silverman, E. S., Du, J., Williams, A. J., Wadgaonkar, R., Drazen, J. M., & Collins, T. (1998). cAMP-response-element-binding-protein-binding protein (CBP) and p300 are transcriptional co-activators of early growth response factor-1 (Egr-1). *Biochem J, 336 (Pt 1)*, 183-189.

277. Simmonds, R. E., & Foxwell, B. M. (2008). Signalling, inflammation and arthritis: NF-kappaB and its relevance to arthritis and inflammation. *Rheumatology (Oxford), 47*(5), 584-590.

278. Sinn, P. L., Sauter, S. L., & McCray, P. B., Jr. (2005). Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors--design, biosafety, and production. *Gene Ther, 12*(14), 1089-1098.

279. Sirven, A., Pflumio, F., Zennou, V., Titeux, M., Vainchenker, W., Coulombel, L., et al. (2000). The human immunodeficiency virus type-1 central DNA flap is a crucial determinant for lentiviral vector nuclear import and gene transduction of human hematopoietic stem cells. *Blood, 96*(13), 4103-4110.

280. Smith, K. P., Liu, B., Scott, C., & Sharp, Z. D. (1995). Pit-1 exhibits a unique promoter spacing requirement for activation and synergism. *J Biol Chem, 270*(9), 4484-4491.

281. Smolen, J. S., & Steiner, G. (2003). Therapeutic strategies for rheumatoid arthritis. *Nat Rev Drug Discov, 2*(6), 473-488.

282. Solomon, L., Robin, G., & Valkenburg, H. A. (1975). Rheumatoid arthritis in an urban South African Negro population. *Ann Rheum Dis, 34*(2), 128-135.

283. Song, Y. W., & Kang, E. H. (2010). Autoantibodies in rheumatoid arthritis: rheumatoid factors and anticitrullinated protein antibodies. *QJM, 103*(3), 139-146.

284. Stanczyk, J., Pedrioli, D. M., Brentano, F., Sanchez-Pernaute, O., Kolling, C., Gay,

R. E., et al. (2008). Altered expression of MicroRNA in synovial fibroblasts and synovial tissue in rheumatoid arthritis. *Arthritis Rheum, 58*(4), 1001-1009.

285. Stemmer, W. P., Crameri, A., Ha, K. D., Brennan, T. M., & Heyneker, H. L. (1995). Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene, 164*(1), 49-53.

286. Subang, M. C., Fatah, R., Bright, C., Blanco, P., Berenstein, M., Wu, Y., et al. (2012). A novel hybrid promoter responsive to pathophysiological and pharmacological regulation. *J Mol Med (Berl), 90*(4), 401-411.

287. Sukhatme, V. P. (1990). Early transcriptional events in cell growth: the Egr family. *J Am Soc Nephrol, 1*(6), 859-866.

288. Sukhatme, V. P., Cao, X. M., Chang, L. C., Tsai-Morris, C. H., Stamenkovich, D., Ferreira, P. C., et al. (1988). A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization. *Cell, 53*(1), 37-43.

289. Sundquist, W. I., & Krausslich, H. G. (2012). HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med, 2*(7), a006924.

290. Suresh, E. (2004). Diagnosis of early rheumatoid arthritis: what the non-specialist needs to know. *J R Soc Med, 97*(9), 421-424.

291. Suzuki, T., Sakurai, F., Nakamura, S., Kouyama, E., Kawabata, K., Kondoh, M., et al. (2008). miR-122a-regulated expression of a suicide gene prevents hepatotoxicity without altering antitumor effects in suicide gene therapy. *Mol Ther, 16*(10), 1719-1726.

292. Svaren, J., Sevetson, B. R., Apel, E. D., Zimonjic, D. B., Popescu, N. C., & Milbrandt, J. (1996). NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by proliferative and differentiative stimuli. *Mol Cell Biol, 16*(7), 3545-3553.

293. Symmons, D., Turner, G., Webb, R., Asten, P., Barrett, E., Lunt, M., et al. (2002). The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology (Oxford), 41*(7), 793-800.

294. Szekanecz, Z., & Koch, A. E. (2000). Cell-cell interactions in synovitis. Endothelial cells and immune cell migration. *Arthritis Res, 2*(5), 368-373.

295. Taganov, K. D., Boldin, M. P., Chang, K. J., & Baltimore, D. (2006). NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc Natl Acad Sci U S A, 103*(33), 12481-12486.

296. Tak, P. P., & Firestein, G. S. (2001). NF-kappaB: a key role in inflammatory diseases. *J Clin Invest, 107*(1), 7-11.

297. Tak, P. P., Gerlag, D. M., Aupperle, K. R., van de Geest, D. A., Overbeek, M.,

Bennett, B. L., et al. (2001). Inhibitor of nuclear factor kappaB kinase beta is a key regulator of synovial inflammation. *Arthritis Rheum, 44*(8), 1897-1907.

298. Takahashi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M., & Chambon, P. (1986). Requirement of stereospecific alignments for initiation from the simian virus 40 early promoter. *Nature, 319*(6049), 121-126.

299. TerMaat, J. R., Pienaar, E., Whitney, S. E., Mamedov, T. G., & Subramanian, A. (2009). Gene synthesis by integrated polymerase chain assembly and PCR amplification using a high-speed thermocycler. *J Microbiol Methods, 79*(3), 295-300.

300. Thomas, M. C., & Chiang, C. M. (2006). The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol, 41*(3), 105-178.

301. Thomas, R. S., Tymms, M. J., McKinlay, L. H., Shannon, M. F., Seth, A., & Kola, I. (1997). ETS1, NFkappaB and AP1 synergistically transactivate the human GM-CSF promoter. *Oncogene, 14*(23), 2845-2855.

302. Tremblay, J. J., & Drouin, J. (1999). Egr-1 is a downstream effector of GnRH and synergizes by direct interaction with Ptx1 and SF-1 to enhance luteinizing hormone beta gene transcription. *Mol Cell Biol, 19*(4), 2567-2576.

303. Treuhaft, P. S., & DJ, M. C. (1971). Synovial fluid pH, lactate, oxygen and carbon dioxide partial pressure in various joint diseases. *Arthritis Rheum, 14*(4), 475-484.

304. Upchurch, K. S., & Kay, J. (2012). Evolution of treatment for rheumatoid arthritis. *Rheumatology (Oxford), 51 Suppl 6*, vi28-36.

305. Urban, M. B., & Baeuerle, P. A. (1990). The 65-kD subunit of NF-kappa B is a receptor for I kappa B and a modulator of DNA-binding specificity. *Genes Dev, 4*(11), 1975-1984.

306. Urlinger, S., Baron, U., Thellmann, M., Hasan, M. T., Bujard, H., & Hillen, W. (2000). Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci U S A, 97*(14), 7963-7968.

307. Valoczi, A., Hornyik, C., Varga, N., Burgyan, J., Kauppinen, S., & Havelda, Z. (2004). Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. *Nucleic Acids Res, 32*(22), e175.

308. van de Loo, F. A., de Hooge, A. S., Smeets, R. L., Bakker, A. C., Bennink, M. B., Arntz, O. J., et al. (2004). An inflammation-inducible adenoviral expression system for local treatment of the arthritic joint. *Gene Ther, 11*(7), 581-590.

309. van de Loo, F. A., & van den Berg, W. B. (2002). Gene therapy for rheumatoid arthritis. Lessons from animal models, including studies on interleukin-4, interleukin-10, and interleukin-1 receptor antagonist as potential disease

modulators. *Rheum Dis Clin North Am, 28*(1), 127-149.

310. van de Loo, A. A., Arntz, O. J., Bakker, A. C., van Lent, P. L., Jacobs, M. J., & van den Berg, W. B. (1995). Role of interleukin 1 in antigen-induced exacerbations of murine arthritis. *Am J Pathol, 146*(1), 239-249.

311. van Vollenhoven, R. F. (2009). Treatment of rheumatoid arthritis: state of the art 2009. *Nat Rev Rheumatol, 5*(10), 531-541.

312. Varley, A. W., Geiszler, S. M., Gaynor, R. B., & Munford, R. S. (1997). A two-component expression system that responds to inflammatory stimuli in vivo. *Nat Biotechnol, 15*(10), 1002-1006.

313. Verma, I. M., & Weitzman, M. D. (2005). Gene therapy: twenty-first century medicine. *Annu Rev Biochem, 74*, 711-738.

314. Wagner, E. F., & Matsuo, K. (2003). Signalling in osteoclasts and the role of Fos/AP1 proteins. *Ann Rheum Dis, 62 Suppl 2*, ii83-85.

315. Walsh, D. A., & Pearson, C. I. (2001). Angiogenesis in the pathogenesis of inflammatory joint and lung diseases. *Arthritis Res, 3*(3), 147-153.

316. Wang, G. L., Jiang, B. H., Rue, E. A., & Semenza, G. L. (1995). Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc Natl Acad Sci U S A, 92*(12), 5510-5514.

317. Wang, J. C. (1979). Helical repeat of DNA in solution. *Proc Natl Acad Sci U S A, 76*(1), 200-203.

318. Wang, G. L., & Semenza, G. L. (1993). Characterization of hypoxia-inducible factor 1 and regulation of DNA binding activity by hypoxia. *J Biol Chem, 268*(29), 21513-21518.

319. Wang, Y., & Stumph, W. E. (1995). RNA polymerase II/III transcription specificity determined by TATA box orientation. *Proc Natl Acad Sci U S A, 92*(19), 8606-8610.

320. Watson, D. K., McWilliams-Smith, M. J., Nunn, M. F., Duesberg, P. H., O'Brien, S. J., & Papas, T. S. (1985). The ets sequence from the transforming gene of avian erythroblastosis virus, E26, has unique domains on human chromosomes 11 and 21: both loci are transcriptionally active. *Proc Natl Acad Sci U S A, 82*(21), 7294-7298.

321. Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature, 447*(7145), 661-678.

322. Wernert, N., Justen, H. P., Rothe, M., Behrens, P., Dreschers, S., Neuhaus, T., et al. (2002). The Ets 1 transcription factor is upregulated during inflammatory angiogenesis in rheumatoid arthritis. *J Mol Med (Berl), 80*(4), 258-266.

323. Wernert, N., Raes, M. B., Lassalle, P., Dehouck, M. P., Gosselin, B., Vandenbunder, B., et al. (1992). c-ets1 proto-oncogene is a transcription factor expressed in endothelial cells during tumor vascularization and other forms of angiogenesis in humans. *Am J Pathol, 140*(1), 119-127.

324. Wernert, N., Stanjek, A., Kiriakidis, S., Hugel, A., Jha, H. C., Mazitschek, R., et al. (1999). Inhibition of Angiogenesis In Vivo by ets-1 Antisense Oligonucleotides-Inhibition of Ets-1 Transcription Factor Expression by the Antibiotic Fumagillin. *Angew Chem Int Ed Engl, 38*(21), 3228-3231.

325. Wightman, B., Burglin, T. R., Gatto, J., Arasu, P., & Ruvkun, G. (1991). Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. *Genes Dev, 5*(10), 1813-1824.

326. Wightman, B., Ha, I., & Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell, 75*(5), 855-862.

327. Wingender, E., Chen, X., Hehl, R., Karas, H., Liebich, I., Matys, V., et al. (2000). TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res, 28*(1), 316-319.

328. Winter, C. A., Risley, E. A., & Nuss, G. W. (1962). Carrageenin-induced edema in hind paw of the rat as an assay for antiiflammatory drugs. *Proc Soc Exp Biol Med, 111*, 544-547.

329. Wittkopp, P. J. (2010). Variable transcription factor binding: a mechanism of evolutionary change. *PLoS Biol, 8*(3), e1000342.

330. Wolff, J. A., Dowty, M. E., Jiao, S., Repetto, G., Berg, R. K., Ludtke, J. J., et al. (1992). Expression of naked plasmids by cultured myotubes and entry of plasmids into T tubules and caveolae of mammalian skeletal muscle. *J Cell Sci, 103 (Pt 4)*, 1249-1259.

331. Wu, G., Wolf, J. B., Ibrahim, A. F., Vadasz, S., Gunasinghe, M., & Freeland, S. J. (2006). Simplified gene synthesis: a one-step approach to PCR-based gene construction. *J Biotechnol, 124*(3), 496-503.

332. Wu, L., & Berk, A. (1988). Constraints on spacing between transcription factor binding sites in a simple adenovirus promoter. *Genes Dev, 2*(4), 403-411.

333. Yi, R., Qin, Y., Macara, I. G., & Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev, 17*(24), 3011-3016.

334. Yin, J. Q., Zhao, R. C., & Morris, K. V. (2008). Profiling microRNA expression with microarrays. *Trends Biotechnol, 26*(2), 70-76.

335. Zhou, Y., Ferguson, J., Chang, J. T., & Kluger, Y. (2007). Inter- and intra-combinatorial regulation by transcription factors and microRNAs. *BMC Genomics, 8*, 396.

336. Zhu, S., Pan, W., Song, X., Liu, Y., Shao, X., Tang, Y., et al. (2012). The microRNA miR-23b suppresses IL-17-associated autoimmune inflammation by targeting TAB2, TAB3 and IKK-alpha. *Nat Med, 18*(7), 1077-1086.

337. Zufferey, R., Donello, J. E., Trono, D., & Hope, T. J. (1999). Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J Virol, 73*(4), 2886-2892.

338. Zufferey, R., Dull, T., Mandel, R. J., Bukovsky, A., Quiroz, D., Naldini, L., et al. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J Virol, 72*(12), 9873-9880.

339. Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L., & Trono, D. (1997). Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol, 15*(9), 871-875.

Web sources

1. Figure 1.1 (Chapter 1): Photograph courtesy of James Heilman, MD, CCFP-EM, Clinical Faculty member of the Department of Emergency Medicine, Faculty of Medicine, University of British Columbia, via Wikimedia Commons. This photograph is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported licence, which grants permission to copy this document under the terms of the GNU Free Documentation License, Version 1.2). The image and licensing can be found at: http://commons.wikimedia.org/wiki/File:Rheumatoid_Arthritis.JPG

2. Figure 1.4 (Chapter 1) was redrawn, with modifications, from the image by: http://www.daviddarling.info/encyclopedia/A/acquired_immune_deficiency_syndrome.html

3. Section 2.7 (Chapter 2): The Assembly PCR method was modified from the online method of Team Heidelberg, International Genetically Engineered Machine (iGEM) Competition, 2009: http://2009.igem.org/Team:Heidelberg/Project_Synthetic_promoters

4. MicroRNA sequences were obtained from the online miRBase repository: www.mirbase.org

5. The pLV.CMVenh.gp91.eGFP.cHS4, plasmid was purchased from Addgene: www.addgene.org; plasmid 30471

6. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor (2010). NICE technology appraisal guidance 195; http://www.nice.org.uk/guidance/ta195

7. Rheumatoid arthritis: The management of rheumatoid arthritis in adults (2009). NICE clinical guideline 79; http://www.nice.org.uk/guidance/cg79

Chapter 10: Appendices

Appendix 1. Table of sequencing primers

The selected constructs were sequenced using one of the sequencing primers listed in Appendix Table 1.

Primer (5'-3')	Sequencing direction	Binding location
GL2 primer 5'-CTTTATGTTTTTGGCGTCTTCCA-3'	Reverse (towards promoter)	Binds to 5'-portion of luciferase, downstream of promoter
Forward pCpG primer 5'-TGGTTTGTCCAAACTCATCAA-3'	Forward (towards promoter)	Binds to pCpG vector, upstream of promoter
Forward Lenti primer 5'-TAGTAGACATAATAGCAACAGAC-3'	Forward (towards promoter)	Binds to lentivector, upstream of promoter
End of Luc$^+$ primer 5'-GGAAAACTCGACGCAAGAAA-3'	Forward (towards miRNA-target)	Binds to 3'-portion of luciferase, upstream of miRNA-target sites

Appendix 2. Sequencing data of the pGL3mCMV construct

The pGL3mCMV construct was sequenced using the GL2 reverse primer. The TATAA box is highlighted in bold and the XhoI (5'-TCGAG) and HindIII (5'-AGCT) overhangs are underlined and represent the start and end of the mCMV promoter.

```
  1 GAGCTGATTT AACAAAAATT TAACGCGAAT TTTAACAAAA TATTAACGCT

 51 TACAATTTGC CATTCGCCAT TCAGGCTGCG CAACYGTTGG GAAGGGCGAT

101 CGGTGCGGGC CTCTTCGCTA TTACGCCAGC CCAAGCTACC ATGATAAGTA

151 AGTAATATTA AGGTACGGGA GGTACTTGGA GCGGCCGCAA TAAAATATCT

201 TTATTTTCAT TACATCTGTG TGTTGGTTTT TTGTGTGAAT CGATAGTACT

251 AACATACGCT CTCCATCAAA ACAAAACGAA ACAAAACAAA CTAGCAAAAT

301 AGGCTGTCCC CAGTGCAAGT GCAGGTGCCA GAACATTTCT CTATCGATAG

351 GTACCGAGCT CTTACGCGTG CTAGCCCGGG CTCGAGGCCT GTAGGCGTGT

401 ACGGTGGGAG GCTTATATAA GCAGAGCTCA AGCTGGCATC CGTACKTGAG

451 CACTGT
```

Appendix 3. Sequencing data of selected pGL3-4bp-composite synthetic promoters with 4 bp space between TFBSs

The pGL3-4bp-composite promoters were sequenced using the GL2 reverse primer. The TFBSs are indicated by a box with an arrow to show the TFBS orientation. The NheI overhang (5'-CTAG) and XhoI site (5'-CTCGAG) is underlined and represents the start and end of the region containing multimerised TFBSs, respectively. The TATAA box is highlighted in bold. Unfortunately, some promoter sequences are incomplete (for unknown reasons). Partial or complete sequences are schematically represented for each construct.

Appendix 3.1. Promoter 24 - pGL3-4bp-composite promoter (complete)

```
  1 GTGAATCGAT AGTACTAACA TACGCTCTCC ATCAAAACAA AACGAAACAA
 51 AACAAACTAG CAAAATAGGC TGTCCCCAGT GCAAGTGCAG GTGCCAGAAC
101 ATTTCTCTAT CGATAGGTAC CGAGCTCTTA CGCGTGCTAG ACGTGGCTAG
151 ACGTGGCTAG CCACGTCTAG GGAATTTCCT AGTGAGTCAC TAGGGAACTT
201 CCGGCTAGTG AGTCACTAGC CACGTCTAGC CCGGGCTCGA GGCCTGTAGG
251 CGTGTACGGT GGGAGGCTTA TATAAGCAGA GCTCAAGCTG GCATCCGTAC
```


promoter 24

Appendix 3.2. Promoter 14 - pGL3-4bp-composite promoter (incomplete)

```
  1 CTAGTGGCCA CTGTATATGC GGCTCCGGAC TTACRCASAW TGCCCAATCT
 51 AGCCCCGTRT ACATTGCGCA ATCTAGCCAC GTCTAGCCAC GTCTAGTGAC
101 TCACTAGACG TGGCTAGGGA ATTTCCTAGG AAATTCCCTA GCCCGGGCTC
151 GAGGCCTGTA GGCGTGTACG GTGGGAGGCT TATATAAGCA GAGCTCAAGC
201 TGGCATCCGT ACKTTGAGCC MCC
```


promoter 14

Appendix 3.3. Promoter 40 - pGL3-4bp-composite promoter (complete)

```
  1 GATAGTACTA ACATACGCTC TCCATCAAAA CAAAACGAAA CAAAACAAAC
 51 TAGCAAAATA GGCTGTCCCC AGTGCAAGTG CAGGTGCCAG AACATTTCTC
101 TATCGATAGG TACCGAGCTC TTACGCGTGC TAGACGTGGC TAGTTGCGTG
151 GGCGTCTAGT GACTCACTAG CCCGGGCTCG AGGCCTGTAG GCGTGTACGG
201 TGGGAGGCTT ATATAAGCAG AGCTCAAGCT GGCATCCGGT ACKTGAGCCA
```


promoter 40

Appendix 3.4. Promoter 134 - pGL3-4bp-composite promoter (incomplete)

```
  1 TTAGGAAATC CCCTTKGACT CAWTAGASTT GGGGACTGAC TCACTAGACG
 51 TGGCTAGATT GCGCAATCTA GCCGGAAGTT CCCTAGCCAC GTCTAGTGAG
101 TCACTAGACG CCCACGCAAC TAGTGACTCA CTAGACGTGG CTAGTGACTC
151 ACTAGACGTG GCTAGTGACT CACTAGCCGG AAGTTCCCTA GACGCCCACG
201 CAACTAGCCC GGGCTCGAGG CCTGTAGGCG TGTACGGTGG GAGGCTTATA
251 TAAGCAGAGC TCAAGCTGGC ATCCGGTACT GTKAGCCMMC CC
```


promoter 134

Appendix 3.5. Promoter 148 - pGL3-4bp-composite promoter (incomplete)

```
  1 CTAAGTAAAC CGTGTTGGAG TTTGAGAAGG AGCGAGGTCY GARAATTTCT
 51 AGACCCATGG GTAKKCRTGT ATYGCSGGAG GATAYYTAGC AGTGATRGGG
101 AATTTCCTAG CCCGGGCTCG AGGCCTGTAG GCGTGTACGG TGGGAGGCTT
151 ATATAAGCAG AGCTCAAGCT GGCATCCGGT ACKTGAGRCA ACCC
```


promoter 148

Appendix 3.6. Promoter 150 - pGL3-4bp-composite promoter (incomplete)

```
  1 YTTTYTTTTT TTGGCGTCTT CCATGGTGGC TTTACCAACA GTACCGGAAT

 51 GCCAAGCTTG AGCTCTGCTT ATATAAGCCT CCCACCGTAC ACGCCTACAG

101 GCCTCGAGCC CGGGCTAGTG ACTCAMTTRG GAACTTCCGG CTAGYCACGC

151 ATAAGTGAGT CTCTGAAATT SSCTAGATTG CGCAATCTAG CCACGTCTAG

201 ATTGCGCAAT CTAGCCACGT CTAGTGACTC ACTAGTGAGT CACTAGCCGG

251 AAGTTCCCTA GTGAGTCACT AGCCCGGGCT CGAGGCCTGT AGGCGTGTAC

301 GGTGGGAGGC TTATATAAGC AGAGCTCAAG CTGGCATCCG TACGTKAGCC
```

← ← ← → ← → ← ← → → →

| AP-1 | Ets1 | NFκB | C/EBPβ | HRE | C/EBPβ | HRE | AP-1 | AP-1 | Ets1 | AP-1 | mCMV | promoter 150 |

Appendix 3.7. Promoter 127 - pGL3-4bp-composite promoter (incomplete)

```
  1 GCGTGGGCGT CTAGTYSCGT GGGCGTCTAG ACGTGGCTAG TGAGTCACTA

 51 GTGAGTCACT AGGGAATTTC CTAGCCACGT CTAGTGACTC ACTAGATTGC

101 GCAATCTAGG GAATTTCCTA GTGACTCACT AGGGAACTTC CGGCTAGTTG

151 CGTGGGCGTC TAGCCACGTC TAGTGACTCA CTAGACGTGG CTAGACGTGG

201 CTAGTGACTC ACTAGCCCGG GCTCGAGGCC TGTAGGCGTG TACGGTGGGA

251 GGCTTATATA AGCAGAGCTC AAGCTGGCAT CCGGTACKTG AGCCMCC
```

→ → → → → ← ← → → ← ← → ← ← → → ←

| Egr1 | HRE | AP-1 | AP-1 | NFκB | HRE | AP-1 | C/EBPβ | NFκB | AP-1 | Ets1 | Egr1 | HRE | AP-1 | HRE | HRE | AP-1 | mCMV | promoter 127 |

Appendix 3.8. Promoter 96 - pGL3-4bp-composite promoter (incomplete)

```
  1 CCTAGCCACG TCTAGGGAAC TTCCGGCTAG TGAGTCACTA GACGTGGCTA

 51 GCCACGTCTA GACGCCCACG CAACTAGGAA ATTCCCTAGT GACTCACTAG

101 CCCGGGCTCG AGGCCTGTAG GCGTGTACGG TGGGAGGCTT ATATAAGCAG

151 AGCTCAAGCT GGCATCCGTA CKTGAGCACG C
```

← ← → → ← ← ← ←

| HRE | Ets1 | AP-1 | HRE | HRE | Egr1 | NFκB | AP-1 | mCMV | promoter 96 |

357

Appendix 3.9. Promoter 78 - pGL3-4bp-composite promoter (complete)

```
  1 AGTACTAACA TACGCTCTCC ATCAAAACAA AACGAAACAA AACAAACTAG

 51 CAAAATAGGC TGTCCCCAGT GCAAGTGCAG GTGCCAGAAC ATTTCTCTAT

101 CGATAGGTAC CGAGCTCTTA CGCGTGCTAG ACGTGGCTAG GAAATTCCCT

151 AGGAAATTCC CTAGCCGGAA GTTCCCTAGC CACGTCTAGG GAATTTCCTA

201 GCCCGGGCTC GAGGCCTGTA GGCGTGTACG GTGGGAGGCT TATATAAGCA

251 GAGCTCAAGC KGGCWYTCKS GGWMGMTMMT GGGWCACCG
```


→ ← ← → ← →
HRE NFκB NFκB Ets1 AP-1 NFκB mCMV   promoter 78

Appendix 3.10. Promoter 173 - pGL3-4bp-composite promoter (incomplete)

```
  1 CTAGWCACGG CTAAWTACGT CTAGCCACGT MTAGACGTGG CTAGGWAATT

 51 CCMTAGGGAA TTTCCTARGA AATTCCCTAG GAAATTCCCT AGACGTGGCT

101 AGGAATTTCC TAGGAAATTC CCTAGTGAGT CACTAGATTS CGCAATMTAG

151 TGACTCACTA GTTGCGTGGG CGTCTAGGGA ATTTCCTARA CGYGGCTAGA

201 CGTGGCTAGC CACGTCTAGA CGCCCACGCA ACTAGCCCGG GCTCGAGGCC

251 TGTAGGCGTG TACGGTGGGA GGCTTATATA AGCAGAGCTC AAGCTGGCAT
```


← → → → ← ← → ← ← → → ← → → → → ← ←
HRE HRE NFκB NFκB NFκB NFκB HRE NFκB NFκB AP-1 C/EBPβ AP-1 Egr1 NFκB HRE HRE HRE Egr1 mCMV   promoter 173

Appendix 3.11. Promoter 186 - pGL3-4bp-composite promoter (incomplete)

```
  1 TTTTTGTGTG AATCGATAGT ACTAACATAC GCTCTCCATC AAAACAAAAC

 51 GAAACAAAAC AAACTAGCAA AATAGGCTGT CCCCAGTGCA AGTGCAGGYG

101 CCAGAACAAC TAGCCACGTC TAGCCMCGTS TWGATTGSGC AATCTAGKGA

151 GTCACTAGAT TGCGCAATCT AGGAAATTCC CTAGTGACTC ACTAGCCCGG

201 GCTCGAGGCC TGTAGGCGTG TACGGTGGGA GGCTTATATA AGCAGAGCTC

251 AAGCTGGCAT CCGTACGTGA GCAMCCC
```


← ← → → → ← ←
HRE HRE C/EBPβ AP-1 C/EBPβ NFκB AP-1 mCMV   promoter 186

Appendix 3.12. Promoter 176 - pGL3-4bp-composite promoter (complete)

```
  1 CAAACTAGCA AAATAGGCTG TCCCCAGTGC AAGTGCAGGT GCCAGAACAT

 51 TTCTCTATCG ATAGGTACCG AGCTCTTACG CGTGCTAGGA AATTCCMTAG

101 TGACTCACTA GACGTGSCTA GCCACGTCTA GACGTGGCTA GATTGCGCAT

151 CTAGTGACTC AYTAGTGAGT CACTAGACGT GGCTAGCCAC GTCTAGGGAA

201 TTTCCTAGTG ACTCACTAGA CGTGGCTAGG AAATTCCCTA GCCCGGGCTC

251 GAGGCCTGTA GGCGTGTACG GTGGGAGGCT TATATAAGCA GAGCTCAAGC

301 TGGCATCCGG TACKTGARCA CTC
```
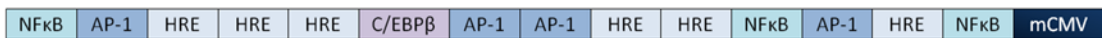

promoter 176

Appendix 3.13. Promoter 41- pGL3-4bp-composite promoter (complete)

```
  1 ATAGGCTGTC CCCAGTGCAA GTGCAGGTGC CAGAACATTT CTCTATCGAT

 51 AGGTACCGAG CTCTTACGCG TGCTAGACGT GGCTAGGGAA CTTCCGGCTA

101 GTGAGTCACT AGTGACTCAC TAGATTGCGC AATCTAGCCA CGTCTAGTGA

151 GTCACTAGCC CGGGCTCGAG GCCTGTAGGC GTGTACGGTG GGAGGCTTAT

201 ATAAGCAGAG CTCAAGCTGG CATCCGTACK TTAAGCCMCC TC
```
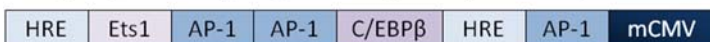

promoter 41

Appendix 3.14. Promoter 19 - pGL3-4bp-composite promoter (complete)

```
  1 CTGTCCCCAG TGCAAGTGCA GGTGCCAGAA CATTTCTCTA TCGATAGGTA

 51 CCGAGCTCTT ACGCGTGCTA GGAAATTCCC TAGGGAATTT CCTAGACGTG

101 GCTAGTGACT CACTAGACGT GGCTAGCCGG AAGTTCCCTA GCCGGAAGTT

151 CCCTAGTGAC TCACTAGACG CCCACGCAAC TAGTGAGTCA CTAGCCGGAA

201 GTTCCCTAGC CACGTCTAGT GACTCACTAG TGACTCACTA GTGAGTCACT

251 AGTGAGTCAC TAGGAAATTC CCTAGTGACT CACTAGCCCG GCTCGAGGC

301 CTGTAGGCGT GTACGGTGGG AGGCTTATAT AAGCAGAGCT CAAGCTGGCA
```
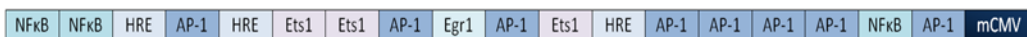

promoter 19

Appendix 3.15. Promoter 125 - pGL3-4bp-composite promoter (incomplete)

```
  1 AACATTTCTC TATCGATAGG TACCGAGYTM TTCCGCGTCC TAGCWCGCGC
 51 AATCTACCCA MGTCCTAGTG AGTCACTAGG AAATTCCCTA GCCCGGTCTA
101 KSCMYSWCTA GATTGCGCAA TCTAGTGAGT CACTAGACGC CCACGCAACT
151 AGTGAGTCAC TAGCCCGGGC TCGAGGCCTG TAGGCGTGTA CGGTGGGAGG
201 CTTATATAAG CAGAGCTCAA GCTGGCATCC GTACKTRATC CAT
```


promoter 125

Appendix 3.16. Promoter 118 - pGL3-4bp-composite promoter (incomplete)

```
  1 GAGCGCGKTG CAATTAATGG GGGYAGCATC CCTCWTCCGG CAAAACAGRA
 51 GTTACCAAAK CRGGGTAGTT CCCTAGTGAC TCACTAGACG TGGCTAGATT
101 GCGCAATCTA GTGAGTCACT AGTGACTCAC TAGCCGGAAG TTCCCTAGTG
151 ACTCACTAGC CCGGGCTCGA GGCCTGTAGG CGTGTACGGT GGGAGGCTTA
201 TATAAGCAGA GCTCAAGCTG GCATCCGTAC GTGAGCCMCG C
```


promoter 118

Appendix 3.17. Promoter 4 - pGL3-4bp-composite promoter (complete)

```
  1 AACAAAACGA AACAAAACAA ACTAGCAAAA TAGGCTGTCC CCAGTGCAAG
 51 TGCAGGTGCC AGAACATTTC TCTATCGATA GGTACCGAGC TCTTACGCGT
101 GCTAGGGAAT TTCCTAGGGA ATTTCCTAGC CGGAAGTTCC CTAGCCCGGG
151 CTCGAGGCCT GTAGGCGTGT ACGGTGGGAG GCTTATATAA GCAGAGCTCA
201 AGCTGGCATC CGGTACKTGA GCACC
```


promoter 4

Appendix 3.18. Promoter 13 - pGL3-4bp-composite promoter (complete)

```
  1 TCCCCAGTGC AAGTGCAGGT GCCAGAACAT TTCTCTATCG ATAGGTACCG
 51 AGCTCTTACG CGTGCTAGTT GCGCAATCTA GCCGGAAGTT CCCTAGCCCG
101 GGCTCGAGGC CTGTAGGCGT GTACGGTGGG AGGCTTATAT AAGCAGAGCT
151 CAAGCTGGCA TCCGTACKTG AGCMCCC
```


C/EBPβ | Ets1 | mCMV   promoter 13

Appendix 3.19. Promoter 17 - pGL3-4bp-composite promoter (incomplete)

```
  1 CGGAAGTTCC CTAGTTGCGT GGGCGTCTAG GAAATTCCCT AGGGAATTTC
 51 CTAGGGAACT TCCGGCTAGT GACTCACTAG GGAATTTCCT AGACGTGGCT
101 AGGAAATTCC CTAGCCCGGG CTCGAGGCCT GTAGGCGTGT ACGGTGGGAG
151 GCTTATATAA GCAGAGCTCA AGCTTGGCAT CCGTACTGTG AGCCACCG
```


Egr1 | NFκB | NFκB | Ets1 | AP-1 | NFκB | HRE | NFκB | mCMV   promoter 17
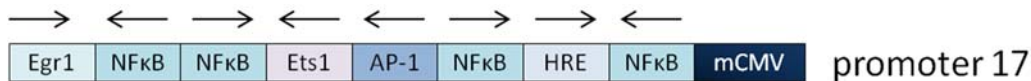
Appendix 3.20. Promoter 21 - pGL3-4bp-composite promoter (complete)

```
  1 CAAACTAGCA AAATAGGCTG TCCCCAGTGC AAGTGCAGGT GCCAGAACAT
 51 TTCTCTATCG ATAGGTACCG AGCTCTTACG CGTGCTAGAT TGCGCAATCT
101 AGCCGGAAGT TCCCTAGTGA CTCACTAGAC GTGGCTAGCC ACGTCTAGCC
151 ACGTCTAGTG ACTCACTAGC CCGGGCTCGA GGCCTGTAGG CGTGTACGGT
201 GGGAGGCTTA TATAAGCAGA GCTCAAGCTK GCATCCCSGG ACCKKGCCCC
```


C/EBPβ | Ets1 | AP-1 | HRE | HRE | HRE | AP-1 | mCMV   promoter 21

Appendix 3.21. Promoter 22 - pGL3-4bp-composite promoter (complete)

```
  1 AAACAAAACG AAACAAAACA AACTAGCAAA ATAGGCTGTC CCCAGTGCAA
 51 GTGCAGGTGC CAGAACATTT CTCTATCGAT AGGTACCGAG CTCTTACGCG
101 TGCTAGGGAA TTTCCTAGCC CGGGCTCGAG GCCTGTAGGC GTGTACGGTG
```

```
151 GGAGGCTTAT ATAAGCAGAG CTCAAGCTGG CATCCGGTAC GTGAGCAACT
```

→
NFκB mCMV promoter 22

Appendix 3.22. Promoter 204 - pGL3-4bp-composite promoter (complete)

```
  1 ATTTCTCTAT CGATAGGTAC CGAGCTCTTA CGCGTGCTAGACGTGCTAGC

 51 CCGGGCTCGA GGCCTGTAGG CGTGTACGGT GGGAGGCTTA TATAAGCAGA

101 GCTCAAGCTG GCATCCGGTA CKTKGAGGCC MMCCC
```

→
HRE mCMV promoter 204

Appendix 3.23. Promoter 207 - pGL3-4bp-composite promoter (complete)

```
  1 GGTACCGAGC TCTTACGCGT GCTAGACGTG GCTAGCCCGG GCTCGAGGCC

 51 TGTASGCGTG TACGGTGGGA GGCTTATATA AGCAGAGCTC AAGCTTGGCA

101 TTCCGGTACT GTTGGTAAAG CCACCATGGA AGACGCCAAA AACATAAAGA
```

→
HRE mCMV promoter 207

Appendix 3.24. Promoter 225 - pGL3-4bp-composite promoter (complete)

```
  1 AAACGAAACA AAACAAACTA GCAAAATAGG CTGTCCCCAG TGCAAGTGCA

 51 GGTGCCAGAA CATTTCTCTA TCGATAGGTA CCGAGCTCTT ACGCGTGCTA

101 GGGAATTTCC TAGTGAGTCA CTAGGGAATT TCCTAGCCGG AAGTTCCCTA

151 GATTGCGCAA TCTAGTGACT CACTAGATTG CGCAATCTAG ACGCCCACGC

201 AACTAGCCAC GTCTAGCCAC GTCTAGGGAA TTTCCTAGAT TGCGCAATCT

251 AGCCACGTCT AGCCACGTCT AGCCACGTCT AGTGAGTCAC TAGACGTGGC

301 TAGGAAATTC CCTAGTTGCG TGGGCGTCCG GAAGTTCCCT AGCCACGTCT

351 AGCCCGGGCT CGAGGCCTGT AGGCGTGTAC GGTGGGAGGC TTATATAAGC

401 AGAGCTCAAG CNGCANCCGN NNGNNNNNNN NNNN
```

→ → → → → ← → ← ← ← → → ← ← ← → → ← → → ←
NFκB AP-1 NFκB Ets1 C/EBPβ AP-1 C/EBPβ Egr1 HRE HRE NFκB C/EBPβ HRE HRE HRE AP-1 HRE NFκB Egr1 Ets1 HRE mCMV promoter 225

Appendix 3.25. Promoter 245 - pGL3-4bp-composite promoter (complete)

```
  1 ATCGATAGTA CTAACATACG CTCTCCATCA AAACAAAACG AAACAAAACA

 51 AACTAGCAAA ATAGGCTGTC CCCAGTGCAA GTGCAGGTGC CAGAACATTT

101 CTCTATCGAT AGGTACCGAG CTCTTACGCG TGCTAGACGT GGCTAGACGT

151 GGCTAGGAAA TTCCCTAGGG AATTCCTAG CCCGGGCTCG AGGCCTGTAG

201 GCGTGTACGG TGGGAGGCTT ATATAAGCAG AGCTCAAGCT GCATCCGNNN
```

HRE → HRE → NFκB ← NFκB → mCMV promoter 245

Appendix 4. Sequencing data of the pCpGmCMV-Luc⁺ construct

The pCpGmCMV-Luc⁺ construct was sequenced using the GL2 reverse primer. The NheI restriction site (5'- GCTAGC; underlined) indicates the end of the pCpGmSEAP fragment (position 1 bp- 387 bp) and start of the pGL3mCMV fragment (position 387 bp – 473 bp).

```
  1 TGRTTTGWTM TAAATYTTGA TATTTAGTGG AACATTYTTT CSCATTTTST

 51 TMTACAAGAA TATTKWTGTT RTYGTCTTTT GGGCTTCTAT ATACATTTTA

101 GAATGAGGTT GGCAAGTTAA CAAACAGYTT TTWTGGGGTG AACATATTGA

151 CTGAATTCCA TACCACATTT GTAGAGGTTT TACTTGCTTT AAAAAACCTC

201 CCACACCTCC CCCTGAACCT GAAACATAAA ATGAATGCAA TTGTTGTTGT

251 TAACTTGTTT ATTGCAGCTT ATAATGGTTA CAAATAAAGC AATAGCATCA

301 CAAATTTCAC AAATAAAGCA TTTTTTTCAC TGCATTCTAG TTGTGGTTTG

351 TCCAAACTCA TCAATGTATC TTATCATGTC TGGCCAGCTA GCCCGGGCTC

401 GAGGCCTGTA GGCGTGTACG GTGGGAGGCT TATATAAGCA GAGCTCAAGC

451 TGGCATCCGT ACGTAGACMC YGC
```

Appendix 5. Sequencing data of the pCpG-variable-6b-NFκB and pCpG-variable-10NFκB constructs

Two pCpG-variable NFκB constructs were sequenced using the GL2 reverse primer. The variable NFκB motifs are indicated by a box and those similar to the fixed NFκB sequence (GGGACTTTCC) are indicated by an asterisk (*).The NheI (5'-GCTAGC) and XhoI (5'-CTCGAG) sites are underlined and represent the start and end of the region containing the multimerised variable NFκB motifs, respectively, and the TATAA box is highlighted in bold.

Appendix 5.1. pCpG-variable-6b-NFκB

```
  1 TAGCATCACA AATTTCACAA ATAAAGCATT TTWTTCAMTG CATTCTAGTT
 51 GTGGTTTGTC CAAACTCATC AATGTATCTT ATCATGTCTG GCCAGCTAGC
101 GTGCCTCTTA TGATCTGGAT CATGTGGGGT CTCCCACAAG ATCTCTGCGA
151 TGAACCTCAC CATGTGGGAG GTCCCACAAG GTGCCTCTTA TGATCTGGAT
201 CATGTGGGAA CTTCC*ACAAG ATCTCTGCGA TGAACCTCAC CATGTGGGAT
251 TTTCC*ACAAG GTGCCTCTTA TGATCTGGAT CATGTGGGAG GTCCCACAAG
301 ATCTCTGCGA TGAACCTCAC CATGTGGGAA TTCC*ACAAGG TGCCTCTTAT
351 GATCTGGATC TCGAGGCCTG TAGGCGTGTA CGGTGGGAGG CTTATATAAG
401 CAGAGCTCAA GCTTGGCATT CCGGTACKKT RAARRMMMMC CT
```


Appendix 5.2. pCpG-variable-10NFκB constructs

```
  1 GTATCTTATC ATGTCTGGCC AGCTAGCGTG CCTCTTATGA TCTGGATCAT
 51 GTGGGATATT CCACAAGATC TCTGCGATGA ACCTCACCAT TGGAGGTCCC
101 ACAAGGTGCC TCTTATGATC TGGATCATGT GGGGCTCTCC ACAAGATCTC
151 TGCGATGAAC CTCACCATGT GGGATACTCC ACAAGGTGCC TCTTATGATC
201 TGGATCATGT GGGGCTTTCC*ACAAGATCTC TGCGATGAAC CTCACCATGT
251 GGGGAGTTCC*CACAAGGTGC CTCTTATGAC CTGGATCATG TGGGAATTCC
301 C*ACAAGATCT CTGCGATGAA CCTCACCATG TGGGATTTCC C*ACAAGGTGC
351 CTCTTATGAT CTGGATCATG TGGGGGCCTC CACAAGATCT CTGCGATGAA
401 CCTCACCATG TGGGGAGCTC CACAAGGTGC CTCTTATGAT CTGGATCTCG
451 AGGCCTGTAG GCGTGTACGG TGGGAGGCTT ATATAAGCAG AGCTCAAGCT
501 KGCATCCGGT ACTKTWAAGS MMMMMCCT
```

Appendix 6. Sequencing data of selected pCpG-NFκB and pCpG-AP-1-responsive promoters with various spacing between TFBSs

The pCpG-NFκB and AP-1-responsive promoters were sequenced using the GL2 reverse primer. The TFBSs are positioned in the forward orientation and are indicated by a box. The NheI (5'-GCTAGC) and XhoI (5'-CTCGAG) sites are underlined and represent the start and end of the region containing multimerised TFBSs, respectively. The TATAA box is highlighted in bold.

Appendix 6.1. Promoter 2 - 15 bp space between 8NFκB motifs

```
  1 AACCTCCCAC ACCTCCCCCT GAACCTGAAA CATAAAATGA ATGCAATTGT
 51 TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA
101 GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT
151 GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGC CAGCTAGCGT
201 GCCTCTTATG ATCGGGACTT TCCTGCGATG AACCTCACGG GACTTTCCGT
251 GCCTCTTATG ATCGGGACTT TCCTGCGATG AACCTCACGG GACTTTCCGT
301 GCCTCTTATG ATCGGGACTT TCCTGCGATG AACCTCACGG GACTTTCCGT
351 GCCTCTTATG ATCGGGACTT TCCTGCGATG AACCTCACGG GACTTTCCGT
401 GCCTCTTATG ATCCTCGAGC CATGTCTGGT CGAGGCCTGT AGGCGTGTAC
451 GGTGGGAGGC TTATATAAGC AGAGCTCAAG CTGGCATCCG GTACKTKAAG
```

Appendix 6.2. Promoter 1 – 20 bp space between 6NFκB motifs

```
  1 CCCCMWGAAC CTGAAACATA AAATGAATGC AATTGTTGTT GTTAACTTGT
 51 TTATTGCAGC TTATAATGGT TACAAATAAA GCAATAGCAT CRCAAATTTC
101 ACAAATAAAG CATTTTTTTC ACTGCATTCT AGTTGTGGTT TGTCCAAACT
151 CATCAATGTA TCTTATCATG TCTGGCCAGC TAGCGTGCCT CTTATGATCT
201 GGATGGGACT TTCCATCTCT GCGATGAACC TCACGGGACT TTCCGTGCCT
251 CTTATGATCT GGATGGGACT TTCCATCTCT GCGATGAACC TCACGGGACT
301 TTCCGTGCCT CTTATGATCT GGATGGGACT TTCCATCTCT GCGATGAACC
351 TCACGGGACT TTCCGTGCCT CTTATGATCT GGATCTCGAG CCATGGTCGA
```

```
401 GGCCTGTAGG CGTGTACGGT GGGAGGCTTA **TATAA**GCAGA GCTCAAGCTG

451 GCATCCGGTA CTKTTAGACC AWWCC
```

Appendix 6.3. Promoter 5 – 30 bp space between 8NFκB motifs

```
  1 AAATTCCCCC CRAGATTTCC AAAAAATAAA RCCTTTTTTT CMCGCCATYM

 51 ARGTGGTGGT TKGYCCAAAC TTCATCAATG TATYTTAYCA TKTCKGGCCA

101 GCTAGCGTGC CTCTTATGAT CTGGATCATG TGGGACTTTC CACAAGATCT

151 CTGCGATGAA CCTCACCATG TGGGACTTTC CACAAGGTGC CTCTTATGAT

201 CTGGATCATG TGGGACTTTC CACAAGATCT CTGCGATGAA CCTCACCATG

251 TGGGACTTTC CACAAGGTGC CTCTTATGAT CTGGATCATG TGGGACTTTC

301 CACAAGATCT CTGCGATGAA CCTCACCATG TGGGACTTT CCACAAGGTG

351 CCTCTTATGA TCTGGATCAT GTGGGACTTT CCACAAGATC TCTGCGATGA

401 ACCTCACCAT GTGGGACTTT CCACAAGGTG CCTCTTATGA TCTGGATCTC

451 GAGTCGAGGC CTGTAGGCGT GTACGGTGGG AGGCTTA**TAT AA**GCAGAGCT

501 CAAGCTGGCA TCCGGTACTK TGAAGGCMMA CCTG
```

Appendix 6.4. Promoter 4 – 35 bp space between 5NFκB motifs

```
  1 TRAAACCTGA AAACATAAAA TGAATGCAAT TGTTGTTGTT AACTTGTTTA

 51 TTGCAGCTTA TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTTCACA

101 AATAAAGCAT TTTTTTCACT GCATTCTAGT TGTGGTTTGT CCAAACTCAT

151 CAATGTATCT TATCATGTCT GGCCAGCTAG CGTGCCTCTT ATGATCCCAT

201 GTGCGATGAA CCTCACGGGA CTTTCCGTGC CTCTTATGAT CCCATGTGCG

251 ATGAACCTCA CGGGACTTTC CGTGCCTCTT ATGATCCCAT GTGCGATGAA

301 CCTCACGGGA CTTTCCGTGC CTCTTATGAT CCCATGTGCG ATGAACCTCA

351 CGGGACTTTC CGTGCCTCTT ATGATCCCAT GTGCGATGAA CCTCACGGGA

401 CTTTCCGTGC CTCTTATGAT CCTCGAGCCA TGTCTGGTCG AGGCCTGTAG

451 GCGTGTACGG TGGGAGGCTT A**TATAA**GCAG AGCTCAAGCT GGCATTCCCG
```

Appendix 6.5. Promoter 2 – 40 bp space between 6NFκB motifs

```
  1 AAACCTCCCA CACCTCCCCC TGAACCTGAA ACATAAAATG AATGCAATTG
 51 TTGTTGTTAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT
101 AGCATCACAA ATTTCACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG
151 TGGTTTGTCC AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGCG
201 TGCCTCTTAT GATCCCATGT CTGTTGCGAT GAACCTCACG GGACTTTCCG
251 TGCCTCTTAT GATCCCATGT CTGTTGCGAT GAACCTCACG GGACTTTCCG
301 TGCCTCTTAT GATCCCATGT CTGTTGCGAT GAACCTCACG GGACTTTCCG
351 TGCCTCTTAT GATCCCATGT CTGTTGCGAT GAACCTCACG GGACTTTCCG
401 TGCCTCTTAT GATCCCATGT CTGTTGCGAT GAACCTCACG GGACTTTCCG
451 TGCCTCTTAT GATCCCATGT CTGTTGCGAT GAACCTCACG GGACTTTCCG
501 TGCCTCTTAT GATCCTCGAG CCATGTCTGG TCGAGGCCTG TAGGCGTGTA
551 CGGTGGGAGG CTTATATAAG CAGAGCTCAA GCTGGCATCC GGTACTGTTA
601 GRCMMCCCT
```

Appendix 6.6. Promoter 1 – 45 bp space between 7NFκB motifs

```
  1 CCMTCMCACA CCTCCCCCTG AACCTGAAAC ATAAAATGAA TGCAATTGTT
 51 GTTGTTTAAC TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA
101 GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT
151 GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGC CAGCTAGCGT
201 GCCTCTTATG ATCTGGATCC ATGATCTCTG CGATGAACCT CACGGGACTT
251 TCCGTGCCTC TTATGATCTG GATCCATGAT CTCTGCGATG AACCTCACGG
301 GACTTTCCGT GCCTCTTATG ATCTGGATCC ATGATCTCTG CGATGAACCT
351 CACGGGACTT TCCGTGCCTC TTATGATCTG GATCCATGAT CTCTGCGATG
401 AACCTCACGG GACTTTCCGT GCCTCTTATG ATCTGGATCC ATGATCTCTG
451 CGATGAACCT CACGGGACTT TCCGTGCCTC TTATGATCTG GATCCATGAT
501 CTCTGCGATG AACCTCACGG GACTTTCCGT GCCTCTTATG ATCTGGATCC
551 ATGATCTCTG CGATGAACCT CACGGGACTT TCCGTGCCTC TTATGATCTG
601 GATCTCGAGC CATGGTCGAG GCCTGTAGGC GTGTACGGTG GGAGGCTTAT
651 ATAAGCAGAG CTCAAGCTGG CATCCGGTWC KGTGAGGAAM CCG
```

367

Appendix 6.7. Promoter 5 – 60 bp space between 4NFκB motifs

```
  1 AAACCTCCCA CACCTCCCCC TGAACCTGAA ACATAAAATG AATGCAATTG
 51 TTGTTGTTAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT
101 AGCATCACAA ATTTCACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG
151 TGGTTTGTCC AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGCG
201 TGCCTCTTAT GATCTGGATC CATGTCTGTA TCTCTGCGAT GAACCTCACC
251 ATGTGGGACT TTCCACAAGG TGCCTCTTAT GATCTGGATC CATGTCTGTA
301 TCTCTGCGAT GAACCTCACC ATGTGGGACT TTCCACAAGG TGCCTCTTAT
351 GATCTGGATC CATGTCTGTA TCTCTGCGAT GAACCTCACC ATGTGGGACT
401 TTCCACAAGG TGCCTCTTAT GATCTGGATC CATGTCTGTA TCTCTGCGAT
451 GAACCTCACC ATGTGGGACT TTCCACAAGG TGCCTCTTAT ATCTGGATCT
501 CGAGTCGAGG CCTGTAGGCG TGTACGGTGG GAGGCTTATA TAAGCAGAGC
551 TCAAGCTGGC ATTCCGGTAC TKTKARRCCM CCC
```

Appendix 6.8. Promoter 4 - 15 bp space between 10AP-1 motifs

```
  1 WRRCMTCCCA CACCTCCCCM TGAACCTGAA ACATAAAATG AATGCAATTG
 51 TTGTTGTTAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT
101 AGCATCACAA ATTTCACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG
151 TGGTTTGTCC AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGCG
201 TGCCTCTTAT GATCTGAGTC ATGCGATGAA CCTCACTGAG TCAGTGCCTC
251 TTATGATCTG AGTCATGCGA TGAACCTCAC TGAGTCAGTG CCTCTTATGA
301 TCTGAGTCAT GCGATGAACC TCACTGAGTC AGTGCCTCTT ATGATCTGAG
351 TCATGCGATG AACCTCACTG AGTCAGTGCC TCTTATGATC TGAGTCATGC
401 GATGAACCTC ACTGAGTCAG TGCCTCTTAT GATCCTCGAG CCATGTCTGG
451 TCGAGGCCTG TAGGCGTGTA CGGTGGGAGG CTTATATAAG CAGAGCTCAA
501 GCTTGGCATC CGGTACTKTG AGCCCCCC
```

Appendix 6.9. Promoter 1 – 20 bp space between 8AP-1 motifs

```
  1 AACCTCCCAC ACCTCCCCCT GAACCTGAAA CATAAAATGA ATGCAATTGT
 51 TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA
101 GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT
151 GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGC CAGCTAGCGT
201 GCCTCTTATG ATCTGGATTG AGTCAATCTC TGCGATGAAC CTCACTGAGT
251 CAGTGCCTCT TATGATCTGG ATTGAGTCAA TCTCTGCGAT GAACCTCACT
301 GAGTCAGTGC CTCTTATGAT CTGGATTGAG TCAATCTCTG CGATGAACCT
351 CACTGAGTCA GTGCCTCTTA TGATCTGGAT TGAGTCAATC TCTGCGATGA
401 ACCTCACTGA GTCAGTGCCT CTTATGATCT GGATCTCGAG CCATGGTSGA
451 GGCCTGTAGG CGTGTACGGT GGGAGGCTTA TATAAGCAGA GCTCAAGCTG
501 GCATCCGGTA CKTGAGCCCC C
```

Appendix 6.10. Promoter 3 - 35 bp space between 7AP-1 motifs

```
  1 AARCSTCCCA CACCTCCCCC TGAACCTGAA ACATAAAATG AATGCAATTG
 51 TTGTTGTTAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT
101 AGCATCACAA ATTTCACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG
151 TGGTTTGTCC AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGCG
201 TGCCTCTTAT GATCCCATGT GCGATGAACC TCACTGAGTC AGTGCCTCTT
251 ATGATCCCAT GTGCGATGAA CCTCACTGAG TCAGTGCCTC TTATGATCCC
301 ATGTGCGATG AACCTCACTG AGTCAGTGCC TCTTATGATC CCATGTGCGA
351 TGAACCTCAC TGAGTCAGTG CCTCTTATGA TCCCATGTGC GATGAACCTC
401 ACTGAGTCAG TGCCTCTTAT GATCCCATGT GCGATGAACC TCACTGAGTC
451 AGTGCCTCTT ATGATCCCAT GTGCGATGAA CCTCACTGAG TCAGTGCCTC
501 TTATGATCCT CGAGCCATGT CTGGTCGAGG CCTGTAGGCG TGTACGGTGG
551 GAGGCTTATA TAAGCAGAGC TCAAGCTTGG CATCCGGTAC TGTGAGMCAM
```

Appendix 6.11. Promoter 3 – 40 bp space between 5AP-1 motifs

```
  1 AAMCSTCCCA CACCTCCCCC TGAACCTGAA ACATAAAATG AATGCAATTG
```

```
 51 TTGTTGTTAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGMAAT
101 AGCATCACAA ATTTCACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG
151 TGGTTTGTCC AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGCG
201 TGCCTCTTAT GATCCCATGT CTGTTGCGAT GAACCTCACT GAGTAAGTGC
251 CTCTTATGAT CCCATGTCTG TTGCGATGAA CCTCACTGAG TCAGTGCCTC
301 TTATGATCCC ATGTCTGTTG CGATGAACCT CACTGAGTCA GTGCCTCTTA
351 TGATCCCATG TCTGTTGCGA TGAACCTCAC TGAGTCAGTG CCTCTTATGA
401 TCCCATGTCT GTTGCGATGA ACCTCACTGA GTCAGTGCCT CTTATGATCC
451 TCGAGCCATG TYTGGKCGAG GCCTGTAGGC GTGTACGGTG GGAGGCTTAT
501 ATAAGCAGAG CTCAAGCTGG CATCCGGTAC TTKRAGCCMC CT
```

Appendix 6.12. Promoter 5 – 45 bp space between 8AP-1 motifs

```
  1 TGTTGWTGTT AACTTGTTTA TTGCAGCTTA TAATGGTTAC AAATAAAGCA
 51 ATAGCATCRC AAATTTCACA AATAAAGCAT TTTTTTCACT GCATTCTAGT
101 TGYGGTTTGT CCAAACTCAT CAATGTATCT TATCATGTCT GGCCAGCTAG
151 CGTGCCTCTT ATGATCTGGA TCCATGATCT CTGCGATGAA CCTCACTGAG
201 TCAGTGCCTC TTATGATCTG GATCCATGAT CTCTGCGATG AACCTCACTG
251 AGTCAGTGCC TCTTATGATC TGGATCCATG ATCTCTGCGA TGAACCTCAC
301 TGAGTCAGTG CCTCTTATGA TCTGGATCCA TGATCTCTGC GATGAACCTC
351 ACTGAGTCAG TGCCTCTTAT GATCTGGATC CATGATCTCT GCGATGAACC
401 TCACTGAGTC AGTGCCTCTT ATGATCTGGA TCCATGATCT CTGCGATGAA
451 CCTCACTGAG TCAGTGCCTC TTATGATCTG GATCCATGAT CTCTGCGATG
501 AACCTCACTG AGTCAGTGCC TCTTATGATC TGGATCCATG ATCTCTGCGA
551 TGAACCTCAC TGAGTCAGTG CCTCTTATGA TCTGGATCTC GAGCCATGGT
601 SGRGGCCTGT AGGCGTGTAC GGTGGGAGGC TTATATAAGC AGAGCTCAAG
651 CTTGGCATCC GGTACKTKAG MCACCGG
```

Appendix 6.13. Promoter 1 – 50 bp space between 4AP-1 motifs

```
  1 CCTCCCACAC CTCCCCCTGA ACCTGAAACA TAAAATGAAT GCAATTGTTG
```

370

```
 51 TTGTTAACTT GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC

101 ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT CTAGTTGTGG

151 TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGGCCA GCTAGCGTGC

201 CTCTTATGAT CTGGATCCAT GTCTGTATCT CTGCGATGAA CCTCACTGAG

251 TCAGTGCCTC TTATGATCTG GATCCATGTC TGTATCTCTG CGATGAACCT

301 CACTGAGTCA GTGCCTCTTA TGATCTGGAT CCATGTCTGT ATCTCTGCGA

351 TGAACCTCAC TGAGTCAGTG CCTCTTATGA TCTGGATCCA TGTCTGTATC

401 TCTGCGATGA ACCTCACTGA GTCAGTGCCT CTTATGATCT GGATCTCGAG

451 CCATGGTCGA GGCCTGTAGG CGTGTACGGT GGGAGGCTTA TATAAGCAGA

501 GCTCAAGCTG GCATCCGTAC KTGARCCMCC T
```

Appendix 6.14. Promoter 2 – 55 bp space between 7AP-1 motifs

```
  1 CCCCCTTKAW CYTGAAAACA TAAAATGAAK GCAATTGTTG TTGTTAAAYT

 51 TGTTTATTGC AGCTTATAAT GGTTACAAAT AAAGCAATAG CATCACAAAT

101 TTCACAAATA AAGCATTTTT TTCACTGCMT TCTAGTTGTG GKTWGTCCAA

151 AYTCATCAAT GTATCTTATC ATGTCTGGCC AGCTAGCGTG CCTCTTATGA

201 TCTGGATCCA TGATCTCTGC GATGAACCTC ACCATGTTGA GTCAACAAGG

251 TGCCTCTTAT GATCTGGATC CATGATCTCT GCGATGAACC TCACCATGTT

301 GAGTCAACAA GGTGCCTCTT ATGATCTGGA TCCATGATCT CTGCGATGAA

351 CCTCACCATG TTGAGTCAAC AAGGTGCCTC TTATGATCTG GATCCATGAT

401 CTCTGCGATG AACCTCACCA TGTTGAGTCA ACAAGGTGCC TCTTATGATC

451 TGGATCCATG ATCTCTGCGA TGAACCTCAC CATGTTGAGT CAACAAGGTG

501 CCTCTTATGA TCTGGATCCA TGATCTCTGC GATGAACCTC ACCATGTTGA

551 GTCAACAAGG TGCCTCTTAT GATCTGGATC CATGATCTCT GCGATGAACC

601 TCACCATGTT GAGTCAACAA GGTGCCTCTT ATGATCTGGA TCTCGAGTCG

651 AGGCCTGTAG GCGTGTACGG TGGGAGGCTT ATATAAGCAG AGCTCAAGCT
```

Appendix 7. Sequencing data of the pCpG-HRE constructs used to create the pCpG-spacer vectors for increased spacing between the proximal TFBS and TATA box

The pCpG-HRE-constructs were sequenced using the GL2 reverse primer and were eventually used to create the cloning vectors with various spacings between the proximal TFBS and the TATA box. The HRE motifs are indicated by a box and have a 60bp, 66bp, 70bp and 74bp space between the proximal HRE and the TATA box. The TATAA box is highlighted in bold. The NheI (5'-GCTAGC) and XhoI (5'-CTCGAG) sites are underlined and represent the start and end of the region containing 4HRE, respectively (in forward and reverse orientations).

Appendix 7.1. pCpG-4HRE-0bp-Sal (60 bp from TATA box)

```
  1 TATCTTATCA TGTCTGGCCA GCTAGCGTGC CTCTTATGAT CTGGATTGCA

 51 CCATCTCTGC GATGAACCTC ACACGTGGGT GCCTCTTATG ATCTGGATTG

101 CACCATSTCT GCGATGAACC TCACACGTGG GTGCCTCTTA TGATCTGGAT

151 CTCGAGWCGR GGCCTGTAGG CGTGTACGGT GGGAGGCTTA TATAAGCAGA

201 GCTCAAGCTK GCATTCCGGG TWSTKTGTTK KTTTCCTCT
```

Appendix 7.2. pCpG-4HRE-5bp-Sal (66 bp from TATA box)

```
  1 RTATCTTATC ATGTCTGGCC AGCTAGCGTG CCTCTTATGA TCTGGATTGC

 51 ACCATCTCTG CGATGAACCT CACACGTGGG TGCCTCTTAT GATCTGGATT

101 GCACCATCTC TGSGATGAAC CTCACACGTG GGTGCCTCTT ATGATCTGGA

151 TCTCGAGCCA TGRTSGRGGC CTGTAGGCGT GTACGGTGGG AGGCTTATAT

201 AAGCAGAGCT CAAGCTKGCA TCCGKTACTG TGGATWAKKC TTTGCT
```

Appendix 7.3. pCpG-4HRE-9bp-Sal (70 bp from TATA box)

```
  1 TATCTTATCA TGTCTGGCCA GCTAGCGTGC CTCTTATGAT CTGGATTGCA
```

372

```
 51 CCATCTCTGC GATGAACCTC ACACGTGGGT GCCTCTTATG ATCTGGATTG

101 CACCATCTCT GCGATGAACC TCACACGTGG GTGCCTCTTA TGATCTGGAT

151 CTCGAGCCAT GTCTGGTCGA GGCCTGTAGG CGTGTACGGT GGGAGGCTTA

201 TATAAGCAGA GCTCAAGCTT GGCATCCGTA CKTGTGAACT TTTATACGG
```

Appendix 7.4. pCpG-4HRE-14bp-Sal (74 bp from TATA box)

```
  1 CAATAGCATC ACAAATTTCA CAAATAAAGC ATTTTTTTCA MTGCATTCTA

 51 GTTGWGGTTT GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGCCAGCT

101 AGCGTGCCTC TTATGATCTG GATTGCACCA TCTCTGCGAT GAACCTCACA

151 CGTGGGTGCC TCTTATGATC TGGATTGCAC CATCTCTGCG ATGAACCTCA

201 CACGTGGGTG CCTCTTATGA TCTGGATCTC GAGCCATGTC TGTATCCGTC

251 GAGGCCTGTA GGCGTGTACG GTGGGAGGCT TATATAAGCA GAGCTCAAGC

301 TGGCATCCGG TACKKTGAGC CCCCWTTT
```

Appendix 8. Sequencing data of the pCpG-NFκB and AP-1-responsive promoters with various spacing between the proximal TFBS and TATA box

The 6NFκB and 8AP-1 motifs were cloned into the different spacer vectors (above; Appendix 7) to allow a 55 bp, 66 bp, 70 bp and 74 bp space between the proximal TFBS and the TATA box. The selected constructs were sequenced using the GL2 reverse primer. The TFBSs are indicated by a box and the TATAA box is highlighted in bold. The NheI site (5'-GCTAGC) and the XhoI overhang (5'-TCGAG) are underlined and represent the start and end of the region containing multimerised TFBS, respectively.

Appendix 8.1. pCpG-8AP-1 (55 bp from TATA box)

```
  1 TGCCTCTTAT GATCTGGATT GAGTCAATCT CTGCGATGAA CCTCACTGAG

 51 TCAGTGCCTC TTATGATCTG GATTGAGTCA ATCTMTGCGA TGAACCTCAC

101 TGAGTCAGTG CCTCTTATGA TCTGGATTGA GTCAATCTCT GCGATGAACC

151 TCACTGAGTC AGTGCCTCTT ATGATCTGGA TTGAGTCAAT YTGTGCGATG

201 AACCTCACTG AGTCAGTGCC TCTTATGATT TGGATCTCGAG GCCTGTAGGC

251 GTGTACGGTG GGAGGCTTAT ATAAGCAGAG CTCAAGCTGG CATCCGTACK
```

Appendix 8.2. pCpG-6NFκB-X-0bp-Sal (60 bp from TATA box)

```
  1 CAAACTCATC AATGTATCTT ATCATGTCTG GCCAGCTAGC GTGCCTCTTA

 51 TGATCTGGAT GGGACTTTCC ATCTCTGCGA TGAACCTCAC GGGACTTTCC

101 GTGCCTCTTA TGATCTGGAT GGGACTTTCC ATCTCTGCGA TGAACCTCAC

151 GGGACTTTCC GTGCCTCTTA TGATCTGGAT GGGACTTTCC ATCTCTGCGA

201 TGAACCTCAC GGGACTTTCC GTGCCTCTTA TGATCTGGAT CTCGTGTCGA

251 GGCCTGTAGG CGTGTACGGT GGGARGCTTA TATAAGCAGA GCTCAAGCTT

301 GGCATCCGGT ACKTGAGMCM CCTAG
```


Appendix 8.3. pCpG-6NFκB-X-5bp-Sal (66 bp from TATA box)

```
  1 AYTCACAATG TATCTTATCA TGTCTGGCCA GCTAGCGTGC CTCTTATGAT

 51 CTGGAKGGGA CTTTCCATCT CTGCGATGAA CCTCACGGGA CTTTCCGTGC

101 CTCTTATGAT CTGGATGYKC CTTAACGACT CTGCGATGAA CCTCACGGGA

151 CTTTCCGTGC CTCTTATGAT CTGGATGGGA CTTTCCATCT CTGCGAGATC

201 GCACTTTCCG TGCCTCTTAT GATCYSGATC TGGAAGACTC TCCCCTGTAG

251 RCGTTACGGY YKKTTTCTYA TATAAGCAGA GCTAMAAGCT TGCATWCCGG
```


Appendix 8.4. pCpG-8AP-1-X-9bp-Sal (70 bp from TATA box)

```
  1 GTATCTTATC ATGTCTGGCC AGCTAGCGTG CCTCTTATGA TCTGGATTGA

 51 GTCAATCTCT GCGATGAACC TCACTGAGTC AGTGCCTCTT ATGATCTGGA

101 TTGAGTCAAT CTCTGCGATG AACCTCACTG AGTCAGTGCC TCTTATGATC

151 TGGATTGAGT CAATCTCTGC GATGAACCTC ACTGAGTCAG TGCCTCTTAT

201 GATCTGGATT GAGTCAATCT CTGCGATGAA CCTCACTGAG TCAGTGCCTC

251 TTATGATCTG GATCTCGAGC CATGTCTGTA TCCTCAGGCC TGTAGGCGTG

301 TACGGGGGAG GCTTATATAAG CAGAGCTCAA GCTGGCATCC GTACKTAGCA

351 TGGGGGGGGA TAACC
```

Appendix 9. Sequencing data of selected pCpG-4bp-composite promoters with a 4 bp space between TFBSs and a 66 bp space between the proximal TFBS and TATA box

The pCpG-4bp-composite promoters were sequenced using the GL2 reverse primer. The TFBSs are indicated by a box and the arrows show the TFBS orientation. The NheI overhang (5'-CTAG) and XhoI site (5'-CTCGAG) is underlined and represents the start and end of the region containing multimerised TFBSs, respectively. The TATAA box is highlighted in bold. All sequences are complete and schematically represented for each construct.

Appendix 9.1. Promoter 4- pCpG-4bp-composite promoter

```
  1 GCGRAAAAAA AAAACAGATA AAAGTAARTC AAACATATAT CCTGCWACWC
 51 GCGCGKATGT CGGTCATCRT GTSYCGCCCM RKGTMCGYST SCWGTRCSTC
101 CCTAGTGACT CACTAGGGGA CTTTCCCTAG GCACGTCTAG GCACGTCTAG
151 TGAGTCACTA GGCACGTCTA GCTCGCGATC TTATGATCTG GATCCATGCT
201 CGAGGCCTGT AGGCGTGTAC GGTGGGAGGC TTATATAAGC AGAGCTCAAG
251 CTGGCATCCG TACKTGAGCC AT
```


promoter 4

Appendix 9.2. Promoter 6- pCpG-4bp-composite promoter

```
  1 GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT
 51 GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGC CAGCTAGACG
101 CGTGCTAGGG AAAGTCCCCT AGACGTGC GG AAAGTCCCCT AGGCACGTCT
151 AGTGAGTCAC TAGGGGACTT TCCCTAGTGA CTCACTAGTG ACTCACTAGC
201 TCGCGATCTT ATGATCTGGA TCCATGCTCG AGGCCTGTAG GCGTGTACGG
251 TGGGAGGCTT ATATAAGCAG AGCTCAAGCT GGCATCCGGT ACKKGRAGCC
```
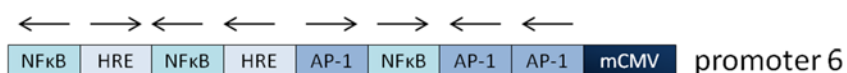

promoter 6

Appendix 9.3. Promoter 7- pCpG-4bp-composite promoter

```
  1 TTATAATGGT TACAAATAAA GCAATAGCAT CACAAATTTC ACAAATAAAG

 51 CATTTTTTTC ACTGCATTCT AGTTGTGGTT TGTCCAAACT CATCAATGTA

101 TCTTATCATG TCTGGCCAGC TAGACGCGTG CTAGGGAAAG TCCCCTAGAC

151 GTGCCTAGGC ACGTCTAGTG AGTCACTAGA CGTGCCTAGA CGTGCCTAGA

201 CGTGCCTAGC TCGCGATCTT ATGATCTGGA TCCATGCTCG AGGCCTGTAG

251 GCGTGTACGG TGGGAGGCTT ATATAAGCAG AGCTCAAGCT TGGCATCCGG
```


Appendix 9.4. Promoter 11- pCpG-4bp-composite promoter

```
  1 TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA

 51 AACTCATCAA TGTATCTTAT CATGTCTGGC CAGCTAGACG CGTGCTAGGG

101 GACTTTCCCT AGGCACGTCT AGTGACTCAC TAGGGGACTT TCCCTAGACG

151 TGCCTAGTGA GTCACTAGCT CGCGATCTTA TGATCTGGAT CCATGCTCGA

201 GGCCTGTAGG CGTGTACGGT GGGAGGCTTA TATAAGCAGA GCTCAAGCTT

251 GGCATCCGGT ACKTKKAGCC ACTA
```


Appendix 9.5. Promoter 15- pCpG-4bp-composite promoter

```
  1 AAACCTCCCA CACCTCCCCC TGAACCTGAA ACATAAAATG AATGCAATTG

 51 TTGTTGTTAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT

101 AGCATCACAA ATTTCACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG

151 TGGTTTGTCC AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGAC

201 GCGTGCTAGG CACGTCTAGT GAGTCACTAG GGGACTTTCC CTAGCTCGCG

251 ATCTTATGAT CTGGATCCAT GCTCGRGGCC TGTAGGCGTG TACGGTGGGA
```

376

301 GGCTTA**TATA A**GCAGAGCTC AAGCTGGCAT CCGGTACKTG AGCCACCC

promoter 15

Appendix 9.6. Promoter 17- pCpG-4bp-composite promoter

   1 AAATAAAGCA ATAGCATCAC AAATTTCACA AATAAAGCAT TTTTTTCACT

 51 GCATTCTAGT TGTGGTTTGT CCAAACTCAT CAATGTATCT TATCATGTCT

101 GGCCAG<u>CTAG</u> ACGCGTGCTA G$\overline{GGGACTTTC\ C}$CTAG$\overline{GCACG\ T}$CTAG$\overline{TGACT}$

151 $\overline{CA}$CTAG$\overline{GGGA\ CTTTCC}$CTAG $\overline{ACGTGC}$CTAG $\overline{TGAGTCA}$CTA GCTCGCGATC

201 TTATGATCTG GATCCATG<u>CT CGAG</u>GCCTGT AGGCGTGTAC GGTGGGAGGC

251 TTA**TATAA**GC AGAGCTCAAG CTTGGCATCC GGTACKTKWA SCCMCCCC

promoter 17

Appendix 10. Sequencing data of the selected pCpG-20bp-composite promoters with a 20bp space between TFBS and a 66bp between the proximal TFBS and TATA box

The selected pCpG-20bp-composite promoters were sequenced using the GL2 reverse primer. The TFBSs are positioned in the forward orientation and are indicated by a box. The NheI (5'-GCTAGC) and XhoI (5'-CTCGAG) sites are underlined and represent the start and end of the region containing multimerised TFBSs, respectively. The TATAA box is highlighted in bold. All sequences are complete and the schematic diagrams are presented for each construct.

Appendix 10.1. Promoter 2 (pCpG-20bp-composite promoter)

   1 CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT

 51 AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA

101 TGTATCTTAT CATGTCTGGC CA<u>GCTAGC</u>GT GCCTCTTATG ATCTGGAT$\overline{TG}$

151 $\overline{AGTCA}$ATCTC TGCGATGAAC CTCAC$\overline{TGAGT\ CA}$GTGCCTCT TATGATCTGG

201 AT$\overline{ACGTGC}$AT CTCTGCCGAT GAACCTCAC$\overline{T\ GAGTCA}$GTGC CTCTTATGAT

251 CTGGAT$\overline{ACGT\ GC}$ATCTCTGC GATGAACCTC AC$\overline{TGAGTCA}$G TGCCTCTTAT

```
301 GATCTGGATT GAGTCAATCT CTGCGATGAA CCTCACTGAG TCAGTGCCTC

351 TTATGATCTG GATTGAGTCA ATCTCTGCGA TGAACCTCAC ACGTGCGTGC

401 CTCTTATGAT CTGGATGGGA CTTTCCATCT CTGCGATGAA CCTCACGGGA

451 CTTTCCGTGC CTCTTATGAT CTGGATGGGA CTTTCCATCT CTGCGATGAA

501 CCTCACTGAG TCAGTGCCTC TTATGATCTG GATCTCGAGC CATGGTCGAG

551 GCCTGTAGGC GTGTACGGTG GGAGGCTTAT ATAAGCAGAG CTCAAGCTTG

601 GCATCCGGTA CKKGAGCCMM CC
```


AP-1 | AP-1 | HRE | AP-1 | HRE | AP-1 | AP-1 | AP-1 | AP-1 | HRE | NFκB | NFκB | NFκB | AP-1 | mCMV — promoter 2

Appendix 10.2. Promoter 9 (pCpG-20bp-composite promoter)

```
  1 TTCTAGTTGT GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGC

 51 CAGCTAGCGT GCCTCTTATG ATCTGGATTG AGTCAATCTC TGCGATGAAC

101 CTCACACGTG CGTGCCTCTT ATGATCTGGA TGGGACTTTC CACCTCTGCG

151 ATGAACCTCA CACGTGCGTG CCTCTTATGA TCTGGATGGG ACTTTCCATC

201 TCTGCGATGA ACCTCACTGA GTCAGTGCCT CTTATGATCT GGATTGAGTC

251 AATCTCTGCG ATGAACCTCA CGGGACTTTC CGTGCCTCTT ATGATCTGGA

301 TACGTGCATC TCTGCGATGA ACCTCACTGA GTCAGTGCCT CTTATGATCT

351 GGATGGGACT TTCCATCTCT GCGATGAACC TCACACGTGC GTGCCTCTTA

401 TGATCTGGAT TGAGTCAATC TCTGCGATGA ACCTCACACG TGCGTGCCTC

451 TTATGATCTG GATCTCGAGC CATGGTCGAG GCCTGTAGGC GTGTACGGTG

501 GGAGGCTTAT ATAAGCAGAG CTCAAGCTGG CATCCGGTAC KTGAGMCMCC
```


AP-1 | HRE | NFκB | HRE | NFκB | AP-1 | AP-1 | NFκB | HRE | AP-1 | NFκB | HRE | AP-1 | HRE | mCMV — promoter 9

Appendix 10.3. Promoter 4 (pCpG-20bp-composite promoter)

```
  1 CATTCTAGTT GTGGTTTGTC CAAACTCATC AATGTATCTT ATCATGTCTG

 51 GCCAGCTAGC GTGCCTCTTA TGATCTGGAT TGAGTCAATC TCTGCGATGA

101 ACCTCACTGA GTCAGTGCCT CTTATGATCT GGATACGTGC ATCTCTGCGA
```

```
151 TGAACCTCAC ACGTGCGTGC CTCTTATGAT CTGGATGGGA CTTTCCATCT

201 CTGCGATGAA CCTCACACGT GCGTGCCTCT TATGATCTGG ATGGGACTTT

251 CCATCTCTGC GATGAACCTC ACGGGACTTT CCGTGCCTCT TATGATCTGG

301 ATGGGACTTT CCATCTCTGC GATGAACCTC ACTGAGTCAG TGCCTCTTAT

351 GATCTGGATG GGACTTTCCA TCTCTGCGAT GAACCTCACG GGACTTTCCG

401 TGCCTCTTAT GATCTGGATA CGTGCATCTC TGCGATGAAC CTCACTGAGT

451 CAGTGCCTCT TATGATCTGG ATACGTGCAT CTCTGCGATG AACCTCACTG

501 AGTCAGTGCC TCTTATGATC TGGATCTCGA GCCATGGTCG AGGCCTGTAG

551 GCGTGTACGG TGGGAGGCTT ATATAAGCAG AGCTCAAGCT TGGCATCCGG
```


AP-1 | AP-1 | HRE | HRE | NFκB | HRE | NFκB | NFκB | NFκB | AP-1 | NFκB | NFκB | HRE | AP-1 | HRE | AP-1 | mCMV   promoter 4

Appendix 10.4. Promoter 6 (pCpG-20bp-composite promoter)

```
  1 AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGCG TGCCTMTTAT

 51 RATCTGGAKT GAGTCAATCT CTGCGATGAA CCTCACACGT GCGTGCCTCT

101 TATGATCTGG ATACGTGCAT CTTTGCGATG AACCTCACGG GACTTTCCGT

151 GCCTCTTATG ATCTGGATAC GTGCATCTCT GCGATGAACC TCACACGTGC

201 GTGCCTCTTA TGATCTGGAT GGGACTTTCC ATCTCTGCGA TGAACCTCAC

251 ACGTGCGTGC CTCTTATGAT CTGGATGGGA CTTTCCATCT CTGCGATGAA

301 CCTCACACGT GCGTGCCTCT TATGATCTGG ATTGAGTCAA TCTCTGCGAT

351 GARCSTCACA CGTGCGTGCC TCTTATGATC TGGATCTCGA GCCATGGTSG

401 AGGCCTGTAG GCGTGTACGG TGGGAGGCTT ATATAAGCAG AGCTCAAGCT

451 GGCATTCCGG TAYKKGKYRR CCKCCTTYTG
```


AP-1 | HRE | HRE | NFκB | HRE | HRE | NFκB | HRE | NFκB | HRE | AP-1 | HRE | mCMV   promoter 6

Appendix 10.5. Promoter 20 (pCpG-20bp-composite promoter)

```
  1 GTATCTTATC ATGTCTGGCC AGCTAGCGTG CCTCTTATGA TCTGGATTGA

 51 GTCAATCTCT GCGATGAACC TCACTGAGTC AGTGCCTCTT ATGATCTGGA
```

379

```
101 TGGGACTTTC CATCTCTGCG ATGAACCTCA CACGTGCGTG CCTCTTATGA

151 TCTGGATACG TGCATCTCTG CGATGAACCT CACTGAGTCA GTGCCTCTTA

201 TGATCTGGAT CTCGAGCCAW GRTSGAGGCC TGTAGGCGTG TACGGTGGGA

251 RGCTTATATA AGCAGAGCTC AAGCTTKGCW TTCCGKAAYC TRYSGTTKKA
```

| AP-1 | AP-1 | NFκB | HRE | HRE | AP-1 | mCMV | promoter 20 |

Appendix 10.6. Promoter 11 (pCpG-20bp-composite promoter)

```
  1 GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGC CAGCTAGCGT

 51 GCCTCTTATG ATCTGGATGG GACTTTCCAT CTCTGCGATG AACCTCACTG

101 AGTCAGTGCC TCTTATGATC TGGATTGAGT CAATCTCTGC GATGAACCTC

151 ACACGTGCGT GCCTCTTATG ATCTGGATAC GTGCATCTCT GCGATGAACC

201 TCACGGGACT TTCCGTGCCT CTTATGATCT GGATGGGACT TTCCATCTCT

251 GCGATGAACC TCACTGAGTC AGTGCCTCTT ATGATCTGGA TCTCGAGCCA

301 TGGTCGAGGC CTGTAGGCGT GTACGGTGGG AGGCTTATAT AAGCAGAGCT

351 CAAGCTGGCA TCCGGTACKK RAGRCCCC
```

| NFκB | AP-1 | AP-1 | HRE | HRE | NFκB | NFκB | AP-1 | mCMV | promoter 11 |

Appendix 10.7. Promoter 12 (pCpG-20bp-composite promoter)

```
 51 CCAGCTAGCG TGCCTCTTAT GATCTGGATT GAGTCAATCT CTGCGATGAA

101 CCTCAGGGGA CTTTCCGTGC CTCTTATGAT CTGGATGGGA CTTTCCATCT

151 CTGCGATGAA CCTCACTGAG TCAGTGCCTC TTATGATCTG GATTGAGTCA

201 ATCTCTGCGA TGAACCTCAC GGGACTTTCC GTGCCTCTTA TGATCTGGAT

251 ACGTGCATCT CTGCGATGAA CCTCACGGGA CTTTCCGTGC CTCTTATGAT

301 CTGGATCTCG AGCCATGGTC GAGGCCTGTA GGCGTGTACG GTGGGAGGCT

351 TATATAAGCA GAGCTCAAGC TGGCATCCGG TACKTGAGAC MCCCTC
```

| AP-1 | NFκB | NFκB | AP-1 | AP-1 | NFκB | HRE | NFκB | mCMV | promoter 12 |

Appendix 10.8. Promoter 14 (pCpG-20bp-composite promoter)

```
  1 AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGCG TGCCTCTTAT
 51 GATCTGGATG GGACTTTCCA TCTCTGCGAT GAACCTCACG GGACTTTCCG
101 TGCCTCTTAT GATCTGGATT GAGTCAATCT CTGCGATGAA CCTCACGGGA
151 CTTTCCGTGC CTCTTATGAT CTGGATACGT GCATCTCTGC GATGAACCTC
201 ACACGTGCGT GCCTCTTATG ATCTGGATGG GACTTTCCAT CTCTGCGATG
251 AACCTCACAC GTGCGTGCCT CTTATGATCT GGATCTCGAG CCATGGTCGA
301 GGCCTGTAGG CGTGTACGGT GGGAGGCTTA TATAAGCAGA GCTCAAGCTG
351 GCATCCGGTA CKTKRAARCC CMCCC
```


NFκB | NFκB | AP-1 | NFκB | HRE | HRE | NFκB | HRE | mCMV   promoter 14

Appendix 10.9. Promoter 3 (pCpG-20bp-composite promoter)

```
  1 CATGTCTGGC CAGCTAGCGT GCCTCTTATG ATCTGGATAC GTGCATCTCT
 51 GCGATGAACC TCACGGGACT TTCCGTGCCT CTTATGATCT GGATACGTGC
101 ATCTCTGCGA TGAACCTCAC ACGTGCGTGC CTCTTATGAT CTGGATTGAG
151 TCAATCTCTG CGATGAACCT CACGGGACTT TCCGTGCCTC TTATGATCTG
201 GATACGTGCA TCTCTGCGAT GAACCTCACT GAGTCAGTGC CTCTTATGAT
251 CTGGATACGT GCATCTCTGC GATGAACCTC ACACGTGCGT GCCTCTTATG
301 ATCTGGATCT CGAGCCATGG TCGAGGCCTG TAGGCGTGTA CGGTGGGAGG
351 CTTATATAAG CAGAGCTCAA GCTGGCATCC GGTACKTGAG CACTTT
```


HRE | NFκB | HRE | HRE | AP-1 | NFκB | HRE | AP-1 | HRE | HRE | mCMV   promoter 3

Appendix 10.10. Promoter 5 (pCpG-20bp-composite promoter)

```
  1 GCAATAGCGT CACAAATTTC ACAAATAAAG CATTTTTTTC ACTKCRTTST
 51 AGTTGWGGTT TGTCCAAACT CATCAATGTA TCTTATCATG TCTGGCCAGC
```

```
101 TAGCGTGCCT CTTATGATCT GGATACGTGC ATCTCTGCGA TGAACCTCAC

151 ACGTGCGTGC CTCTTATGAT CTGGATACGT GCATCTCTGC GATGAACCTC

201 ACTGAGTCAG TGCCTCTTAT GATCTGGATC TCGAGCCAWG GKSGAGGCCT

251 GTAGGCGTGT ACGGTGGGAG GCTTATATAA GCAGAGCTCA AGCTTGGCAT

301 CCGGTACTGT GARCCWYCTT
```


Appendix 10.11. Promoter 15 (pCpG-20bp-composite promoter)

```
  1 ATTTCACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG TGGTTTGTCC

 51 AAACTCATCA ATGTATCTTA TCATGTCTGG CCAGCTAGCG TGCCTCTTAT

101 GATCTGGATG GGACTTTCCA TCTCTGCGAT GAACCTCACG GGACTTTCCG

151 TGCCTCTTAT GATCTGGATC TCGAGCCATG GKCGAGGCCT GTAGGCGTGT

201 ACGGTGGGAG GCTTATATAA GCAGAGCTCA AGCTGGCATC CGGTACTKTG
```


Appendix 10.12. Promoter 18 (pCpG-20bp-composite promoter)

```
  1 TGTATCTTAT CATGTCTGGC CAGSTAGCGT GCCTCTTATG ATCTGGATGG

 51 GACTTTCCAT CTCTGCGATG AACCTCACAC GTGCGTGCCT CTTATGATCT

101 GGATGGGACT TTCCATCTCT GCGATGAACC TCACGGGACT TTCCGTGCCT

151 CTTATGATCT GGATTGAGTC AATCTCTGCG ATGAACCTCA CGGGACTTTC

201 CGTGCCTCTT ATGATCTGGA TGGGACTTTC CATCTCTGCG ATGAACCTCA

251 CACGTGCGTG CCTCTTATGA TCTGGATCTC GAGCCATGGT CGAGGCCTGT

301 AGGCGTGTAC GGTGGGAGGC TTATATAAGC AGAGCTCAAG CTGGCATCCG

351 GTACKTTAGC ACTGGCA
```


Appendix 10.13. Promoter 19 (pCpG-20bp-composite promoter)

```
  1 TGTCCAAACT CATCAATGTA TCTTATCATG TCTGGCCAGC TAGCGTGCCT

 51 CTTATGATCT GGATTGAGTC AATCTCTGCG ATGAACCTCA CGGGACTTTC

101 CGTGCCTCTT ATGATCTGGA TACGTGCATC TCTGCGATGA ACCTCACTGA

151 GTCAGTGCCT CTTATGATCT GGATGGGACT TTCCATCTCT GCGATGAACC

201 TCACTGAGTC AGTGCCTCTT ATGATCTGGA TGGGACTTTC CATCTCTGCG

251 ATGAACCTCA CACGTGCGTG CCTCTTATGA TCTGGATGGG ACTTTCCATC

301 TCTGCGATGA ACCTCACTGA GTCAGTGCCT CTTATGATCT GGATACGTGC

351 ATCTCTGCGA TGAACCTCAC ACGTGCGTGC CTCTTATGAT CTGGATCTCG

401 AGCCATGGTC GAGGCCTGTA GGCGTGTACG GTGGGAGGCT TATATAAGCA

451 GAGCTCAAGC TGGCATCCGG TACKTGAGCA CTTG
```

| AP-1 | NFκB | HRE | AP-1 | NFκB | AP-1 | NFκB | HRE | NFκB | AP-1 | HRE | HRE | mCMV | promoter 19

Appendix 11. Sequencing data of selected pCpG-clustered-composite promoters

The selected pCpG-cluster-composite promoters were sequenced using the Forward pCpG primer. The TFBSs are indicated by a box and the arrows show the orientation of the TFBS cluster. The NheI (5'-GCTAGC) and XhoI (5'-CTCGAG) sites are underlined and represent the start and end of the region containing clustered TFBSs, respectively. The TATAA box is highlighted in bold. Schematic diagrams are presented for each construct.

Appendix 11.1. Promoter 4 (pCpG-cluster-proximal 6NFκB)

```
241    TTTCACAAAT AAAGCATTTN TTTCACTGCA TTCTAGTTGT

281    GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGC

321    CAGCTAGCGT GCCTCTTATG ATCTGGATTG AGTCAATCTC

361    TGCGATGAAC CTCACTGAGT CAGTGCCTCT TATGATCTGG

401    ATTGAGTCAA TCTCTGCGAT GAACCTCACT GAGTCAGTGC

441    CTCTTATGAT CTGGATTGAG TCAATCTCTG CGATGAACCT

481    CACTGAGTCA GTGCCTCTTA TGATCTGGAT TGAGTCAATC
```

```
521   TCTGCGATGA ACCTCAC TGA GTCA GTGCCT CTTATGATCT

561   GGAT CTCGAG ATCCAGATCA TAAGAGGCAC  GCACGT GTGA

601   GGTTCATCGC AGAGAT GCAC GT ATCCAGAT CATAAGAGGC

641   AC GCACGT GT GAGGTTCATC GCAGAGAT GC ACGT ATCCAG

681   ATCATAAGAG GCAC GCACGT  GTGAGGTTCA TCGCAGAGAT

721    GCACGT ATCC AGATCATAAG AGGCAC GCTA GCGTGCCTCT

761   TATGATCTGG AT GGGACTTT CC ATCTCTGC GATGAACCTC

801   AC GGGACTTT CC GTGCCTCT TATGATCTGG AT GGGACTTT

841    CC ATCTCTGC GATGAACCTC AC GGGACTTT CC GTGCCTCT

881   TATGATCTGG AT GGGACTTT CC ATCTCTGC GATGAACCTC

921   AC GGGACTTT CC GTGCCTCT TATGATCTGG AT CTCGAG CC

961   ATGGTCGAGG CCTGTAGGCG TGTACGGTGG GAGGCTTA**TA**

1001  **TAA**AGCAGAG CTCAAGCTGC ATCCCNGTAC
```


8AP-1 | 6HRE | 6NFκB | mCMV   Promoter 4

Appendix 11.2. Promoter 5 (pCpG-cluster-proximal 6NFκB)

```
1     CTTATAATGG TTACAAATAA AGCAATAGCA TCACAAATTT

41    CACAAATAAA GCATTTTTTT CACTGCATTC TAGTTGTGGT

81    TTGTCCAAAC TCATCAATGT ATCTTATCAT GTCTGGCCA G

121    CTAGC GTGCC TCTTATGATC TGGAT ACGTG C ATCTCTGCG

161   ATGAACCTCA C ACGTGC GTG CCTCTTATGA TCTGGAT ACG

201    TGC ATCTCTG CGATGAACCT CAC ACGTGC G TGCCTCTTAT

241   GATCTGGAT A  CGTGC ATCTC TGCGATGAAC CTCAC ACGTG

281    C GTGCCTCTT ATGATCTGGA T CTCGAG ATC CAGATCATAA

321   GAGGCAC TGA CTCA GTGAGG TTCATCGCAG AGAT TGACTC
```

```
361    AATCCAGATC ATAAGAGGCA CTGACTCAGT GAGGTTCATC

401    GCAGAGATTG ACTCAATCCA GATCATAAGA GGCACTGACT

441    CAGTGAGGTT CATCGCAGAG ATTGACTCAA TCCAGATCAT

481    AAGAGGCACT GACTCAGTGA GGTTCATCGC AGAGATTGAC

521    TCAATCCAGA TCATAAGAGG CACGCTAGCG TGCCTCTTAT

561    GATCTGGATG GGACTTTCCA TCTCTGCGAT GAACCTCACG

601    GGACTTTCCG TGCCTCTTAT GATCTGGATG GGACTTTCCA

641    TCTCTGCGAT GAACCTCACG GGACTTTCCG TGCCTCTTAT

681    GATCTGGATG GGACTTTCCA TCTCTGCGAT GAACCTCACG

721    GGACTTTCCG TGCCTCTTAT GATCTGGATC TCGAGCCATG

761    GTCGAGGCCT GTAGGCGTGT ACGGTGGGAG GCTTA**TATAA**

801    AGCAGAGCTC AAGCTGCATC CGTACCT
```

←

| 6HRE | 8AP-1 | 6NFκB | mCMV | Promoter 5 |

Appendix 11.3. Promoter 11 (pCpG-cluster-proximal 8AP-1)

```
1      GCATCACAAA TNTCACAAAT AAAGCATTTN TTTCACTGCA

41     TTCTAGTTGT GGTTTGTCCA AACTCATCAA TGTATCTTAT

81     CATGTCTGGC CAGCTAGCGT GCCTCTTATG ATCTGGATGG

121    GACTTTCCAT CTCTGCGATG AACCTCACGG GACTTTCCGT

161    GCCTCTTATG ATCTGGATGG GACTTTCCAT CTCTGCGATG

201    AACCTCACGG GACTTTCCGT GCCTCTTATG ATCTGGATGG

241    GACTTTCCAT CTCTGCGATG AACCTCACGG GACTTTCCGT

281    GCCTCTTATG ATCTGGATCT CGAGATCCAG ATCATAAGAG

321    GCACGCACGT GTGAGGTTCA TCGCAGAGAT GCACGTATCC

361    AGATCATAAG AGGCACGCAC GTGTGAGGTT CATCGCAGAG

401    ATGCACGTAT CCAGATCATA AGAGGCACGC ACGTGTGAGG

441    TTCATCGCAG AGATGCACGT ATCCAGATCA TAAGAGGCAC

481    GCTAGCGTGC CTCTTATGAT CTGGATTGAG TCAATCTCTG

521    CGATGAACCT CACTGAGTCA GTGCCTCTTA TGATCTGGAT
```

```
561   TGAGTCAATC TCTGCGATGA ACCTCACTGA GTCAGTGCCT

601   CTTATGATCT GGATTGAGTC AATCTCTGCG ATGAACCTCA

641   CTGAGTCAGT GCCTCTTATG ATCTGGATTG AGTCAATCTC

681   TGCGATGAAC CTCACTGAGT CAGTGCCTCT TATGATCTGG

721   ATCTCGAGCC ATGGTCGAGG CCTGTAGGCG TGTACGGTGG

761   GAGGCTTATA TAAAGCAGAG CTCAAGCTGC ATNCCCCGTA
```

←

| 6NFκB | 6HRE | 8AP-1 | mCMV | Promoter 11

Appendix 11.4. Promoter 15 (pCpG-cluster-proximal 8AP-1)

```
1     CACAAATAAA GCATTTTTTT CACTGCATTC TAGTTGTGGT

41    TTGTCCAAAC TCATCAATGT ATCTTATCAT GTCTGGCCAG

81    CTAGCGTGCC TCTTATGATC TGGATACGTG CATCTCTGCG

121   ATGAACCTCA CACGTNNGTG CCTCTTATGA TCTGGATACG

161   TGCATCTCTG CGATGAACCT CACACGTGCG TGCCTCTTAT

201   GATCTGGATA CNNGCATCTC TGCGATGAAC CTCACACGTG

241   CGTGCCTCTT ATGATCTGGA TCTCGAGATC CAGATCATAA

281   GAGGCACGGA AAGTCCCGTG AGGTTCATCG CAGAGATGGA

321   AAGTCCCATC CAGATCATAA GAGGCACGGA AAGTCCCGTG

361   AGGTTCATCG CAGAGATGGA AAGTCCCATC CAGATCATAA

401   GAGGCACGGA AAGTCCCGTG AGGTTCATCG CAGAGATGGA

441   AAGTCCCATC CAGATCATAA GAGGCACGCT AGCGTGCCTC

481   TTATGATCTG GATTGAGTCA ATCTCTGCGA TGAACCTCAC

521   TGAGTCAGTG CCTCTTATGA TCTGGATTGA GTCAATCTCT

561   GCGATGAACC TCACTGAGTC AGTGCCTCTT ATGATCTGGA

601   TTGAGTCAAT CTCTGCGATG AACCTCACTG AGTCAGTGCC

641   TCTTATGATC TGGATTGAGT CAATCTCTGC GATGAACCTC

681   ACTGAGTCAG TGCCTCTTAT GATCTGGATC TCGAGCCATG
```

```
721    GTCGAGGCCT GTAGGCGTGT ACGGTGGGAG GCTTATATAA

761    AGCAGAGCTC AAAGCTGCAT CCGTACC
```

←

| 6HRE | 6NFκB | 8AP-1 | mCMV |

Promoter 15

Appendix 11.5. Promoter 5 (pCpG-cluster-proximal 6HRE)

```
1      GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGCCAGCT

41     AGCGTGCCTC TTATGATCTG GATGGGACTT TCCATCTCTG

81     CGATGAACCT CACGGGACTT TCCGTGCCTC TTATGATCTG

121    GATGGGACTT TCCATCTCTG CGATGAACCT CACGGGACTT

161    TCCGTGCCTC TTATGATCTG GATGGGACTT TCCATCTCTG

201    CGATGAACCT CACGGGACTT TCCGTGCCTC TTATGATCTG

241    GATCTCGAGA TCCAGATCAT AAGAGGCACT GACTCAGTGA

281    GGTTCATCGC AGAGATTGAC TCAATCCAGA TCATAAGAGG

321    CACTGACTCA GTGAGGTTCA TCGCAGAGAT TGACTCAATC

361    CAGATCATAA GAGGCACTGA CTCAGTGAGG TTCATCGCAG

401    AGATTGACTC AATCCAGATC ATAAGAGGCA CTGACTCAGT

441    GAGGTTCATC GCAGAGATTG ACTCAATCCA GATCATAAGA

481    GGCACGCTAG CGTGCCTCTT ATGATCTGGA TACGTGCATC

521    TCTGCGATGA ACCTCACACG TGCGTGCCTC TTATGATCTG

561    GATACGTGCA TCTCTGCGAT GAACCTCACA CGTGCGTGCC

601    TCTTATGATC TGGATACGTG CATCTCTGCG ATGAACCTCA

641    CACGTGCGTG CCTCTTATGA TCTGGATCTC GAGCCATGGT

681    CGAGGCCTGT AGGCGTGTAC GGTGGGAGGC TTATATAAAG

721    CAGAGCTCAA GCTGCATCCG TACC
```

←

| 6NFκB | 8AP-1 | 6HRE | mCMV |

Promoter 5

Appendix 11.6. Promoter 4 (pCpG-cluster-proximal 6HRE)

```
1     CATTTTNTTC ACTGCATTCT AGTTGTGGTT TGTCCAAACT

41    CATCAATGTA TCTTATCATG TCTGGCCAGC TAGCGTGCCT

81    CTTATGATCT GGATTGAGTC AATCTCTGCG ATGAACCTCA

121   CTGAGTCAGT GCCTCTTATG ATCTGGATTG AGTCAATCTC

161   TGCGATGAAC CTCACTGAGT CAGTGCCTCT TATGATCTGG

201   ATTGAGTCAA TCTCTGCGAT GAACCTCACT GAGTCAGTGC

241   CTCTTATGAT CTGGATTGAG TCAATCTCTG CGATGAACCT

281   CACTGAGTCA GTGCCTCTTA TGATCTGGAT CTCGAGATCC

321   AGATCATAAG AGGCACGGAA AGTCCCGTGA GGTTCATCGC

361   AGAGATGGAA AGTCCCATCC AGATCATAAG AGGCACGGAA

401   AGTCCCGTGA GGTTCATCGC AGAGATGGAA AGTCCCATCC

441   AGATCATAAG AGGCACGGAA AGTCCCGTGA GGTTCATCGC

481   AGAGATGGAA AGTCCCATCC AGATCATAAG AGGCACGCTA

521   GCGTGCCTCT TATGATCTGG ATACGTGCAT CTCTGCGATG

561   AACCTCACAC GTGCGTGCCT CTTATGATCT GGATACGTGC

601   ATCTCTGCGA TGAACCTCAC ACGTGCGTGC CTCTTATGAT

641   CTGGATACGT GCATCTCTGC GATGAACCTC ACACGTGCGT

681   GCCTCTTATG ATCTGGATCT CGAGCCATGG TCGAGGCCTG

721   TAGGCGTGTA CGGTGGGAGG CTTA**TATAA**G CAGAGCTCAA

761   GCTGGCATTC CGTA
```

←

8AP-1 | 6NFκB | 6HRE | mCMV   Promoter 4

Appendix 12. Sequencing data of plasmid-miR-23b and lentiviral-miR-23b-target constructs

The plasmid-miR-23b and lentiviral-miR-23b-target constructs were sequenced using the End of Luc⁺ forward sequencing primer. For the pcLuc⁺-miR-23b-target constructs, the XbaI

overhang (5'-TCTAG) and the ApaI site (5'-GGGCCC) is underlined and represents the start and end of the miR-23b target sequences, respectively. The internal XbaI restriction site (5'-TCTAGA) is also underlined and the miR-23b target sequences (5'-GGTAATCCCTGGCAATGTGAT-3') are indicated by boxes.

Appendix 12.1. pcLuc⁺-miR-23b-2T

```
 801 AAAGAGATCG TGGATTACGT CGCCAGTCAA GTAACAACCG CGAAAAAGTT
 851 GCGCGGAGGA GTTGTGTTTG TGGACGAAGT ACCGAAAGGT CTTACCGGAA
 901 AACTCGACGC AAGAAAAATC AGAGAGATCC TCATAAAGGC CAAGAAGGGC
                      XbaI              miR-23b target
 951 GGAAAGATCG CCGTGTAATT CTAGGGTAAT CCCTGGCAAT GTGATCGATG
        miR-23b target          XbaI          ApaI
1001 GTAATCCCTG GCAATGTGAT TCTAGAATTC GGGCCCTATT CTATAGTGTC
1051 ACCTAAATGC TAGAGCTCGC TGATCAGCCT CGACTGTGCC TTCTAGTTGC
1101 CAGCCATCTG TTGTTTGCCC CTCCCCCGTG CCTTCCTTGA CCCTGGAAGG
1151 TGCCACTCCC ACTGTCCTTC CTAATAAACK AGTTTSCT
```

Appendix 12.2. pcLuc⁺-miR-23b-4T

```
 751 AACTTCCCGC CGCCGTTGTT GTTTTGGAGC ACGGAAAGAC GATGACGGAA
 801 AAAGAGATCG TGGATTACGT CGCCAGTCAA GTAACAACCG CGAAAAAGTT
 851 GCGCGGAGGA GTTGTGTTTG TGGACGAAGT ACCGAAAGGT CTTACCGGAA
 901 AACTCGACGC AAGAAAAATC AGAGAGATCC TCATAAAGGC CAAGAAGGGC
                      XbaI              miR-23b target
 951 GGAAAGATCG CCGTGTAATT CTAGGGTAAT CCCTGGCAAT GTGATCGATG
        miR-23b target                 miR-23b target
1001 GTAATCCCTG GCAATGTGAT TCTAGGGTAA TCCCTGGCAA TGTGATCGAT
        miR-23b target         XbaI          ApaI
1051 GGTAATCCCT GGCAATGTGA TTCTAGAATT CGGGCCCTAT TCTATAGTGT
1101 CACCTAAATG CTAGAGCTCG CTGATCAGCC TCGACTGTGC CTTCTAGTTG
```

```
1151 CCAGCCATCT GTTGTTTGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG
1201 GTGCCACTCC CACTGTCCTT CCTAATAAAC KAGTTTSCT
```

For the lentiviral-miR-23b target constructs, the MfeI overhang (5'-AATTG) and the KpnI site (5'-GGTACC) is underlined and represents the start and end of the miR-23b target sequences, respectively. The miR-23b target sequences (5'-GGTAATCCCTGGCAATGTGAT-3') are indicated by boxes.

Appendix 12.3. LV-SFFV-miR-23b-2T

```
451 GCTGACGTCC TTTCCATGGC TGCTCGCCTG TGTTGCCACC TGGATTCTGC
501 GCGGGACGTC CTTCTGCTAC GTCCCTTCGG CCCTCAATCC AGCGGACCTT
551 CCTTCCCGCG GCCTGCTGCC GGCTCTGCGG CCTCTTCCGC GTCTTCGCCT
                                                          MfeI
601 TCGCCCTCAG ACGAGTCGGA TCTCCCTTTG GGCCGCCTCC CCGCCTGGAA
                        miR-23b target              miR-23b target
651 TTGGTAATTC TAGGGTAATC CTGGCAATGT GATCGATGGT AATCCCTGGC
                                    KpnI
701 AATGTGATTC TAGAATTCGG GCCCTATTCG GTACCTTTAA GACCAATGAC
751 TTACAAGGCA GCTGTAGATC TTAGCCACTT TTTAAAAGAA AAGGGGGGAC
801 TGGAAGGGCT AATTCACTCC CAACGAAGAC AAGATGTGAT CCTCCTAGAG
851 GGACAGCCCC CCCCCAAAGC CCCCAGGGAT GTAATTACGT CCCTCCYCCG
```

Appendix 12.4. LV-SFFV-miR-23b-4T

```
501 CGCGGGACGT CCTTCTGCTA CGTCCCTTCG GCCCTCAATC CAGCGGACCT
551 TCCTTCCCGC GGCCTGCTGC CGGCTCTGCG GCCTCTTCCG CGTCTTCGCC
601 TTCGCCCTCA GACGAGTCGG ATCTCCCTTT GGGCCGCCTC CCCGCCTGGA
     MfeI                  miR-23b target            miR-23b target
651 ATTGGTAATT CTAGGGTAAT CCTGGCAATG TGATCGATGG TAATCCCTGG
                        miR-23b target              miR-23b target
701 CAATGTGATT CTAGGGTAAT CCCTGGCAAT GTGATCGATG GTAATCCCTG
                                     KpnI
751 GCAATGTGAT TCTAGAATTC GGGCCCTATT CGGTACCTTT AAGACCAATG
801 ACTTACAAGG CAGCTGTAGA TCTTAGCCAC TTTTTAAAAG AAAAGGGGGG
851 ACTGGAAGGG CTAATTCACT CCCACGAAGA CAAGATGTGA TCCTCCTAGA
901 GGGACAGCCC CCCCCCAAAG CCCCCAGGGA TGTAATTACC TCCCTCCCCC
```

For the 6NFκB-miR-23b-target construct, the XbaI overhang (5'-TCTAG) and the FseI site (5'-GGCCGGCC) is underlined and represents the start and end of the miR-23b target sequence, respectively. The internal XbaI restriction site (5'-TCTAGA) is also underlined and the miR-23b target sequences (5'-GGTAATCCCTGGCAATGTGAT-3') are indicated by boxes.

Appendix 12.5. 6NFkB-miR-23b-2T

```
                                                            XbaI
  1 CRRAAAWYYT CMWAAAAGCC AAGAAGGGCG GAAAGATCGC CGTGTAATTC
              miR-23b target                  miR-23b target
 51 TAGGGTAATC CTGGCAATGT GATCGATGGT AATCCCTGGC AATGTGATTC
    XbaI           FseI
101 TAGAATTCGG GCCGGCCGCT TCGAGCAGAC ATGATAAGAT ACATTGATGA
151 GTTTGGACAA ACCACAACTA GAATGCAGTG AAAAAAATGC TTTATTTGTG
201 AAATTTGTGA TGCTATTGCT TTATTTGTAA CCATTATAAG CTGCAATAAA
```

Appendix 13. Sequencing data of lentiviral synthetic promoters expressing the luciferase gene

Appendix 13.1. LV-CMV-Luc⁺

The cloning vector LV-CMV-Luc⁺ was sequenced using the GL2 reverse primer (Appendix 13.1.1) and the End of Luc⁺ forward primer (Appendix 13.1.2). A BLAST search was performed to compare the sequence of the LV-CMV-Luc⁺ to that of the backbone lentiviral plasmid pLV.CMVenh.gp91.eGFP.cHS4 (Addgene plasmid 30471), using NCBI BLAST http://blast.ncbi.nlm.nih.gov/Blast.cgi.

Appendix 13.1.1. LV-CMV-Luc⁺ (sequenced using the GL2 reverse primer)

```
  1 ATATCAACCA TTMCAATTGT TAATACTTTT MTCAGCAAGG CTATGAATGM
 51 TGTTCCAGCS TGTCAAAATC ACACCTGTTT AATGTGTTTT ACCCAGCACG
101 AAGTCATGTS TAGTTGAGTG GSTTAAAAAT TGTGATCAAA TAGSTGGTTA
151 GTTAAAAAGT TATTTCACTG TGTAAAATAC ATCCCTTAAA ATGCACTGTT
201 ATTWATCTCT TAGTTGTAGA AATTGGTTTC ATTTTCCACT ATGTTTAATT
251 GTGACTGGAT CATTATAGAC CCTTTTTTTG TAGTTGTTGA GGTTTAAAGA
301 TTTAAGTTTG TTATGGATGC AAGCTTTTCA GTTGACCAAT GATTATTAGC
351 CAATTTSTGA TAAAAGAAAA GGAAACCGAT TGCCCCAGGG CTGCTRWTKK
401 CATTTCCTCA TTGGAAGAAG AAGCATAGTA TAGAAGAAAA GGCAAACACA
451 ACACATCAAC YTYYKGCCMC CCCSWMMRAW YYYCA
```

Score	Expect	Identities	Gaps	Strand
785 bits(425)	0.0	445/461(97%)	2/461(0%)	Plus/Plus

```
Query 3457  ATATCAACCATTACAATTGTTACTACTTTTCTCAGCAAGGCTATGAATGCTGTTCCAGCC  3516
            |||||||||| ||||||||| |||||| ||||||||||||||||||| |||||||||||
Sbjct 1     ATATCAACCATTMCAATTGTTAATACTTTTMTCAGCAAGGCTATGAATGMTGTTCCAGCS  60

Query 3517  TGTCAAAATCACACCTGTTTAATGTGTTTTACCCAGCACGAAGTCATGTCTAGTTGAGTG  3576
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct 61    TGTCAAAATCACACCTGTTTAATGTGTTTTACCCAGCACGAAGTCATGTSTAGTTGAGTG  120

Query 3577  GCTTAAAAATTGTGATCAAATAGCTGGTTAGTTAAAAAGTTATTTCACTGTGTAAAATAC  3636
            | ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct 121   GSTTAAAAATTGTGATCAAATAGSTGGTTAGTTAAAAAGTTATTTCACTGTGTAAAATAC  180

Query 3637  ATCCCTTAAAATGCACTGTTATTTATCTCTTAGTTGTAGAAATTGGTTTCATTTTCCACT  3696
            ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Sbjct 181   ATCCCTTAAAATGCACTGTTATTWATCTCTTAGTTGTAGAAATTGGTTTCATTTTCCACT  240

Query 3697  ATGTTTAATTGTGACTGGATCATTATAGACCCtttttttGTAGTTGTTGAGGTTTAAAGA  3756
            |||||||||||||||||||||||||||||||||        ||||||||||||||||||||
Sbjct 241   ATGTTTAATTGTGACTGGATCATTATAGACCCTTTTTTTGTAGTTGTTGAGGTTTAAAGA  300

Query 3757  TTTAAGTTTGTTATGGATGCAAGCTTTTCAGTTGACCAATGATTATTAGCCAATTTCTGA  3816
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Sbjct 301   TTTAAGTTTGTTATGGATGCAAGCTTTTCAGTTGACCAATGATTATTAGCCAATTTSTGA  360

Query 3817  TAAAAGAAAAGGAAACCGATTGCCCCAGGGCTGCTGTTTTCATTTCCTCATTGGAAGAAG  3876
            |||||||||||||||||||||||||||||||||||   ||| |||||||||||||||||||
Sbjct 361   TAAAAGAAAAGGAAACCGATTGCCCCAGGGCTGCTRWTKKCATTTCCTCATTGGAAGAAG  420

Query 3877  AAGCATAGTATAGAAGAAA-GGCAAACACAACACATTCAAC  3916
            ||||||||||||||||||| |||||||||||||||| ||||
Sbjct 421   AAGCATAGTATAGAAGAAAAGGCAAACACAACACAT-CAAC  460
```

Appendix 13.1.1. BLAST output displaying 97% sequence homology between the LV-CMV-Luc⁺ and the backbone pLV.CMVenh.gp91.eGFP.cHS4 plasmid. The LV-CMV-Luc⁺ sequence (subject) aligns at position 3457bp of the pLV.CMVenh.gp91.eGFP.cHS4 (query), which corresponds to the region upstream of the eGFP reporter gene in the latter plasmid.

Appendix 13.1.2. LV-CMV-Luc⁺ (sequenced using the End of Luc⁺ forward primer)

```
  1 TMSKKKKKTC YYMWWWAARG CCMAGAAAGG GCGGAAGATC GCCGTGTAAT
 51 TCTAGAGTCG ACAATCAACC TCTGGATTAC AAAATTTGTG AAAGATTGAC
101 TGGTATTCTT AACTATGTTG CTCCTTTTAC GCTATGTGGA TACGCTGCTT
151 TAATGCCTTT GTATCATGCT ATTGCTTCCC GTATGGCTTT CATTTTCTCC
201 TCCTTGTATA AATCCTGGTT GCTGTCTCTT TATGAGGAGT TGTGGCCCGT
251 TGTCAGGCAA CGTGGCGTGG TGTGCACTGT GTTTGCTGAC GCAACCCCCA
301 CTGGTTGGGG CATTGCCACC ACCTGTCAGC TCCTTTCCGG GACTTTCGCT
351 TTCCCCCTCC CTATTGCCAC GGCGGAACTC ATCGCCGCCT GCCTTGCCCG
401 CTGCTGGACA GGGGCTCGGC TGTTGGGCAC TGACAAT TCC GTGGTGTTGT
451 CGGGGAAGCT GACGTCCTTT CCATGGCTGC TCGCCTGTGT TGCCACCTGG
501 ATTCTGCGCG GGACGTCCTT CTGCTACGTC CCTTCGGCCC TCAATCCAGC
551 GGACCTTCCT TCCCGCGGCC TGCTGCCGGC TCTGCGGCCT CTTCCGCGTC
601 TTCGCCTTCG CCCTCAGACG AGTCGGATCY CCCTTTGGGC CGCCTCCCCG
```

Score	Expect	Identities	Gaps	Strand
1107 bits(599)	0.0	600/601(99%)	0/601(0%)	Plus/Plus

```
Query  4677  GTCGACAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTAT  4736
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    57  GTCGACAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTAT   116

Query  4737  GTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCT  4796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   117  GTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCT   176

Query  4797  TCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAG  4856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   177  TCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAG   236

Query  4857  GAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACC  4916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   237  GAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACC   296

Query  4917  CCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCC  4976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   297  CCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCC   356

Query  4977  CTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCT  5036
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   357  CTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCT   416

Query  5037  CGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGG  5096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   417  CGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGG   476

Query  5097  CTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCG  5156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   477  CTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCG   536

Query  5157  GCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG  5216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   537  GCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG   596

Query  5217  CGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGGA  5276
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct   597  CGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCYCCCTTTGGGCCGCCTCCCCGCCTGGA   656
```

Appendix 13.1.2. BLAST output displaying 99% sequence homology between the LV-CMV-Luc⁺ and the backbone pLV.CMVenh.gp91.eGFP.cHS4 plasmid. The LV-CMV-Luc⁺ sequence (subject) aligns at position 4677 bp of the pLV.CMVenh.gp91.eGFP.cHS4 (query), which corresponds to the region immediately downstream of the eGFP gene in the latter plasmid.

Lentiviral-synthetic promoter- luciferase constructs

Unless stated otherwise, all lentiviral-synthetic promoter-luciferase constructs were sequenced using the GL2 reverse primer. TFBS are positioned in the forward orientation and are indicated by a box. The NheI (5'-GCTAGC) and XhoI (5'-CTCGAG) sites are underlined and represent the start and end of the region containing multimerised TFBSs, respectively. The TATAA box is highlighted in bold. Schematic diagrams of complete DNA sequences are presented for each construct

Appendix 13.2. LV-2-Luc⁺

```
201 TGCAGGGGAA AGAATAGTAG ACATAATAGC AACAGACATA CAAACTAAAG
251 AATTACAAAA ACAAATTACA AAAATTCAAA ATTTTATCGA TCACGAGACT
301 AGCTCGAGAA GCTTGATGAT CCGTTTCGAT GGCCAGCTAG CGTGCCTCTT
351 ATGATCTGGA TTGAGTCAAT CTCTGCGATG AACCTCACTG AGTCAGTGCC
401 TCTTATGATC TGGATACGTG CATCTCTGCC GATGAACCTC ACTGAGTCAG
451 TGCCTCTTAT GATCTGGATA CGTGCATCTC TGCGATGAAC CTCACTGAGT
501 CAGTGCCTCT TATGATCTGG ATTGAGTCAA TCTCTGCGAT GAACCTCACT
551 GAGTCAGTGC CTCTTATGAT CTGGATTGAG TCAATCTCTG CGATGAACCT
601 CACACGGCGT GCCTCTTATG ATCTGGATGG GACTTTCCAT CTCTGCGATG
651 AACCTCACGG GACTTTCCGT GCCTCTTATG ATCTGGATGG GACTTTCCAT
701 CTCTGCGATG AACCTCACTG AGTCAGTGCC TCTTATGATC TGGATCTCGA
751 SCCATRGTSG RGGCCTGTAG GCGTGTACGG TGGGAGGCTT ATATAAGCAG
801 AGCTCAAGCT GGCATCCGTA CKTGAGAMAA YTT
```

| AP-1 | AP-1 | HRE | AP-1 | HRE | AP-1 | AP-1 | AP-1 | AP-1 | HRE | NFκB | NFκB | NFκB | AP-1 | mCMV | Luc⁺ | LV-2-Luc⁺ |

Appendix 13.3. LV-9-Luc⁺

```
351 AAGAATAGTA GACATAATAG CAACAGACAT ACAAACTAAA GAATTACAAA
401 AACAAATTAC AAAAATTCAA AATTTTATCG ATCACGAGAC TAGCTCGAGA
451 AGCTTGATGA TCCGTTTCGA TGGCCAGCTA GCGTGCCTCT TATGATCTGG
501 ATTGAGTCAA TCTCTGCGAT GAACCTCACA CGTGCGTGCC TCTTATGATC
551 TGGATGGGAC TTTCCACCTC TGCGATGAAC CTCACACGTG CGTGCCTCTT
601 ATGATCTGGA TGGGACTTTC CATCTCTGCG ATGAACCTCA CTGAGTCAGT
651 GCCTCTTATG ATCTGGATTG AGTCAATCTC TGCGATGAAC CTCACGGGAC
701 TTTCCGTGCC TCTTATGATC TGGATACGTG CATCTCTGCG ATGAACCTCA
751 CTGAGTCAGT GCCTCTTATG ATCTGGATGG GACTTTCCAT CTCTGCGATG
```

```
801 AACCTCACAC GTGCGTGCCT CTTATGATCT GGATTGAGTC AATCTCTGCG
851 ATGAACCTCA CACGTGCGTG CCTCTTATGA TCTGGATCTC GAGCCATGGT
901 SGAGGCCTGT AGGCGTGTAC GGTGGGAGGC TTATATAAGC AGAGCTCAAG
```


LV-9-Luc⁺

Appendix 13.4. LV-11-Luc⁺

```
 751 GGGGATTGGG GGGTACAGTG CAGGGGAAAG AATAGTAGAC ATAATAGCAA
 801 CAGACATACA AACTAAAGAA TTACAAAAAC AAATTACAAA AATTCAAAAT
 851 TTTATCGATC ACGAGACTAG CTCGAGAAGC TTGATGATCC GTTTCGATGG
 901 CCAGCTAGCG TGCCTCTTAT GATCTGGATG GGACTTTCCA TCTCTGCGAT
 951 GAACCTCACT GAGTCAGTGC CTCTTATGAT CTGGATTGAG TCAATCTCTG
1001 CGATGAACCT CACACGTGCG TGCCTCTTAT GATCTGGATA CGTGCATCTC
1051 TGCGATGAAC CTCACGGGAC TTTCCGTGCC TCTTATGATC TGGATGGGAC
1101 TTTCCATCTC TGCGATGAAC CTCACTGAGT CAGTGCCTCT TATGATCTGG
1151 ATCTCGAGCC ATGGTCGAGG CCTGTAGGCG TGTACGGTGG GAGGCTTATA
1201 TAAGCAGAGC TCAAGCTGGC ATCCGGTACK TKRAAGSCCM CCCC
```


LV-11-Luc⁺

Appendix 13.5. LV-12-Luc⁺

```
 751 GGGAAAGAAT AGTAGACATA ATAGCAACAG ACATACAAAC TAAAGAATTA
 801 CAAAAACAAA TTACAAAAAT TCAAAATTTT ATCGATCACG AGACTAGCTC
 851 GAGAAGCTTG ATGATCCGTT TCGATGGCCA GCTAGCGTGC CTCTTATGAT
 901 CTGGATTGAG TCAATCTCTG CGATGAACCT CAGGGGACTT TCCGTGCCTC
 951 TTATGATCTG GATGGGACTT TCCATCTCTG CGATGAACCT CACTGAGTCA
1001 GTGCCTCTTA TGATCTGGAT TGAGTCAATC TCTGCGATGA ACCTCACGGG
1051 ACTTTCCGTG CCTCTTATGA TCTGGATACG TGCATCTCTG CGATGAACCT
1101 CACGGGACTT TCCGTGCCTC TTATGATCTG GATCTCGAGC CATGGTCGAG
1151 GCCTGTAGGC GTGTACGGTG GGAGGCTTAT ATAAGCAGAG CTCAAGCTTG
1201 GCATCCGTAC KTRAAGRMCM MCC
```


LV-12-Luc⁺

Appendix 13.6. LV-14-Luc⁺

```
151 TACAAAAACA AATTACAAAA ATTCAAAATT TTATCGATCA CGAGACTAGC
201 TCGAGAAGCT TGATGATCCG TTTCGATGGC CAGCTAGCGT GCCTCTTATG
251 ATCTGGATGG GACTTTCCAT CTCTGCGATG AACCTCACGG GACTTTCCGT
```

```
301 GCCTCTTATG ATCTGGAT TG AGTCA ATCTC TGCGATGAAC CTCAC GGGAC

351 TTTCC GTGCC TCTTATGATC TGGAT ACGTG C ATCTCTGCG ATGAACCTCA

401 C ACGTGC GTG CCTCTTATGA TCTGGAT GGG ACTTTCC ATC TCTGCGATGA

451 ACCTCAC ACG TGC GTGCCTC TTATGATCTG GAT CTCGAGC CATGGTCGAG

501 GCCTGTAGGC GTGTACGGTG GGAGGCTTA T ATAA GCAGAG CTCAAGCTGG

551 CATCCGGTAC KTTKAAGCCC C
```

NFκB | NFκB | AP-1 | NFκB | HRE | HRE | NFκB | HRE | mCMV | Luc⁺ | LV-14-Luc⁺

Appendix 13.7. LV-6-Luc⁺ (Incomplete)

```
  1 TTCTATWMTT TTGGCTCTTC CATGGTGGGT TTAMCAACAG TAMCCGGAAK
 51 GCCAAGCTKG AGCMCCGCYK AWATTAGCCT CWACCTTTAC GCSTMCKGCA
101 GTGACCATWG TGGWGATTTA GATCATGGAR GAAAGAACGT GTGAGGTYCG
151 TCGCRGWGAT GRACGTGRCC MGTTATAWGA GGTTCGTACC TGAGMWGRAA
201 TCGCCCATAC AGATCATACC AGGCAC GGAA AGTCCC GTGA GGTTCATCGC
251 AGAGAK GGAA AGTCCC ATCC AGATCATAAG AGGCAC GCTA GCGTGCCTCT
301 TATGATCTGG AT ACGTGC AT CTCTGCGATG AACCTCAC AC GTGC GTGCCT
351 CTTATGATCT GGAT ACGTGC ATCTCTGCGA TGAACCTCAC ACGTGC GTGC
401 CTCTTATGAT CTGGAT ACGT GC ATCTCTGC GATGAACCTC AC ACGTGC GT
451 GCCTCTTATG ATCTGGAT CT CGAG CCATGG TCGAGGCCTG TAGGCGTGTA
501 CGGTGGGAGG CTTA TATAA G CAGAGCTCAA GCTGGCATCC GGTACYKTKR
551 AGRMCAMCCC
```

← ←
NFκB | NFκB | HRE | HRE | HRE | HRE | HRE | HRE | mCMV | Luc⁺ | LV-6-Luc⁺

Based on the cloning strategy, this construct should have a 6NFκB cluster in the reverse orientation, another cluster (either 6NFκB or 8AP-1) and a proximal 6HRE cluster, which was not successfully sequenced.

Appendix 13.8. LV-245-Luc⁺

```
551 AGGGGGGATT GGGGGGGTACA GTGCAGGGGA AAGAATAGTA GACATAATAG
601 CAACAGACAT ACAAACTAAA GAATTACAAA AACAAATTAC AAAAATTCAA
651 AATTTTATCG ATCACGAGAC TAGCTCGAGA AGCTTGATGA TCCGTTTCGA
701 TCGCGTGCTA G ACGTGG CTA G ACGTGG CTA G GAAATTCC C TAG GGAATTT
751 C CTAGCCCGG G CTCGAGGCC TGTAGGCGTG TACGGTGGGA GGCTTA TATA
801 A GCAGAGCTC AAGCTGGCAT CCGGTACKGT KWAASMCMMC
```

←
HRE | HRE | NFκB | NFκB | mCMV | Luc⁺ | LV-245-Luc⁺

Appendix 13.9. LV-4NFκB-Luc⁺

```
651 ACATAATAGC AACAGACATA CAAACTAAAG AATTACAAAA ACAAATTACA
701 AAAATTCAAA ATTTTATCGA TCACGAGACT AGCTCGAGAA GCTTGATGAT
751 CCGTTTCGAT GGCCAGCTAG CGTGCCTCTT ATGATCTGGA TGGGACTTTC
801 CATCTCTGCG ATGAACCTCA CGGGACTTTC CGTGCCTCTT ATGATCTGGA
851 TGGGACTTTC CATCTCTGCG ATGAACCTCA CGGGACTTTC CGTGCCTCTT
901 ATGATCTGGA TCTCGAGCCA TGGTCGAGGC CTGTAGGCGT GTACGGTGGG
951 AGGCTTATAT AAGCAGAGCT CAAGCTGGCA TCCGGTACGT GAGCACCTAT
```

NFκB | NFκB | NFκB | NFκB | mCMV | Luc⁺   LV-4NFκB-Luc⁺

Appendix 13.10. LV-mCMV-Luc⁺

The start and end of the mCMV promoter is indicated by XhoI (5'-TCGA) and HindIII (5'-AGCT), overhangs respectively (underlined). The TATAA box is highlighted in bold. The start of the luciferase gene is indicated by the NcoI restriction site 5'-CCATGG (underlined).

```
  1 AMCAGGTTRC TCAAARACMR TTACRAAAAT TCARAATTTT ATCGATCACG
 51 AGACTAGCTC GAGAAGCTTG ATGATCCGTT TCGATGCGTG CTAGCCCGGG
101 CTCGAGGCCT GTAGGCGTGT ACGGTGGGAG GCTTATATAA GCAGAGCTCA
151 AGCTTGGCAT TCCGGTACTG TTGGTAAAGC CACCATGGAA GACGCCAAAA
201 ACATAAAGAA AGGCCCGGCG CCATTCTATC CGCTGGAAGA TGGAACCGCT
251 GGAGAGCAAC TGCATAAGGC TATGAAGAGA TACGCCCTGG TTCCTGGAAC
```

mCMV | Luc⁺   LV-mCMV-Luc⁺

Appendix 13.11. LV-SFFV-Luc⁺ (sequenced using the Forward Lenti primer)

The start of the SFFV promoter is indicated by an EcoRI overhang 5'-AATTC (underlined). The start of the luciferase gene is indicated by the NcoI restriction site 5'-CCATGG (underlined).

```
  1 CWWRGGATTA MAAAACAAAT TACAAAAATT CAAAATTTTA TCGATCACGA
 51 GACTAGCTCG AGAAGCTTGA TGATCCGTTT AATTCCTGCA GCCCCGATAA
101 AATAAAAGAT TTTATTTAGT CTCCAGAAAA AGGGGGGAAT GAAAGACCCC
151 ACCTGTAGGT TTGGCAAGCT AGCTGCAGTA ACGCCATTTT GCAAGGCATG
201 GAAAATACC AAACCAAGAA TAGAGAAGTT CAGATCAAGG GCGGGTACAT
251 GAAAATAGCT AACGTTGGGC CAAACAGGAT ATCTGCGGTG AGCAGTTTCG
301 GCCCCGGCCC GGGGCCAAGA ACAGATGGTC ACCGCAGTTT CGGCCCCGGC
```

```
351 CCGAGGCCAA GAACAGATGG TCCCCAGATA TGGCCCAACC CTCAGCAGTT
401 TCTTAAGACC CATCAGATGT TTCCAGGCTC CCCCAAGGAC CTGAAATGAC
451 CCTGCGCCTT ATTTGAATTA ACCAATCAGC CTGCTTCTCG CTTCTGTTCG
501 CGCGCTTCTG CTTCCCGAGC TCTATAAAG AGCTCACAAC CCCTCACTCG
551 GCGCGCCAGT CCTCCGACAG ACTGAGTCGC CCGGGGGGGA TCTGCGATCT
601 AAGTAAGCTT GGCATTCCGG TACTGTTGGT AAAGCCA͟C͟C͟A͟ ͟T͟G͟GAAGACGC
651 CAAAAACATA AAGAAAGGCC CGGCGCCATT CTATCCGCTG GAAGATGGAA
701 CCGCTGGAGA GCAACTGCAT AAGGCTATGA AGAGATACGC CCTGGTTCCT
751 GGAACAATTG CTTTTACAGA TGCACATATC GAGGTGGACA TCACTTACGC
801 TGAGTACTTC GAAATGTCCG TTCGGTTGGC AGAAGCTATG AAACGATATG
851 GGCTGAATAC AAATCACAGA ATCGTCGTAT GCAGTGAAAA CTCTCTTCAA
901 TTCTTTATGC CGGTGTTGGG CGCGTTATTT ATCGGAGTTG CAGTTGCGCC
951 SCGAAACGAC WTTTATAATG AACGTGAATT GCTCAACAGT ATGGGCATTT
```

SFFV Luc⁺   LV-SFFV-Luc⁺

Appendix 14. Sequencing data of lentiviral synthetic promoters expressing the therapeutic mTNFRII-Fc or hIL-1Ra gene

Unless stated otherwise, the lentiviral-synthetic promoter-IL-1Ra and mTNFRII-Fc constructs were sequenced using the Forward Lenti primer which binds upstream of the synthetic promoter. The TFBSs are positioned in the forward orientation and are indicated by a box. The NheI (5'-GCTAGC) and XhoI (5'-CTCGAG) sites are underlined and represent the start and end of the region containing multimerised TFBSs, respectively. The TATAA box is highlighted in bold. The start of the therapeutic genes mTNFRII-Fc or hIL-1Ra is indicated by the underlined BamHI restriction site (5'-GGATCC). Schematic diagrams are presented with the sequencing data for each construct

Appendix 14.1. LV-2-mTNFRII-Fc

```
  1 ATWAGATTAC AAAAACAAAT TACAAAAATT CAAAATTTTA TCGATCACGA
 51 GACTAGCTCG AGAAGCTTGA TGATCCGTTT TAGGCCA͟G͟C͟T͟ ͟A͟GCGTGCCTC
101 TTATGATCTG GAT TGAGTCA ATCTCTGCGA TGAACCTCAC TGAGTCA GTG
151 CCTCTTATGA TCTGGAT ACG TGC ATCTCTG CCGATGAACC TCAC TGAGTC
201 A GTGCCTCTT ATGATCTGGA T ACGTGC ATC TCTGCGATGA ACCTCAC TGA
251 GTCA GTGCCT CTTATGATCT GGAT TGAGTC A ATCTCTGCG ATGAACCTCA
301 C TGAGTCA GT GCCTCTTATG ATCTGGAT TG AGTCA ATCTC TGCGATGAAC
351 CTCAC ACGGC GTGCCTCTTA TGATCTGGAT  GGGACTTTCC  ATCTCTGCGA
```

```
401 TGAACCTCAC GGGACTTTCC GTGCCTCTTA TGATCTGGAT GGGACTTTCC

451 ATCTCTGCGA TGAACCTCAC TGAGTCAGTG CCTCTTATGA TCTGGATCTC

501 GAGCCATGGT CGAGGCCTGT AGGCGTGTAC GGTGGGAGGC TTATATAAGC

551 AGAGCTCAGG ATCCATGTAC AGGATGCAAC TCCTGTCTTG CATTGCACTA

601 AGTCTTGCAC TTGTCACGAA TTCCACCATG GCGCCCGCCG CCCTCTGGGT

651 CGCGCTGGTC TTCGAACTGC AGCTGTGGGC CACCGGGCAC ACAGTGCCCG

701 CCCAGGTTGT CTTGACACCC TACAAACCGG AACCTGGGTA CGAGTGCCAG

751 ATCTCACAGG AATACTATGA CAGGAAGGCT CAGATGTGCT GTGCTAAGTG

801 TCCTCCTGGC CAATATGTGA AACATTTCTG CAACAAGACC TCAGACACCG
```

AP-1 | AP-1 | HRE | AP-1 | HRE | AP-1 | AP-1 | AP-1 | AP-1 | HRE | NFκB | NFκB | NFκB | AP-1 | mCMV | mTNFRII-Fc   LV-2-mTNFRII-Fc

Appendix 14.2. LV-2-hIL-1-Ra

```
  1 GAWAGATTAC AAAAACAAAT TACAAAAATT CAAAATTTTA TCGATCACGA

 51 GACTAGCTCG AGAAGCTTGA TGATCCGTTT GTAGGCCAGC TAGCGTGCCT

101 CTTATGATCT GGATTGAGTC AATCTCTGCG ATGAACCTCA CTGAGTCAGT

151 GCCTCTTATG ATCTGGATAC GTGCATCTCT GCCGATGAAC CTCACTGAGT

201 CAGTGCCTCT TATGATCTGG ATACGTGCAT CTCTGCGATG AACCTCACTG

251 AGTCAGTGCC TCTTATGATC TGGATTGAGT CAATCTCTGC GATGAACCTC

301 ACTGAGTCAG TGCCTCTTAT GATCTGGATT GAGTCAATCT CTGCGATGAA

351 CCTCACACGG CGTGCCTCTT ATGATCTGGA TGGGACTTTC CATCTCTGCG

401 ATGAACCTCA CGGGACTTTC CGTGCCTCTT ATGATCTGGA TGGGACTTTC

451 CATCTCTGCG ATGAACCTCA CTGAGTCAGT GCCTCTTATG ATCTGGATCT

501 CGAGCCATGG TCGAGGCCTG TAGGCGTGTA CGGTGGGAGG CTTATATAAG

551 CAGAGCTCAG GATCCATGGA AATCTGCAGA GGCCTCCGCA GTCACCTAAT

601 CACTCTCCTC CTCTTCCTGT TCCATTCAGA GACGATCTGC CGACCCTCTG

651 GGAGAAAATC CAGCAAGATG CAAGCCTTCA GAATCTGGGA TGTTAACCAG

701 AAGACCTTCT ATCTGAGGAA CAACCAACTA GTTGCTGGAT ACTTGCAAGG

751 ACCAAATGTC AATTTAGAAG AAAAGATAGA TGTGGTACCC ATTGAGCCTC
```

AP-1 | AP-1 | HRE | AP-1 | HRE | AP-1 | AP-1 | AP-1 | AP-1 | HRE | NFκB | NFκB | NFκB | AP-1 | mCMV | hIL-1Ra   LV-2-hIL-1Ra

Appendix 14.3. LV-9-mTNFRII-Fc

```
  1 ACWTAAGAAT ACAAAAACAA TTACAAAAAT TCAAAATTTT ATCGATCACG

 51 AGACTAGCTC GAGAAGCTTG ATGATCCGTT TGTAGGCCAG CTAGCGTGCC

101 TCTTATGATC TGGATTGAGT CAATCTCTGC GATGAACCTC ACACGTGCGT

151 GCCTCTTATG ATCTGGATGG GACTTTCCAC CTCTGCGATG AACCTCACAC

201 GTGCGTGCCT CTTATGATCT GGATGGGACT TTCCATCTCT GCGATGAACC
```

```
251 TCACTGAGTC AGTGCCTCTT ATGATCTGGA TTGAGTCAAT CTCTGCGATG
301 AACCTCACGG GACTTTCCGT GCCTCTTATG ATCTGGATAC GTGCATCTCT
351 GCGATGAACC TCACTGAGTC AGTGCCTCTT ATGATCTGGA TGGGACTTTC
401 CATCTCTGCG ATGAACCTCA CACGTGCGTG CCTCTTATGA TCTGGATTGA
451 GTCAATCTCT GCGATGAACC TCACACGTGC GTGCCTCTTA TGATCTGGAT
501 CTCGAGCCAT GGTCGAGGCC TGTAGGCGTG TACGGTGGGA GGCTTATATA
551 AGCAGAGCTC AGGATCCATG TACAGGATGC AACTCCTGTC TTGCATTGCA
601 CTAAGTCTTG CACTTGTCAC GAATTCCACC ATGGCGCCCG CCGCCCTCTG
651 GGTCGCGCTG GTCTTCGAAC TGCAGCTGTG GGCCACCGGG CACACAGTGC
701 CCGCCCAGGT TGTCTTGACA CCCTACAAAC CGGAACCTGG GTACGAGTGC
751 CAGATCTCAC AGGAATACTA TGACAGGAAG GCTCAGATGT GCTGTGCTAA
801 GTGTCCTCCT GGCCAATATG TGAAACATTT CTGCAACAAG ACCTCAGACA
```

AP-1 | HRE | NFκB | HRE | NFκB | AP-1 | AP-1 | NFκB | HRE | AP-1 | NFκB | HRE | AP-1 | HRE | mCMV | mTNFRII-Fc   LV-9-mTNFRII-Fc

Appendix 14.4. LV-9-hIL-1-Ra

```
  1 CCWAAGGAAT TAMAAAAACA AATTACAAAA ATTCAAAATT TTATCGATCA
 51 CGAGACTAGC TCGAGAAGCT TGATGATCCG TTTTAGGCCA GCTAGCGTGC
101 CTCTTATGAT CTGGATTGAG TCAATCTCTG CGATGAACCT CACACGTGCG
151 TGCCTCTTAT GATCTGGATG GGACTTTCCA CCTCTGCGAT GAACCTCACA
201 CGTGCGTGCC TCTTATGATC TGGATGGGAC TTTCCATCTC TGCGATGAAC
251 CTCACTGAGT CAGTGCCTCT TATGATCTGG ATTGAGTCAA TCTCTGCGAT
301 GAACCTCACG GGACTTTCCG TGCCTCTTAT GATCTGGATA CGTGCATCTC
351 TGCGATGAAC CTCACTGAGT CAGTGCCTCT TATGATCTGG ATGGGACTTT
401 CCATCTCTGC GATGAACCTC ACACGTGCGT GCCTCTTATG ATCTGGATTG
451 AGTCAATCTC TGCGATGAAC CTCACACGTG CGTGCCTCTT ATGATCTGGA
501 TCTCGAGCCA TGGTCGAGGC CTGTAGGCGT GTACGGTGGG AGGCTTATAT
551 AAGCAGAGCT CAGGATCCAT GGAAATCTGC AGAGGCCTCC GCAGTCACCT
601 AATCACTCTC CTCCTCTTCC TGTTCCATTC AGAGACGATC TGCCGACCCT
651 CTGGGAGAAA ATCCAGCAAG ATGCAAGCCT TCAGAATCTG GGATGTTAAC
701 CAGAAGACCT TCTATCTGAG GAACAACCAA CTAGTTGCTG GATACTTGCA
751 AGGACCAAAT GTCAATTTAG AAGAAAGAT AGATGTGGTA CCCATTGAGC
801 CTCATGCTCT GTTCTTGGGA ATCCATGGAG GGAAGATGTG CCTGTCCTGT
```

AP-1 | HRE | NFκB | HRE | NFκB | AP-1 | AP-1 | NFκB | HRE | AP-1 | NFκB | HRE | AP-1 | HRE | mCMV | hIL-1Ra   LV-9-hIL-1Ra

Appendix 14.5. LV-mCMV-mTNFRII-Fc

The start of the mCMV promoter is indicated by the SnabI overhang 5'-GTA (underlined). The start of mTNFRII-Fc gene is indicated by the BamHI restriction site 5'- GGATCC (underlined). The TATAA box is highlighted in bold.

```
  1 AAMMMWWARG RWTTWCAAAA AMAAATTACA AAAAATTCAA AATTTTATCG
 51 ATCACGAKAC TAGCTCGAGA AGCTTGATGA TCCGTTTGTA GCCTGTAGGC
101 GTGTACGGTG GGAGGCTTAT ATAACWTAGC TCGGATCCAT GTACAGGATG
151 CAACTCCTGT CTTGCATTGC ACTAAGTCTT GCACTTGTCA CGAATTCCAC
201 CATGGCGMCC GCCGYCCTCT GGGTCGCGCT GGTCTTCRAA CTGCAGCTGT
251 GGGCCACCGG GCACACAGTG CCCGCCCAGG TTGTCTTGAC ACCCT
```

mCMV mTNFRII-Fc   LV-mCMV-mTNFRII-Fc


Appendix 14.6. LV-mCMV-hIL-1Ra

The start of the mCMV promoter is indicated by the SnabI overhang 5'-GTA (underlined). The start of hIL-1Ra gene is indicated by the BamHI restriction site 5'- GGATCC (underlined). The TATAA box is highlighted in bold.

```
  1 TWMYYWWRGR AWTTCCAAAA AACAAATTAC AAAAAATTCA AAATTTTATC
 51 GATCACGAKA CTAGCTCGAG AAGCTTGATG ATCCCGTTTG TAGCCTGTAG
101 GCGTGTACGG TGGGAGGCTT ATATAACCMT AGCTCGGATC CATGGAAATC
151 TGCAGAGGCC TCCGCAGCCY C
```

mCMV hIL-1Ra   LV-mCMV-hIL-1Ra

Appendix 14.7. LV-SFFV-mTNFRII-Fc

The start of the SFFV promoter is indicated by the SnabI overhang 5'-GTA (underlined). The start of mTNFRII-Fc gene is indicated by the BamHI restriction site 5'- GGATCC (underlined).

```
  1 AYWWKRGGRA TTAMAAAAAM AAATTACAAA AATTCAAAAT TTTATCGATC
 51 ACGAGACTAG CTCGAGAAGC TTGATGATCC GTTTGTAAAT TCCTGCAGCC
101 CCGATAAAAT AAAAGATTTT ATTTAGTCTC CAGAAAAAGG GGGGAATGAA
151 AGACCCCACC TGTAGGTTTG GCAAGCTAGC TGCAGTAACG CCATTTTGCA
201 AGGCATGGAA AAATACCAAA CCAAGAATAG AGAAGTTCAG ATCAAGGGCG
251 GGTACATGAA AATAGCTAAC GTTGGGCCAA ACAGGATATC TGCGGTGAGC
301 AGTTTCGGCC CCGGCCCGGG GCCAAGAACA GATGGTCACC GCAGTTTCGG
351 CCCCGGCCCG AGGCCAAGAA CAGATGGTCC CCAGATATGG CCCAACCCTC
```

```
401 AGCAGTTTCT TAAGACCCAT CAGATGTTTC CAGGCTCCCC CAAGGACCTG
451 AAATGACCCT GCGCCTTATT TGAATTAACC AATCAGCCTG CTTCTCGCTT
501 CTGTTCGCGC GCTTCTGCTT CCCGAGCTCT ATAAAAGAGC TCACAACCCC
551 TCACTCGGCG CGCCAGTCCT CCGACAGACT GAGTCGCCCG GGGGGGATCT
601 GCGATCTAAG TAAGCTTGGC ATTCCGGTAC TGTTGGTAAA GCCACCGGAT
651 CCATGTACAG GATGCAACTC CTGTCTTGCA TTGCACTAAA GTCTTGCACT
701 TGTCACSAAT TCCACCATGG CR
```

SFFV | mTNFRII-Fc   LV-SFFV-mTNFRII-Fc

Appendix 14.8. LV-SFFV-hIL-1Ra

The start of the SFFV promoter is indicated by the SnabI overhang 5'-GTA (underlined). The start of hIL1-Ra gene is indicated by the BamHI restriction site 5'- GGATCC (underlined).

```
  1 AWYCTWAAAG YATTRCAAAA CAAATTACAA AAATTCAAAA TTTTATCGAT
 51 CACGAGACTA GCTCGAGAAG CTTGATGATC CGTTTGTAAA TTCCTGCAGC
101 CCCGATAAAA TAAAAGATTT TATTTWKTCT CCAGAAAAAG GGGGGAATGA
151 AAGACCCCAC CTGTAGGTTT GGCAAGCTAG CTGCAGTAAC GCCATTTTGC
201 AAGGCATGGA AAAATACCAA ACCAAGAATA GAGAAGTTCA GATCAAGGGC
251 GGGTACATGA AAATAGCTAA CGTTGGGCCA AACAGGATAT CTGCGGTGAG
301 CAGTTTCGGC CCCGGCCCGG GGCCAAGAAC AGATGGTCAC CGCAGTTTCG
351 GCCCCGGCCC GAGGCCAAGA ACAGATGGTC CCCAGATATG GCCCAACCCT
401 CAGCAGTTTC TTAAGACCCA TCAGATGTTT CCAGGCTCCC CCAAGGACCT
451 GAAATGACCC TGCGCCTTAT TTGAATTAAC CAATCAGCCT GCTTCTCGCT
501 TCTGTTCGCG CGCTTCTGCT TCCCGAGCTC TATAAAAGAG CTCACAACCC
551 CTCACTCGGC GCGCCAGTCC TCCGACAGAC TGAGTCGCCC GGGGGGGATC
601 TGCGATCTAA GTAAGCTTGG CATTCCGGTA CTGTTGGTAA AGCCACCGGA
651 TCCATGGAAA TCTGCAGAGG CCTCCGCAGT CACCTAATCA CTCTCCTCCT
701 CTTCCTGTTC CATTCAGAGA CGATCTGCCG ACCCTCTGGG AGAAAATCCA
751 GCAAGATGCA AGCCTTCAGA ATCTGGGATG TTAACCAGAA GACCTTTCTA
801 TCTGAGGAAC AACCAACTAG TTGCTGGATA CTTGCAAGGA CCAAATGTCA
```

SFFV | hIL-1Ra   LV-SFFV-hIL-1Ra

Appendix 15. Absolute Real-Time qPCR standard curve plot, amplification plot and dissociation curve of miR-23b and control miRNAs (miR-16, U6, miR-17-5p, miR-103, miR-191). All standard curve plots show very good R^2 values of 0.98- 1.0 (two decimal places). Amplification plots show that miRNA PCR products (within samples) are within the standards.
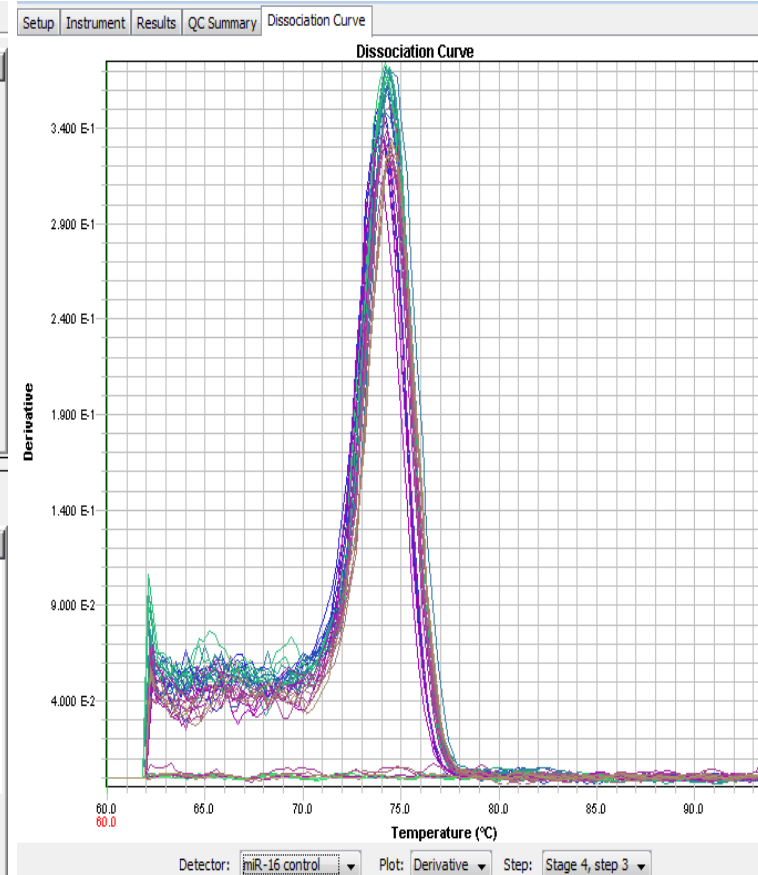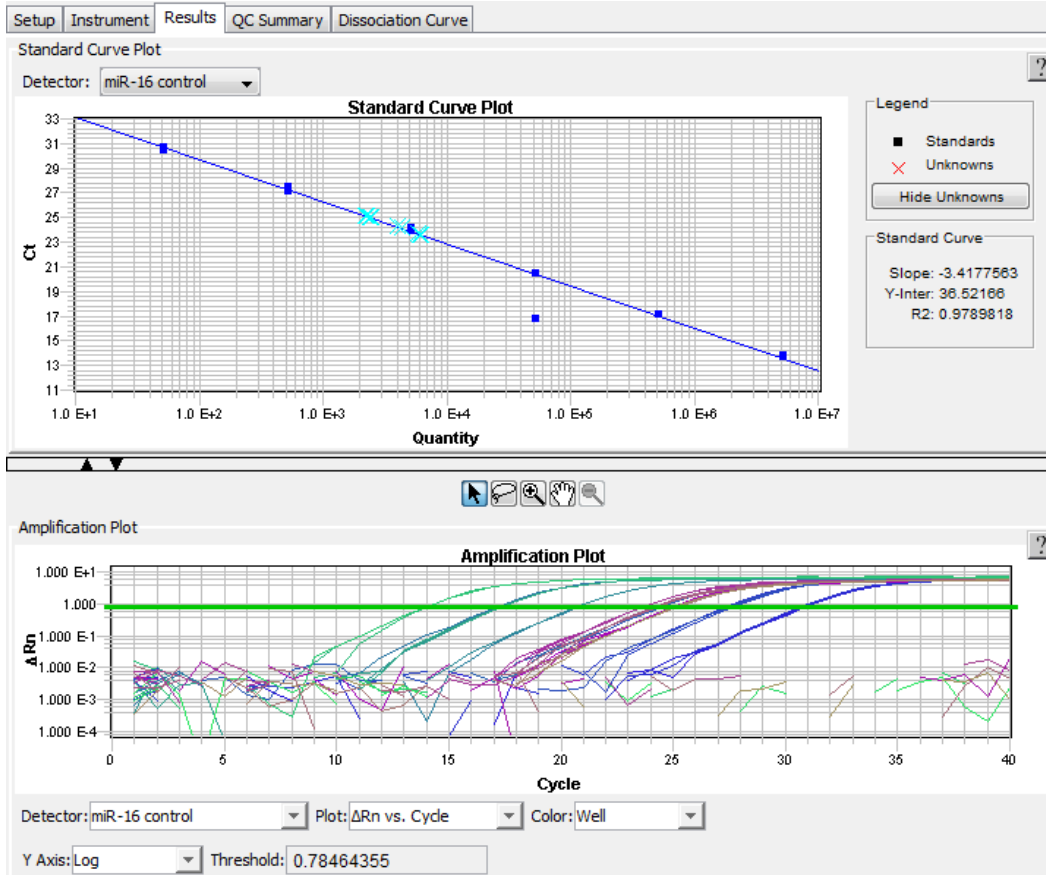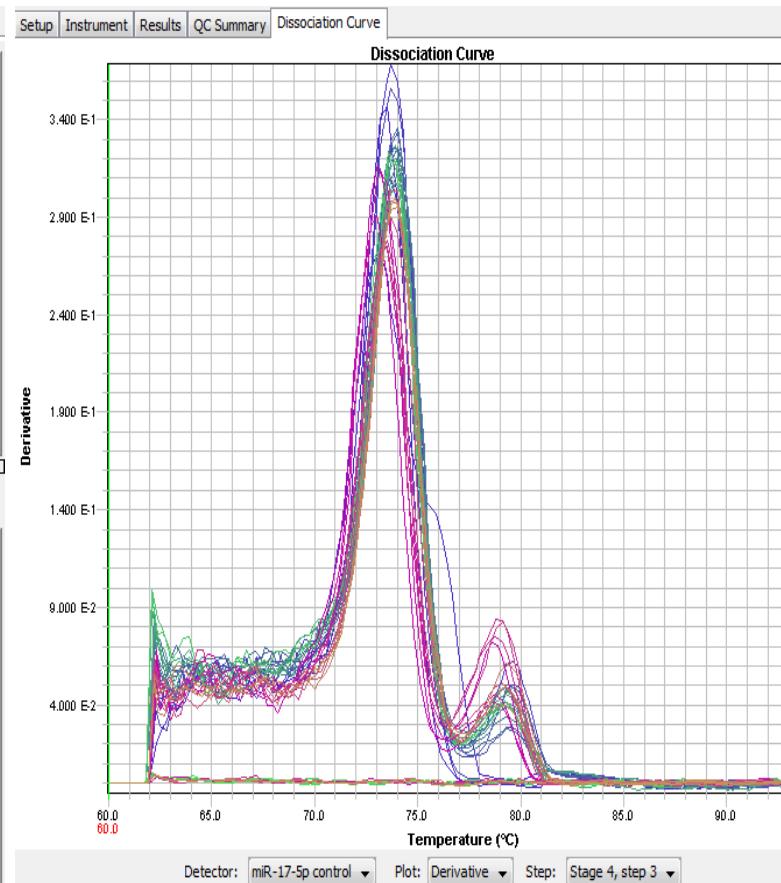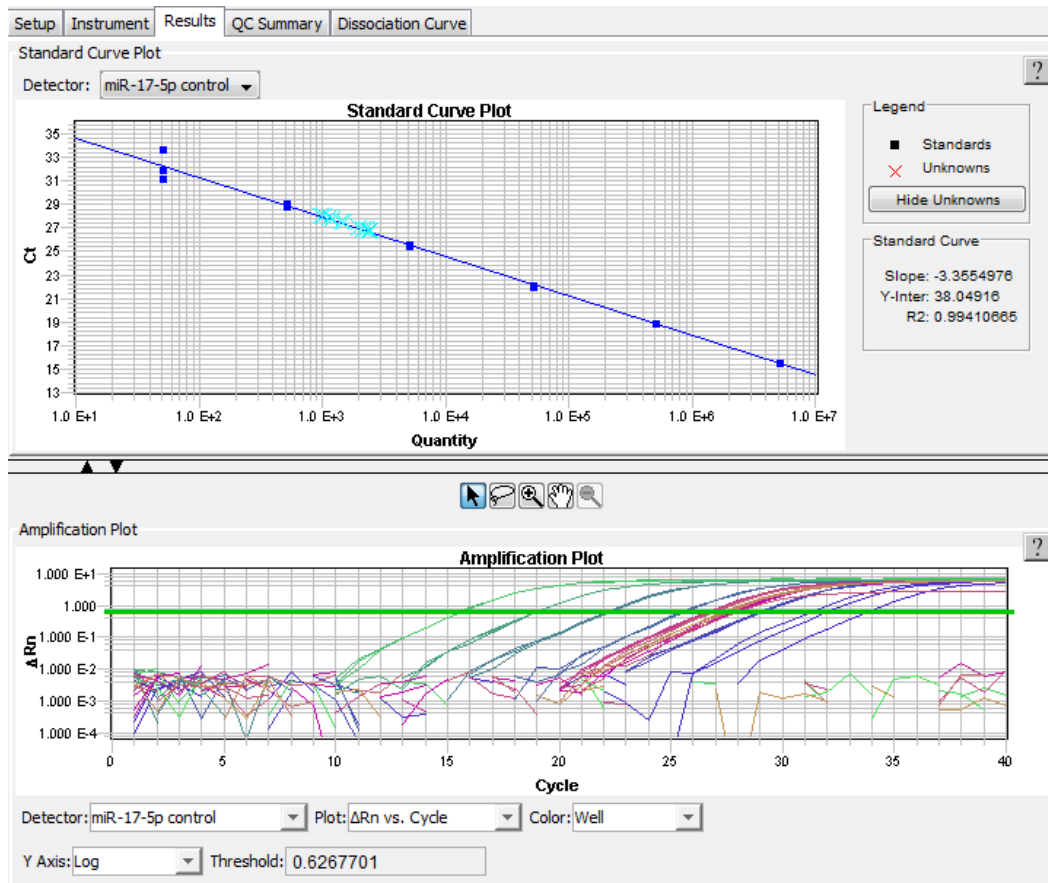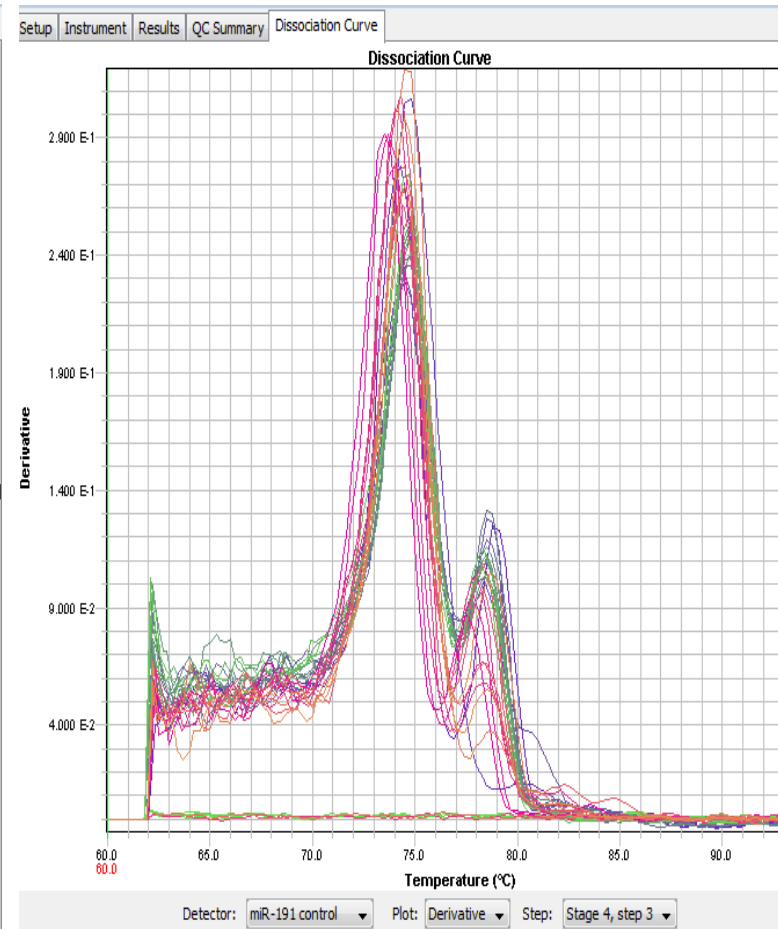
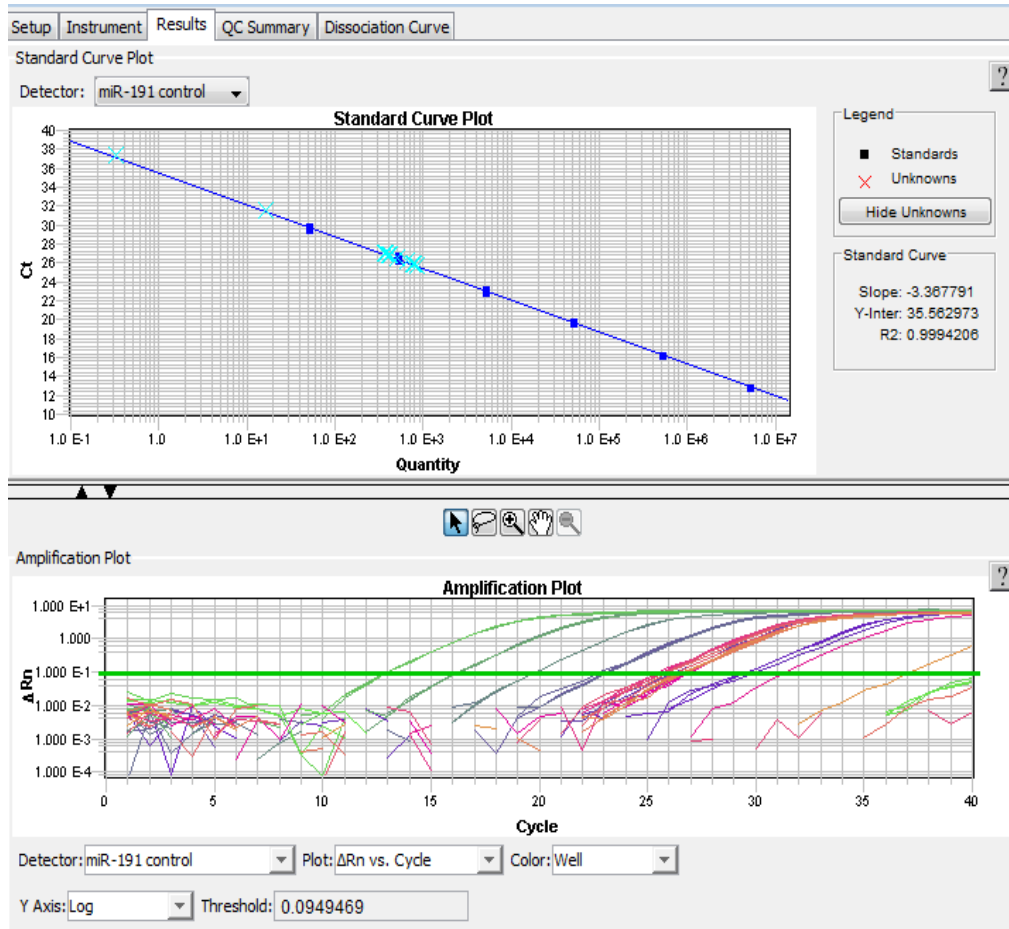<u>**miR-23b**</u>
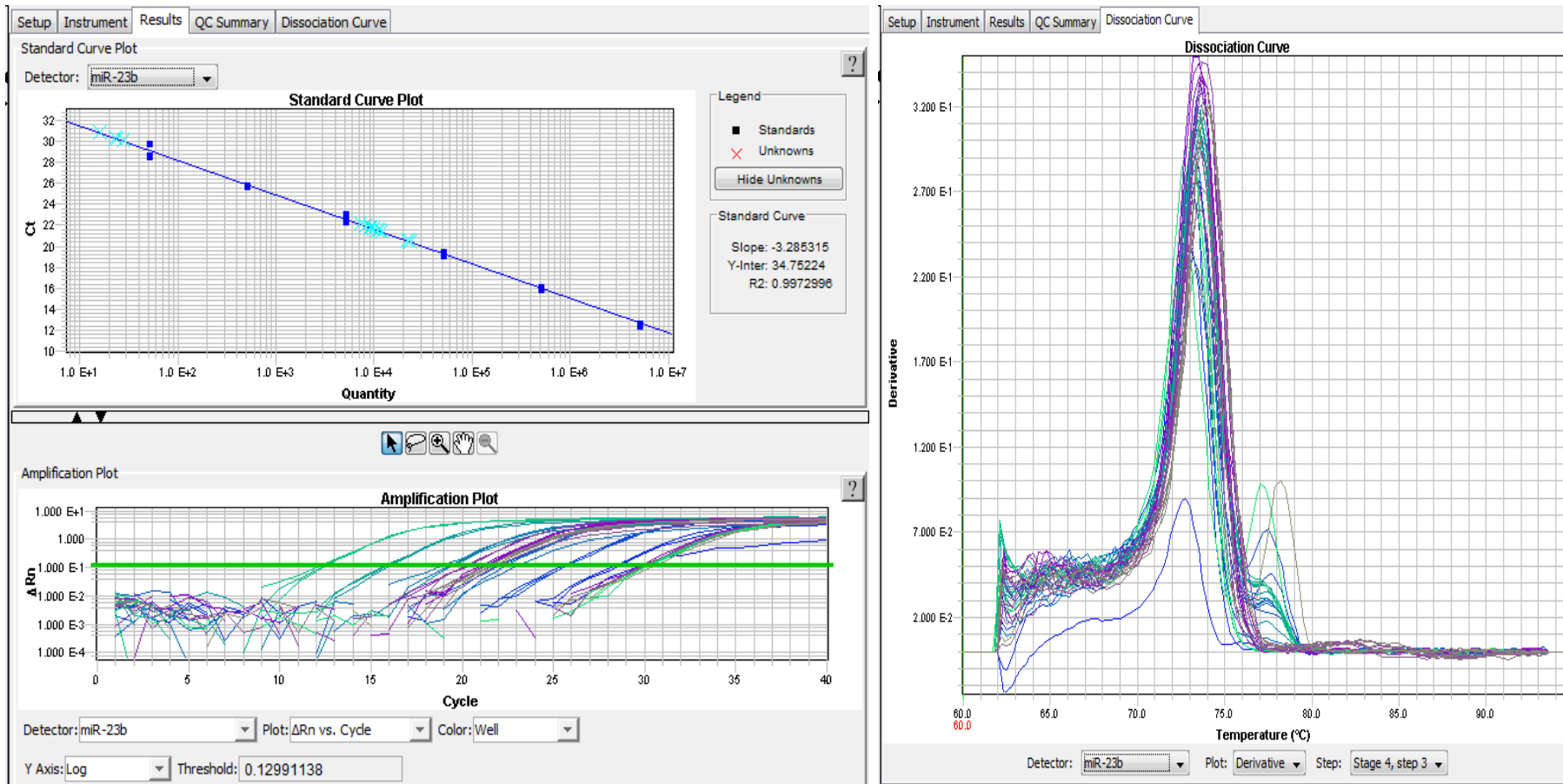
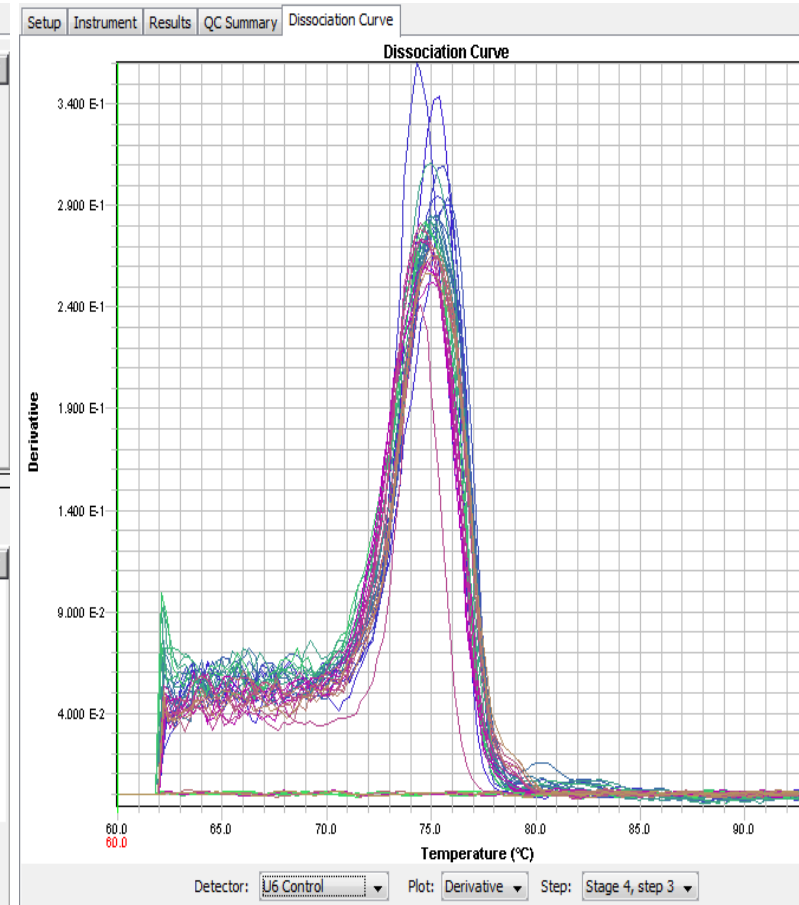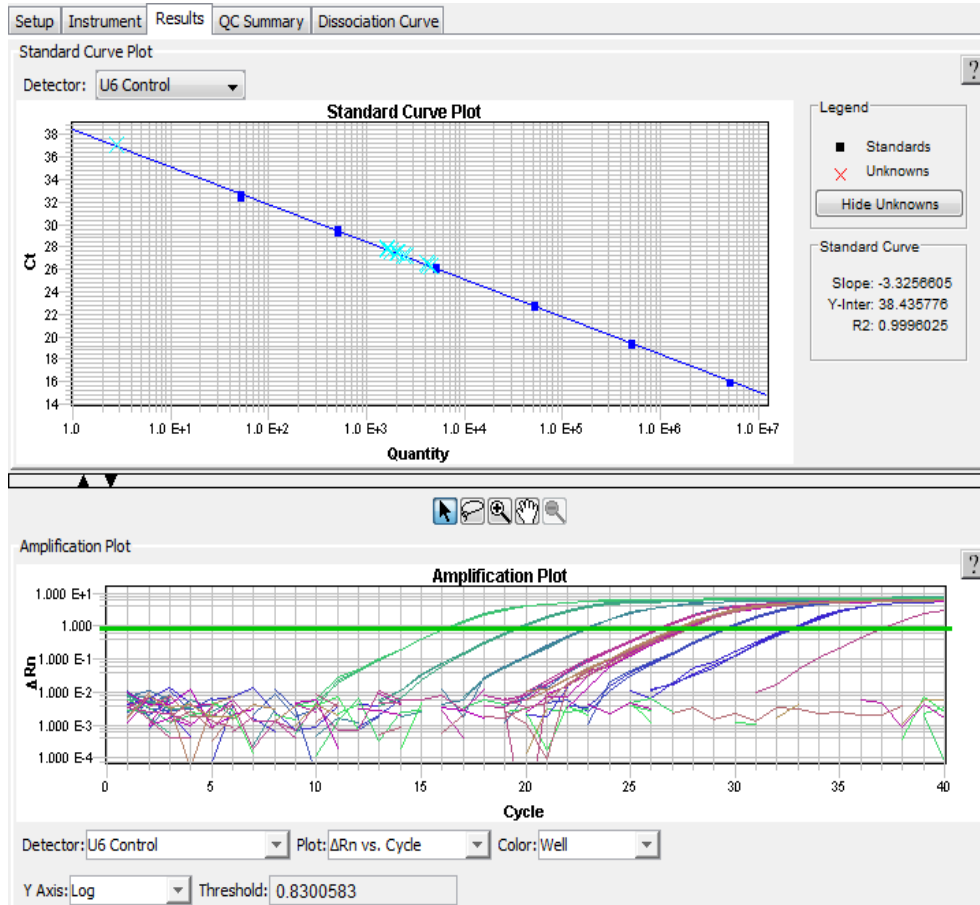## miR-103 control
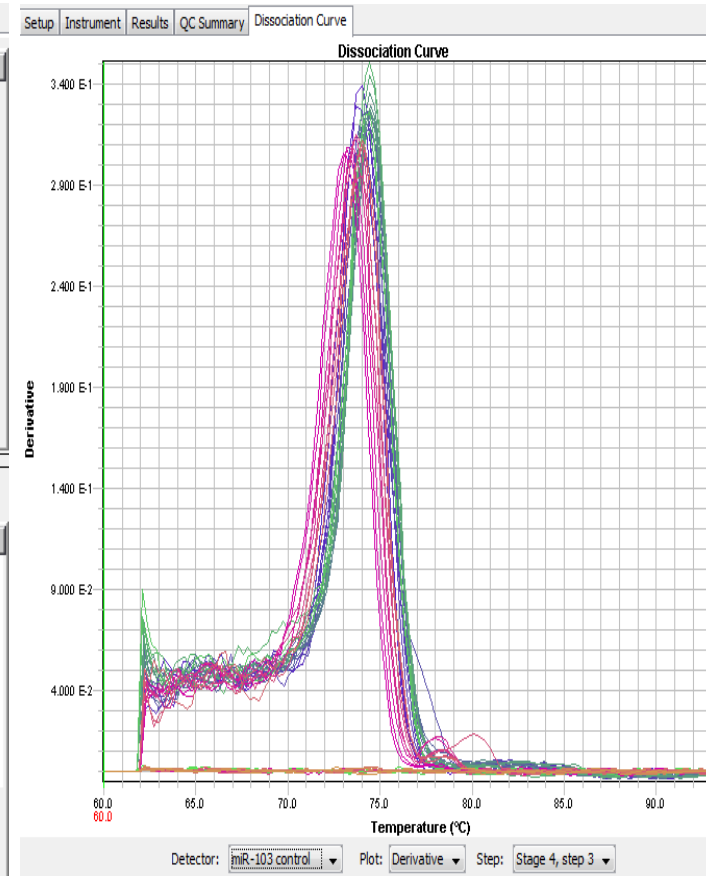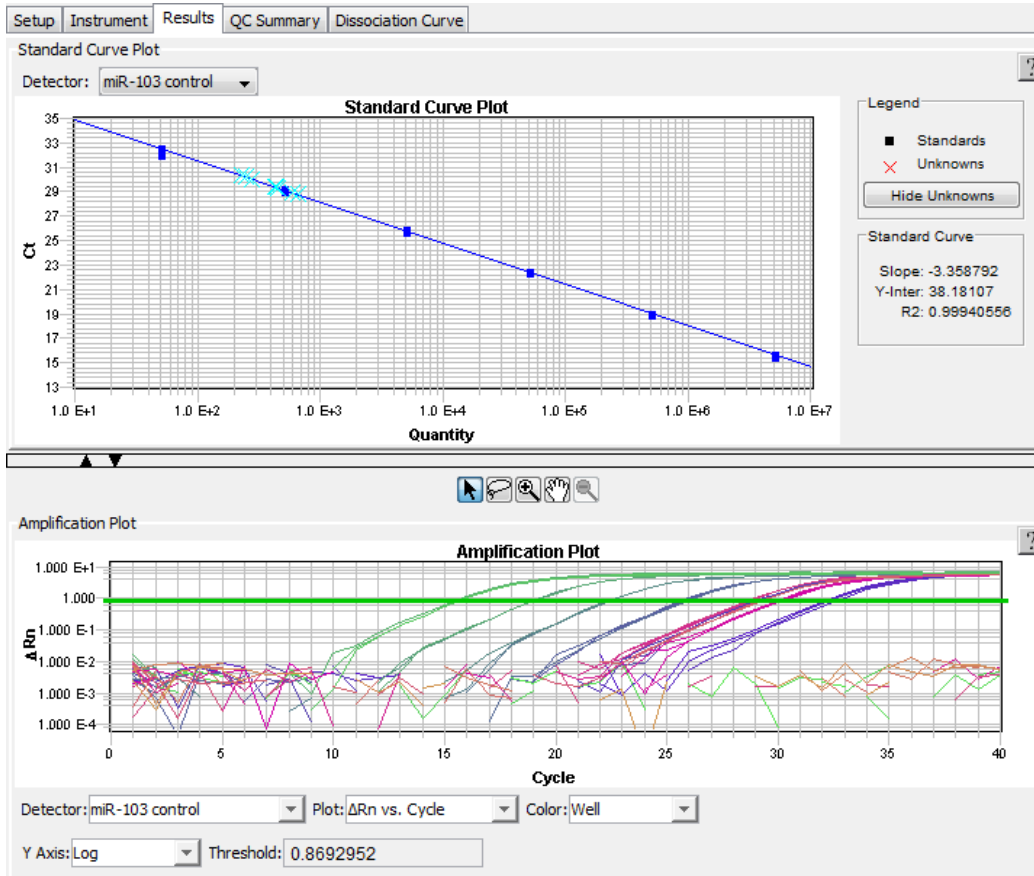
miR-16 control

miR-17-5p control

miR-191 control

U6 control

Scientific Communications

Poster presentations:

1. Hodan H.A. Mohamed, Yuti Chernajovsky and David J. Gould. 'Generation and Characterisation of Synthetic Inflammation-Responsive Promoters'. William Harvey Day, Queen Mary University of London, UK, 2011.

2. Hodan H.A. Mohamed, Yuti Chernajovsky and David J. Gould. 'Generation and Characterisation of Synthetic Inflammation-Responsive Promoters'. 6th Meeting of Gene Therapy of Arthritis and Related Disorders, Queen Mary University of London, UK, 2011.

3. Hodan H.A. Mohamed, Yuti Chernajovsky and David J. Gould. 'Development of Novel Transcriptionally Responsive Promoters for Therapeutic Gene Expression in Direct Response to Disease Activity'. William Harvey Annual Review Day, Queen Mary University of London, UK, 2012.

4. Hodan H.A. Mohamed, Yuti Chernajovsky and David J. Gould. 'Development of Novel Transcriptionally Responsive Promoters for Therapeutic Gene Expression during Disease Activity'. BSGCT Conference, Royal Holloway University of London, UK, 2013.

5. Hodan H.A. Mohamed, Yuti Chernajovsky and David J. Gould. 'Development of Novel Transcriptionally Responsive Promoters for Therapeutic Gene Expression during Disease Activity'. BSGCT Conference, Institute of Child Health, UCL, London, UK, 2014.

Oral presentations:

1. Hodan H.A. Mohamed, Yuti Chernajovsky and David J. Gould. 'Development of Novel Transcriptionally Responsive Synthetic Promoters for Therapeutic Gene Expression during Disease Activity'. William Harvey Annual Review Day, Queen Mary University of London, UK, July 2013.

Nominated for Young Investigator Award. Oral presentation at the William Harvey Annual Review Day, January 2014:

2. Hodan H.A. Mohamed, Yuti Chernajovsky and David J. Gould. 'Development of Novel Transcriptionally Responsive Synthetic Promoters for Therapeutic Gene Expression during Disease Activity'. William Harvey Annual Review Day, Queen Mary University of London, UK, January 2014.

Publication:

Mohamed HHA, Chernajovsky Y & Gould DJ. Compact multi-responsive synthetic promoters built by an optimised assembly PCR approach display synergistic activation and are suited to gene therapy application. Submitted for publication July 2014.