By Petra Disterer

A thesis submitted in partial fulfilment of the requi

for the degree of Doctor of Philosophy

Royal Free and University College Medical Sch

Department of Medicine, University College Lo

2008

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Oligonucleotide-mediated gene editing of Apolipoprotein A-I

Apolipoprotein A-I (ApoA-I) is the major protein constituent of high density lipoprotein (HDL) and controls reverse cholesterol transport, an important process in preventing atherosclerosis. A natural point mutation, ApoA-I_{Milano} (ApoA-I_M) enhances the atheroprotective potential of HDL. Here, I attempt to introduce this specific modification into the genome of mammalian cells using the gene therapy strategy of oligonucleotide-mediated gene editing. I showed successful *APOA-I* gene editing in recombinant Chinese hamster ovary (CHO-AI) and human hepatocellular carcinoma (HepG2) cells by polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP) analysis and sequencing. However, I was unable to isolate gene-edited cell clones or quantify gene editing. Therefore, I established a recombinant CHO cell line expressing mutated non-fluorescent enhanced green fluorescent protein (mEGFP) that demonstrates successful gene editing through restoration of fluorescence.

Using flow cytometry, I studied the influence of transfection reagents and oligonucleotide design on gene editing efficiency and viability of gene-edited cells. I found that Lipofectamine 2000 generated higher initial editing efficiencies with phosphorothioate (PTO) or locked nucleic acid (LNA) modifications, but unmodified oligonucleotides produced significantly more gene-edited clones. I also investigated ways of increasing editing efficiencies or selecting for gene-edited cells independent of the target gene and demonstrated that I had introduced three specific nucleotide alterations by Southern blotting of genomic DNA from gene-edited cell clones. I established that the phenotype in these clones was unstable due to epigenetic down-regulation which was not specific to the gene-edited allele. Variegated gene expression in three mEGFP cell lines demonstrated that gene editing was positively associated with target gene expression. Isolation of ApoA-I_M gene-edited cells using flow cytometry failed because an ApoA-I_M homodimer antibody was not specific.

In summary, I demonstrated that oligonucleotide-mediated gene editing produces stable gene-edited cells in a reporter gene system. However, further research is needed before this method can be applied to non-selectable genes.

Declaration

I, Petra Disterer, confirm that the work presented in this thesis is my own and where information has been derived from other sources, I confirm that this has been indicated in the thesis. No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or institute of learning.

Petra Disterer

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

In particular, I would like to thank my supervisor Professor Jim Owen for providing inspiration, encouraging my scientific independence and always being available for guidance, support or scientific discussion. My secondary supervisor, Dr. Paul Simons, has given excellent technical and scientific advice and provided additional guidance during my studies.

I am greatly indebted to Prof. Stefan Krauss from the Cellular and Genetic Therapy research group in Oslo, Norway, for the opportunity to spend several months in his laboratory and to all the members of his research group for making my stay highly enjoyable as well as very educational. In particular, I would like to thank Dr. Petter Angell Olsen for his continuing help and advice. Also, Dr. Rafael Yáñez from the Royal Holloway, University of London, provided very helpful constructive criticism of my work during my upgrade review which led to the results in chapters 7 and 8. Thanks go to Dr. Mark Lowdell and Prof. Vince Emery from the Royal Free and University College Medical School for permitting me the use of their flow cytometers. Dr. Mark Lowdell and Dr. Derek Davies, Institute of Cancer Research provided invaluable technical expertise in flow cytometry, while Dr. Laura Bazely and Carolyn Koh performed the fluorescence activated cell sorting. Dr. Maya Thanou from Imperial College London collaborated on the transfection reagent optimization.

During my studies, I have also benefited enormously from the friendship, help and support of my lab mates Dr. Vanessa Evans, Eyman Osman and Dr. Ioannis Papaioannou as well as Dr. Meleri Jones, Sherri-Ann Chalmers and all other members of the Mike Jacobs and liver groups. This work would not have been possible without the financial support of the Medical Research Council and the Centre of Hepatology, Royal Free and University College Medical School.

Finally, I would like to thank my friends for being interested and supportive when all I talked about was my work during the few times I saw or spoke to them in the last two years; my family, especially my father, Josef Disterer, for encouraging me to be what I wanted to be and giving me a little push when I needed it; my long-suffering partner David Woods for his emotional support, tolerating my long work hours, feeding me whenever I did turn up, giving me back rubs and helping with figure preparation.

ABBREVIATIONS

2A	oligonucleotide containing two mismatches to target sequence
3A	oligonucleotide containing three mismatches to target sequence
21G	21 gauge
AAV	adeno-associated virus
ABCA1	ATP-binding cassette transporter A1
acLDL	acetylated low density lipoprotein
ApoA-I	apolipoprotein A-I
ApoA-I_M	apolipoprotein A-I_{Milano}
ApoB	apolipoprotein B
ApoE	apolipoprotein E
ATM	ataxia telangiectasia mutated
ATP	adenosine triphosphate
ATR	ataxia telangiectasia and Rad3 related
BA	bile acid
bp	base pairs
BSA	bovine serum albumin
CDAN	N-1-cholesteryloxycarbonyl-3,7-diazanononane-1,9-diamine
CE	cholesteryl ester
CETP	cholesteryl ester transfer protein
CHD	coronary heart disease
CHO	Chinese hamster ovary cells
CHO-AI	Chinese hamster ovary cells expressing apolipoprotein A-I
CHO-AIM	Chinese hamster ovary cells expressing apolipoprotein A-I_{Milano}
CMV	cytomegalovirus
cpt-cAMP	8-(4-chlorophenylthio)adenosine 3',5'-cyclic monophosphate
Ctr	control oligonucleotide with no mismatches to the target sequence
CVD	cardiovascular disease
DAPI	4',6-diamidino-2-phenylindole
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DH5α	*E.coli* strain with genotype *fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*
dhfr	dihydrofolate reductase
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNA-PKcs	DNA-dependent protein kinase catalytic subunit
dNTP	deoxynucleotide triphosphate
DOPE	1,2-dioleoyl-phosphatidylethanolamine
DOTAP	1,2-dioleoyl-3-trimethylammonium propane
DOTMA	N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride
DraI	restriction endonuclease I isolated from *Deinococcus radiophilus*
ds	double-stranded
DSB	double-strand break
dTTP	deoxythymidine triphosphate
edEGFP	gene-edited enhanced green fluorescent protein (fluorescent)
EDTA	ethylenediaminetetraacetate

EGFP	enhanced green fluorescent protein
ERCC	excision repair cross-complementation group
FACS	fluorescence activated cell sorting
FBS	fetal bovine serum
FC	free cholesterol
FokI	restriction endonuclease I isolated from *Flavobacterium okeanokoites*
FSC	forward scatter
HaeII	restriction endonuclease II isolated from *Haemophilus aegyptius*
HDL	high density lipoprotein
HEK293T	human embryonic kidney cells containing the SV40 large T antigen
HEPES	2-hydroxyethyl-1-piperazine ethane sulphonic acid
HepG2	human hepatocellular carcinoma cells
HIV	human immunodeficiency virus
HSV-1	herpes simplex virus-1
HR	homologous recombination
HT	hypoxanthine and thymidine
hu	human
IgG	immunoglobulin class G
IMT	intima media thickness
IPTG	isopropyl β-D-1-thiogalactopyranoside
IRES	internal ribosome entry site
kb	kilo bases
lacZ	β-galactosidase
LB	lysogeny broth
LCAT	lecithin:cholesterol acyltransferase
LDL	low density lipoprotein
LDLR	LDL receptor
LNA	locked nucleic acid
LPEI	linear polyethylenimine
LPL	lipoprotein lipase
mEGFP	mutated enhanced green fluorescent protein (non-fluorescent)
MEM	minimum essential medium (Eagle's)
MfeI	restriction endonuclease I isolated from *Mycoplasma fermentas*
MFI	mean fluorescent intensity (geometric mean)
MLH1	mutL homolog 1
mmLDL	minimally modified low density lipoprotein
MMR	mismatch repair
mRNA	messenger RNA
miRNA	micro RNA
MSH2	mutS homolog 2
NADPH	reduced form of nicotinamide adenine dinucleotide phosphate
NdeI	restriction endonuclease I isolated from *Neisseria denitrificans*
NER	nucleotide excision repair
NHEJ	non-homologous end joining
NT	oligonucleotide complementary to the non-transcribed strand
OD	optical density
oxLDL	oxidized low density lipoprotein
PBS	phosphate-buffered saline
PC	phosphatidylcholine
PCR	polymerase chain reaction
PEI	polyethylenimine

Pfu	DNA polymerase isolated from *Pyrococcus furiosus*
pH	pondus hydrogeni or $-\log_{10}[H]^+$
PI	propidium iodide
PL	phospholipid
PLTP	phospholipid transfer protein
POPC	1-palmitoyl-2-oleoyl phosphatidylcholine
PPARα	peroxisome proliferators-activated receptor type α
PTO	phosphorothioate
RCT	reverse cholesterol transport
RFLP	restriction fragment length polymorphism
RFP	red fluorescent protein
rHDL	reconstituted high density lipoprotein
RNA	ribonucleic acid
RNAi	RNA interference
SceI	endonuclease I isolated from *Saccharomyces cerevisiae*
SD	standard deviation
SDS	sodium dodecyl sulphate
SFHR	small fragment homologous replacement
siRNA	small interfering ribonucleic acid
SNP	single nucleotide polymophism
SOC	super optimal catabolite repression broth
SR-A	scavenger receptor class A
SR-B1	scavenger receptor class B type I
SSA	single-strand annealing
SSC	side scatter
ssODN	single-stranded oligodeoxyribonucleotide
T	oligonucleotide complementary to the transcribed strand
Taq	DNA polymerase isolated from *Thermus aquaticus*
TBE	tris-borate ethylenediaminetetraacetate
TFO	triplex-forming oligonucleotide
TG	triglyceride
TMO	trimethylated oligomeric chitosan
UV	ultraviolet
VLDL	very low density lipoprotein
XbaI	restriction endonuclease I isolated from *Xanthomonas badrii*
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
X-SCID	X-linked severe combined immunodeficiency
ZFN	zinc finger nuclease

CHAPTER 1: GENERAL INTRODUCTION

1.1 Cardiovascular disease and atherosclerosis

According to the British Heart Foundation,[1] in 2004, cardiovascular disease (CVD) accounted for nearly 40% of deaths in the UK, 10% more than all forms of cancer taken together. The World Health Organization estimates that 30% of all deaths worldwide are due to CVD.[2] Half of these deaths are caused by coronary heart disease (CHD) and a quarter by stroke.

Both have the same cause – a blockage in an artery cuts off the blood supply to parts of the heart muscle or brain, resulting in oxygen and nutrient deprivation that can lead to irreversible tissue damage. The most common reason for arterial blockage is the build-up of fatty streaks and their subsequent development into fibrous lesions on the inner walls of blood vessels in a process known as atherosclerosis. In a few cases the fibrous lesion itself grows large enough to block the blood flow in the artery, but usually it is the rupture or erosion of such a lesion that leads to the formation of an occluding blood clot.[3]

Atherosclerosis can be described as a chronic inflammatory response to injury of the endothelial cell layer coating the luminal side of the artery wall.[4] Lesion formation is initiated by decreased haemodynamic shear stress and increased turbulence in the blood flow caused by arterial branches and curvatures.[5] In these areas the endothelial cell layer displays greater permeability to macromolecules such as low density lipoprotein (LDL) and raised production of reactive oxygen species primarily by NADPH oxidases. A simultaneous drop in endothelial nitric oxide synthase reduces the amount of available nitric oxide radicals normally responsible for scavenging reactive oxygen species.[6]

Interaction between reactive oxygen species and LDL trapped in the intima by aggregate formation and proteoglycan binding leads to oxidation of phospholipids on the LDL. This minimally modified LDL (mmLDL) induces overlying endothelial cells to express chemotactic proteins and vascular cell adhesion molecules that recruit monocytes and T lymphocytes to the intima.[7] The monocytes mature into macrophages and proliferate as a result of their exposure to growth factors such as colony-stimulating factor secreted by the activated endothelial cells and differentiated macrophages themselves.

Myeloperoxidase, sphingomyelinase and secretory phospholipase further modify mmLDL to highly oxidized LDL (oxLDL) which is recognized by the macrophage scavenger receptor class A (SR-A) and CD36. Receptor-mediated internalization of oxLDL and subsequent formation of membrane-bound lipid droplets in the cytoplasm of macrophages triggers their differentiation into foam cells (Figure 1-1). CD40/CD40L cross talk between macrophages and T cells stimulates secretion of interferon-γ from the latter which up-regulates the expression of the SR-A and CD36 receptors on the surface of macrophages.

Autocrine activation by tumour necrosis factor α also plays a role in this positive feedback loop.[8] Ultimately the bloated foam cells succumb to programmed cell death and the resulting lipid-rich cell debris forms the beginning of a necrotic core. Smooth muscle cells emigrate from the media into the intima as a result of the cytokines and growth factors released by both T cells and macrophages. There they secrete extra-cellular matrix proteins and give rise to a fibrous cap covering the necrotic core.

As the lesion progresses from a fatty streak to a fibrous necrotic plaque, the blood flow in the artery is further disturbed. This leads to continuous recruitment and infiltration of monocytes and T cells at the lesion shoulders. The newly arrived macrophages produce proteases such as collagenases, gelatinases, stromolysin and cathepsins that digest the extra-cellular matrix proteins.[9] Ultimately, however, these macrophages also differentiate into foam cells and increase the lesion volume.

Lesion rupture frequently occurs at these foam cell rich shoulders, perhaps due to instability caused by a higher ratio of foam cells to smooth muscle cells in areas already suffering from a weakened extra-cellular matrix. Exposure of the necrotic core contents to the luminal blood promotes coagulation and thrombosis, which may lead to a local blockage. Alternatively, the blood clot might be swept away from the injury site and obstruct a smaller blood vessel downstream.

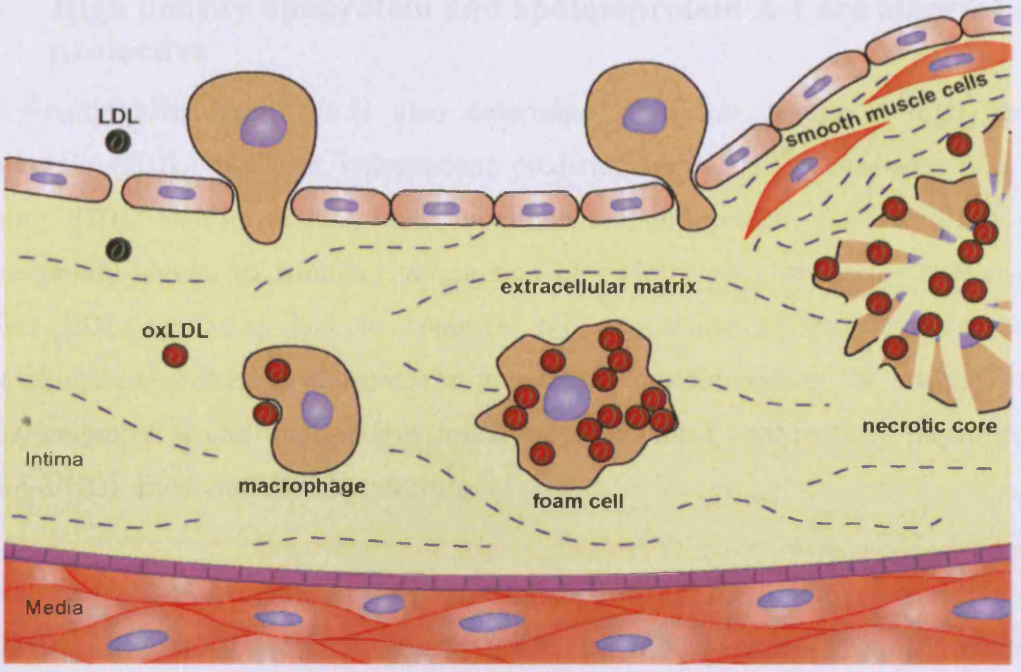


Figure 1-1 Early steps in atherosclerosis development

LDL particles can break through the endothelial cell layer coating the luminal side of the artery wall in injured areas. Once in the intima, LDL particles are modified into oxLDL which attract macrophages. These engulf oxLDL and differentiate into cholesterol-laden foam cells. Ultimately, foam cells succumb to lipid overload and die, forming the cholesterol-rich necrotic core of an advanced fatty streak. Smooth muscle cells emigrate from the media into the intima and secrete extra-cellular matrix proteins that give rise to a fibrous cap covering the necrotic core.

Even though atherosclerosis is a dynamic process, already evident in the unborn foetus,[7] not every arterial wall injury progresses to a complicated and potentially fatal necrotic plaque. However, in the Framingham Heart Study smoking, high LDL-cholesterol levels, physical inactivity, obesity, diabetes mellitus and hypertension were all linked to increased risk of coronary heart disease.

1.2 High density lipoprotein and apolipoprotein A-I are athero-protective

The Framingham Heart Study also determined that low levels of high density lipoprotein (HDL) were an independent predictor for CHD.[10] Consistent with this finding, HDL and its main protein component apolipoprotein A-I (ApoA-I) were subsequently shown to stimulate production of endothelial nitric oxide synthase, to inhibit LDL oxidation and to suppress the expression of monocyte attracting chemokines and adhesion molecules in the smooth muscle cells of the arterial wall.[11] Nevertheless, it is still unclear how much these functions contribute to physiological ApoA-I/HDL-mediated atheroprotection.

It is widely believed that the key anti-atherogenic role of HDL is the removal of excess cholesterol from cells and its transport back to the liver, a concept originally proposed by Glomset[12] and subsequently termed "reverse cholesterol transport" (RCT). A detailed explanation of RCT is provided in Figure 1-2 and Figure 1-3.

Based on the role of ApoA-I in HDL maturation and reverse cholesterol transport it is not surprising that deletion of the *APOA-I* gene results in undetectable levels of HDL in humans.[13] Moreover, very little ApoA-I or HDL is present in the plasma of patients with Tangier disease or familial HDL deficiency, genetic conditions caused by mutations in the *ABCA1* gene.[14] ApoA-I production is normal in Tangier disease patients, but it is catabolized extremely rapidly and thus very few mature HDL particles are formed.[15] Both *APOA-I* deletion and *ABCA1* mutations are strongly linked with early-onset cardiovascular disease.[13,16]

Loss of lecithin:cholesterol acyltransferase (LCAT) function as seen in familial LCAT deficiency and fish-eye disease also results in reduced ApoA-I/HDL levels. Recent data based on measurement of intima media thickness (IMT), a validated surrogate marker for atherosclerotic vascular disease, suggest a moderately increased CVD risk for carriers of *LCAT* mutations.[17] Indeed, IMT progression over family controls in carriers of *LCAT* mutations was 0.0055 mm/year compared with 0.0082 for *APOA-I* deletion or 0.0073 for *ABCA1* mutations.[18]

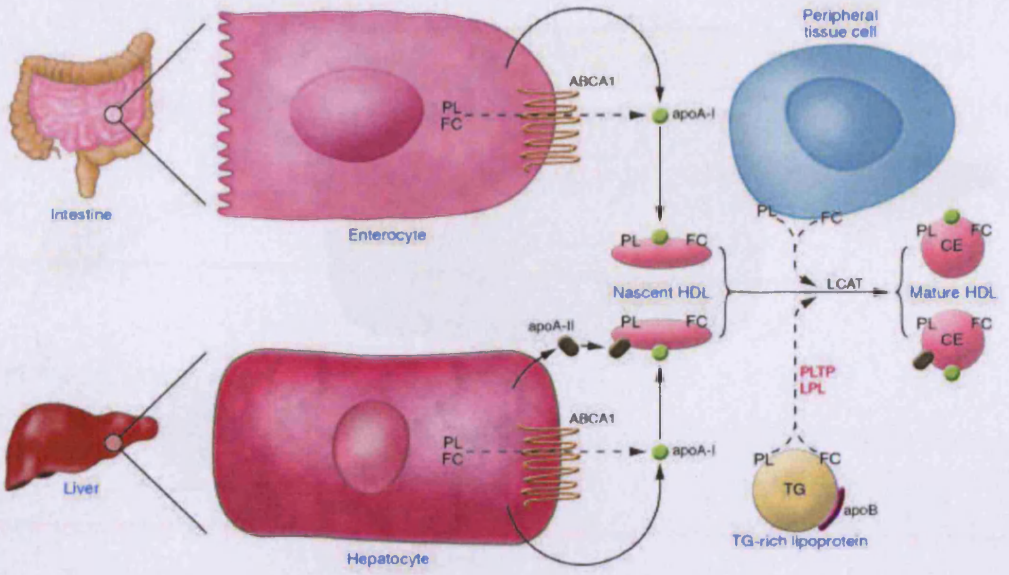


Figure 1-2 Reverse Cholesterol Transport – Maturation of HDL

Lipid-poor monomolecular ApoA-I is secreted from enterocytes and hepatocytes and is immediately lipidated via ATP-binding cassette transporter A1 (ABCA1)-mediated efflux of phospholipid (PL) and free cholesterol (FC). This generates disc-shaped HDL particles consisting of a PL bilayer in the centre bordered by anti-parallel ApoA-I dimers in a looped belt formation.[19] The HDL-associated LCAT is activated by the presence of discoidal HDL and catalyzes the transfer of 2-acyl groups from lecithin to free cholesterol on the surface of these particles.[12]

The cholesteryl esters (CE) created in this way are hydrophobic and consequently form an inner core, swelling the discoidal HDL initially into small spherical HDL_3 and then larger HDL_2 particles.[20] During this expansion the particles also acquire a third and fourth ApoA-I molecule. The HDL_3 and HDL_2 particles can then act as acceptors for further FC and PL efflux from peripheral tissues through the ATP-binding cassette transporters G1 and G4[21] or via gradient-dependent scavenger receptor B1 (SR-B1)-mediated FC exchange (not shown).

Transfer of CE and PL by cholesteryl ester transfer protein (CETP) and phospholipid transfer protein (PLTP) during lipoprotein lipase (LPL)-mediated lipolysis of the triglyceride (TG)-rich lipoproteins, chylomicrons and very low density lipoproteins (VLDL), also contributes to the CE and PL load of mature HDL particles. Figure taken from Rader.[22]

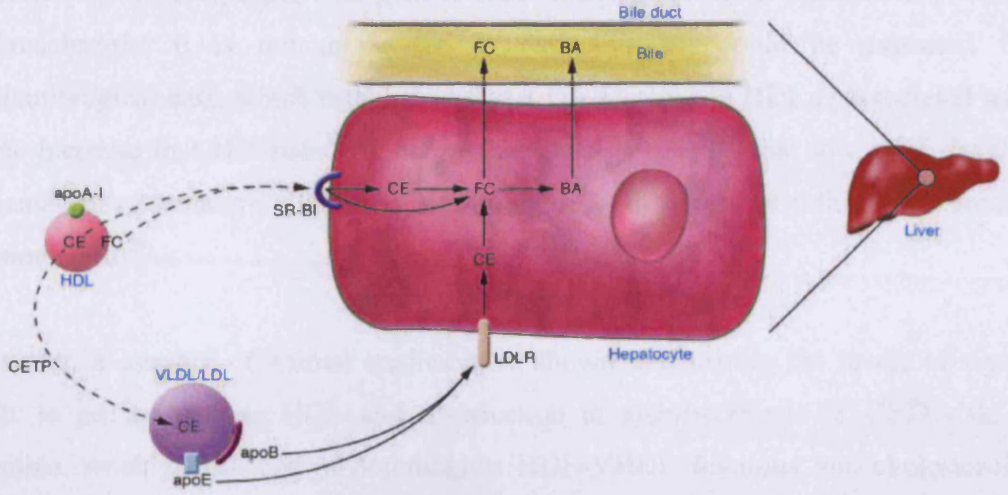


Figure 1-3 Reverse Cholesterol Transport – HDL catabolism and regeneration
Three possible mechanisms of HDL regeneration have been proposed.[23] In the first option, large spherical HDL particles travel to the liver where they are taken up selectively through ApoA-I binding to SR-B1.[24] Following removal of CE the depleted HDL is secreted and thus discoidal HDL particles are regenerated.

In an alternative pathway, HDL CE are exchanged for TG from ApoB-containing VLDL or LDL particles through the action of CETP. At least *in vitro* the efflux of CE to VLDL/LDL is higher than the influx of TG thus reducing the core lipid content and HDL particle size. The resulting excess of surface components leads to shedding of lipid-poor ApoA-I. The CE in VLDL/LDL can then be taken up by the liver via the LDL receptor (LDLR). In hepatocytes CE is hydrolyzed to free cholesterol which can either be directly excreted into the bile duct or is converted to bile acid (BA) before excretion.

Lastly, lipid-poor ApoA-I and discoidal HDL particles are side-products of the already mentioned remodelling of TG-rich chylomicrons and VLDL by LPL and PLTP (not shown). Regenerated lipid-poor ApoA-I may be lost from the HDL lifecycle through excretion by the kidneys or it may be relipidated and so retained in the plasma. Relipidation is achieved either by fusion with existing HDL particles during LCAT-mediated maturation or by absorption of CE and PL as described in Figure 1-2. Figure taken from Rader.[22]

Despite the fact that people with genetic HDL deficiencies show evidence of premature atherosclerosis, it is not as severe or extensive as would be expected from epidemiological data, which established that a 1% decrease in HDL is associated with a 1-2% increase in CHD risk.[25] Hence, it has been suggested that low HDL levels in patients with premature CHD may not be cause and effect but rather a concomitant phenomenon.[26]

However, a number of animal studies have shown that raising the levels of ApoA-I leads to an increase in HDL and a reduction in atherosclerosis or CHD risk. For example, weekly infusions of homologous HDL-VHDL fractions into cholesterol-fed rabbits markedly inhibited development of aortic fatty streaks and lipid deposition in the arterial walls without changing plasma lipid levels[27] and also reduced the extent of established aortic atherosclerosis by 50%.[28] Over-expression of human ApoA-I in transgenic rabbits[29] and mice[30,31] resulted in an increase in HDL concentration and protection against both diet-induced atherosclerosis[29,30] and spontaneous atherosclerosis in the ApoE-deficient mouse.[31]

At the same time, several clinical trials were using fibrates such as gemfibrozil or bezafibrate to increase HDL levels in humans. Both the Helsinki Heart Study[32] and the Veterans Affairs High density lipoprotein cholesterol Intervention Trial (VA-HIT)[33] demonstrated some clinical benefit that was dependent on elevated HDL levels using gemfibrozil. Even though results from the bezafibrate infarction prevention study[34] were not significant despite an 18% increase in HDL concentration, a post-hoc analysis of patients with TG levels ≥200 mg/dL suggested a substantial benefit in this subgroup.

Since fibrates are amphipathic carboxylic acids that act as peroxisome proliferator-activated receptor α (PPARα) agonists, they boost the synthesis of ApoA-I, but also ApoA-II, LPL, ABCA1 and SR-B1,[35] they reduce LDL and especially triglyceride levels in addition to the modest increase in HDL concentration. Thus, it was difficult to determine which changes in the lipid profile were responsible for the clinical benefits seen and the question of whether low HDL levels are cause and effect of CHD risk, or a concomitant phenomenon, could not be answered.

Another class of drugs intended to raise HDL levels in patients are CETP inhibitors such as torcetrapib, anacetrapib and R1658 (also known as JTT-705). These drugs were developed after it was noted that genetic CETP deficiency, a defect common in the Japanese population, is associated with markedly increased levels of HDL and accumulation of HDL_2 particles.[36] This change in lipid profile was attributed to decreased removal of HDL CE and slower catabolism of ApoA-I.[37]

Inconsistent epidemiological evidence hinted that CETP deficiency may protect against atherosclerosis, provided that it is coupled with elevated HDL levels and that the HDL particles are functional.[38] Results from studies in mice were also confusing, but given that mice normally lack CETP, this was not surprising. Studies in cholesterol-fed rabbits on the other hand were unequivocally positive, with a more than threefold increase in HDL levels and a 60% reduction in atherosclerosis.[39] In addition, early clinical trials using torcetrapib were extremely encouraging, with significant increases in HDL (up to 100% at the highest dose) and concurrent but smaller decreases in LDL levels.[40,41]

Therefore, both surrogate marker and end-point studies were commenced at the same time. Due to marketing considerations all studies combined torcetrapib with atorvastatin. Statins are widely used for the reduction of high LDL levels as they inhibit 3-hydroxy-3-methyglutaryl-coenzyme A reductase, an enzyme that is responsible for the biosynthesis of cholesterol. Unfortunately, the "Investigation of Lipid Level management to Understand its iMpact IN ATherosclerotic Events" (ILLUMINATE) trial had to be terminated early because of a 61% excess of deaths from cardiovascular and noncardiovascular causes in the torcetrapib/atorvastatin over the atorvastatin only group.[42] Increased risk of death was associated with changes in serum electrolyte and aldosterone levels as well as increased blood pressure.[43]

Three other torcetrapib studies using intravascular ultrasound and carotid artery IMT as surrogate markers also disappointed as there was no treatment effect on atheroma volume or intimal media thickness. All studies showed a 50-60% increase in HDL and a 20% reduction in LDL levels.[43-46] It is not yet clear if torcetrapib failed due to side-effects caused by the drug itself or by combination with atorvastatin, or if the mechanism of raising HDL by CETP inhibition is to blame. However, data from clinical trials with anacetrapib and R1658 showed similar modulation of HDL and LDL levels without significant blood pressure increases.[47-49]

What seems clear is that increasing HDL levels by any means does not necessarily correlate with protection against atherosclerosis. Apparently, our understanding of the intricate remodelling processes involved in the HDL lifecycle is still too incomplete for successful manipulation. Perhaps more direct methods need to be considered first.

One such direct method of raising HDL is based on intravenous infusion of reconstituted HDL-like particles (rHDL) made from recombinant ApoA-I or purified plasma ApoA-I and soybean phosphatidylcholine (PC). In the earliest studies Carlson[50] showed that a single injection of rHDL, corresponding to about 1.6 g of recombinant ApoA-I, improved HDL concentration by 20% in four patients with low plasma HDL. The same treatment transiently increased ApoA-I and HDL levels in four patients with heterozygous familial hypercholesterolaemia and was associated with excess cholesterol excretion of around 500 mg/d for at least nine days.[51]

Nanjee and colleagues also demonstrated that intravenous infusion of ApoA-I/PC discs increased small HDL particle concentration and stimulated RCT in healthy male volunteers,[52,53] whereas lipid-free ApoA-I did not.[54] Endothelial function in hyper-cholesterolaemic men[55] and familial hypoalphalipoproteinaemia patients[56] was restored by a single infusion of rHDL corresponding to 80 mg/kg bodyweight of plasma derived ApoA-I. No adverse side effects were reported in any of the studies.

These early clinical studies provided enough evidence to encourage a larger, randomized controlled trial - the "Effect of rHDL on Atherosclerosis - Safety and Efficacy" (ERASE) study.[57] In this trial patients were randomly assigned to receive placebo (saline), 40 mg/kg or 80 mg/kg rHDL, but the higher dose group was discontinued early due to abnormalities in liver function tests. Although there was a 3.4% or 5.3 mm^3 reduction in total atheroma volume versus patient baseline the changes were not significant compared to the control group.[57] However, both plaque characterization index on intravascular ultrasound and changes in coronary score on quantitative coronary angiography were substantially different between groups.[57]

Considering that the ERASE differences were potentially confounded because 90% of enrolled patients, including those in the control group, were also taking LDL-lowering medication, this outcome was very encouraging.

1.3 Apolipoprotein A-I_{Milano}, a natural variant of ApoA-I, is super-atheroprotective

A similar clinical trial,[58] where patients were injected with 15 mg/kg or 45 mg/kg of recombinant apolipoprotein A-I_{Milano} (ApoA-I_M) and 1-palmitoyl-2-oleoyl phosphate-idylcholine (POPC) complexes (ETC-216) once per week for five weeks, demonstrated a 4.2% decrease or 14.1 mm^3 reduction in total atheroma volume that was statistically significant versus patient baseline, but not between treatment and control groups. In *in vivo* experiments, ETC-216 induced plaque regression in rabbits by 5-6%[59,60] and also in mice, where it was attributed to a reduction in plaque lipid content.[61,62]

ApoA-I_M is a naturally occurring variant of ApoA-I believed to be super-atheroprotective. It was originally described in a man presenting with a high-risk lipid profile, specifically hypertriglyceridaemia and markedly decreased HDL, but no sign of atherosclerosis.[63] Further investigation revealed that this protein contained an arginine to cysteine substitution at position 173, allowing formation of homodimers as well as heterodimers with ApoA-II via a disulphide bridge.[64,65]

The patient originally hailed from the small northern Italian village Limone sul Garda, an isolated community with a distinct genetic and clinical conformity due to high consanguinity. Sampling of the entire population of this village found 32 additional carriers, who all were descendants of a couple living in the second half of the eighteenth century.[66] Inheritance was autosomal dominant. All carriers were heterozygotes and displayed the characteristic high-risk lipid profile of moderate hypertriglyceridaemia combined with very low HDL levels.[66,67]

When the IMTs of ApoA-I_M carriers were compared with control subjects from the same kindred who did not have the mutation, there was no difference. Clinic patients and blood donors, both with hypoalphalipoproteinaemia, on the other hand had maximum IMT measurements up to twice as large.[67] Prevalence of atherosclerotic plaques too was significantly higher in these two groups than in ApoA-I_M carriers or kindred controls.[67] Relative to controls ApoA-I_M carriers displayed a 17% higher cardiac output, possibly suggesting improved ventricular distensibility.[67]

It is clear from these data that despite rather low HDL levels ApoA-I_M carriers do not have an increased risk of CVD, quite unlike patients with genetic ApoA-I deficiency or Tangier disease. How then does the ApoA-I_M mutation mediate this increased atheroprotection?

Several explanations for this increased atheroprotection have been investigated in the literature. For example, studies have shown that the increased atheroprotection of ApoA-I_M carriers is due to increased cholesterol efflux to the mutant protein.[68-70] Other studies containing ostensibly contradicting data[71-75] were also published, but careful reading and analysis allows two conclusions to be drawn. Firstly, ApoA-I and ApoA-I_M exhibit similar cholesterol efflux ability in lipid-free and rHDL form[70,72,73] and, secondly, ApoA-I_M containing serum from human carriers or transgenic mice[68,70,72] is significantly better than ApoA-I serum at both SR-B1- and ABCA1-mediated cholesterol efflux. Favari *et al.*[70] attribute this finding to the presence of a unique protease-sensitive small HDL particle containing an ApoA-I_M dimer in addition to normal amounts of lipid-poor ApoA-I.

Another potential reason for the apparent atheroprotection of ApoA-I_M carriers is increased metabolism of the mutant protein and thus more rapid RCT. Two papers[76,77] have investigated this possibility and agree in their findings. ApoA-I_M was almost equally distributed between monomers, homodimers and heterodimers with ApoA-II in human carriers and all forms were more rapidly metabolized than ApoA-I monomers. Even though production rate was normal, wild-type protein constituted only about 30% of the total ApoA-I protein level. Hence wild-type protein too was metabolized faster in ApoA-I_M carriers than in control subjects, perhaps because ApoA-I and ApoA-I_M molecules often co-localize in the same particle. One explanation for increased HDL catabolism may be the reduced reactivity of ApoA-I_M with LCAT,[78] which would inhibit HDL maturation.

Higher anti-oxidant and anti-inflammatory capabilities[79] leading to decreased endothelial dysfunction have also been proposed to explain enhanced ApoA-I_M atheroprotection, but this has been disputed.[75] Nevertheless, a recent paper[80] has shown that ApoA-I_M HDL isolated from human carriers stimulated higher expression and activation of endothelial nitric oxide synthase than HDL from kindred controls and was also more effective in down-regulating vascular cell adhesion molecule-1.

Whatever the mode of action, the evidence supports the idea that ApoA-I_M has beneficial effects on the outcome of CVD above that of ApoA-I. However, a recombinant protein therapy requiring tens of grams per treatment is hardly practical in terms of cost and ease of manufacture for widespread chronic treatment of cardiovascular disease. Gene therapy would be an ideal alternative.

1.4 Gene Therapy

Gene therapy was first proposed nearly thirty years ago as a method of manipulating cells at the molecular level in order to cure monogenetic diseases like cystic fibrosis, haemophilia and phenylketonuria.[81] Since then gene therapy has moved into treatment of acquired disorders such as cancer and cardiovascular disease with 66.5% and 9.1% of gene therapy clinical trials to date focused on these areas.[82] Other targets include infectious disease (6.5% of trials), particularly human immunodeficiency virus infection, neurological diseases such as multiple sclerosis, Alzheimer's and Parkinson's (1.5%) and ocular diseases including age-related macular degeneration (0.9%).[82]

Gene therapy can be applied to either somatic or germ line cells, but even though germ line gene therapy has the advantage of being inherited by future generations, it is currently prohibited in human beings due to ethical issues. Somatic gene therapy, on the other hand, avoids the ethical implications of germ line alterations, but has to be applied anew in each generation. *Ex vivo* gene therapy manipulates cells outside the body and then transplants the treated cells into the body, whereas *in vivo* gene therapy treats cells *in situ*.

Gene therapy can be divided into three broad classes: gene addition, gene targeting and gene modulation. Gene addition inserts extra functional copies of the target gene into affected cells to complement the loss of functionality caused by a defective endogenous gene, whereas gene targeting replaces the non-functional endogenous gene with a functional copy. A subclass of gene targeting, gene editing, aims to repair the loss-of-function mutation *in situ* and to return the non-functional endogenous gene to normal function in this way. Gene modulation alters the expression level of the target gene or changes splicing patterns and generally works at the transcriptional or translational level.

Up to now, the overwhelming majority of gene therapy clinical trials have used gene addition strategies,[82] although several gene modulation studies are underway and one gene modulation treatment has been approved.[83] Currently, gene targeting is widely used in the production of transgenic mouse lines, but no clinical trials employing this method are underway. Gene editing is still in the basic research phase, though some data from *in vivo* experiments are available.

Each type of gene therapy offers certain advantages, but also suffers from specific limitations. Thus, gene therapy methods have to be carefully chosen depending on the intended application. In the following sections, each type of gene therapy including respective advantages and limitations will be discussed in more detail.

1.4.1 Gene addition

As stated in the section above, gene addition aims to introduce exogenous functional copies in addition to the non-functional endogenous gene. This can be achieved by either viral or non-viral vectors. Two thirds of clinical trials to date have used viral vectors,[82] in particular adenoviral and retroviral vectors, in an attempt to exploit the high efficiency of viral transduction and relative specificity of viral tissue targeting.

However, the efficiency of viral vectors may be inhibited by pre-existing immunity to the virus and viral proteins induce a cellular immune response against the transduced cells. Viral vectors may also have limits in transgene capacity and are expensive to produce. Non-viral vectors on the other hand, can be produced cost-effectively, induce limited host immune responses, have a larger capacity, but are very inefficiently delivered and thus require a delivery system. Both viral and non-viral gene addition strategies have encountered problems with sustaining transgene expression over longer periods of time. If integrating viral or non-viral vectors are used, insertional mutagenesis may occur.

A problem with all gene addition strategies is loss of transgene expression due to epigenetic gene silencing or the development of host immune responses against the therapeutic gene product. Epigenetic silencing is mediated via *de novo* methylation of cytosines at CpG dinucleotides in the DNA sequence and may be induced by multiple copy transgene arrays,[84] retro- or lentiviral sequences[85] or bacterial sequences.[86,87] Solutions to the problem of epigenetic silencing include the use of tissue-specific or synthetic promoters instead of viral promoters,[88] optimizing bacterial DNA sequences for human codon usage[89] or excision of the transgene from bacterial sequences after gene transfer.[90] Anti-transgene immune responses may be modulated by exploiting the microRNA (miRNA) regulatory pathways of cells.[91]

1.4.1.1 Viral vectors

Viral vectors have been employed for gene therapy because viruses have evolved highly efficient mechanisms for infection of their target cells and have a clearly defined tissue tropism. Generally, viruses are made replication incompetent by replacing specific viral genes with the transgene cassette containing the gene of interest and a promoter.[92] Additional viral genes are removed to reduce the host cellular immune response, while genes that are required for *in cis* functions, such as viral capsid packaging or DNA integration, are left intact. During production of recombinant vector particles, the deleted viral genes are provided *in trans* by a separate helper construct which is either co-transfected with the vector genome or provided by packaging cells.[93]

Currently, there are five main groups of viral vectors based on retroviruses, lentiviruses, adenoviruses, adeno-associated viruses (AAV) and herpes simplex virus-1 (HSV-1).[92] They all have varying properties (see Table 1-1 for a comparison) and are, therefore, suitable for different applications. One major distinction between the different groups is the ability to either integrate into the host genome or persist in the cell nucleus as extra-chromosomal episomes.

Vectors based on oncoretroviruses, for example Moloney murine leukaemia virus, were the first to be developed.[92] Oncoretrovirus vectors permit stable transgene expression in dividing cells because they integrate into the host genome, but integration carries the risk of insertional mutagenesis. The extent of this risk was demonstrated when 4 out of 20 children with X-linked severe immunodeficiency treated with such a vector[94,95] developed a leukaemia-like disease[96] because the vector had integrated near and activated an oncogene.[97]

In contrast to oncoretroviruses, lentiviral vectors derived from human immuno-deficiency virus (HIV) or non-human lentiviruses can transduce non-dividing cells[98,99] as they are actively transported through nuclear pores.[100] Self-inactivating lentiviral vectors are considered particularly safe because a deletion in the downstream long-terminal repeat sequence has minimized the risk of creating replication-competent retrovirus.[92] So far, lentiviral vectors have been used in a single phase I clinical trial where a 937 bp long antisense sequence to the HIV *env* protein was expressed.[101]

Table 1-1 The five main groups of viral vectors

Comparison of the main features of the most commonly used viral vectors for gene therapy. Table taken from Thomas *et al.*[92]

Herpes simplex virus-1 is a human pathogen and particularly interesting as a viral vector because it can persist in a latent state as an episome for extended periods of time and its natural tropism is for neurons.[102] It also has a packaging capacity of up to 40 kb and simultaneous expression of multiple transgenes is possible.[103] However, there is a trade-off between long-term high-level transgene expression and cytotoxicity.[104] Conversely, cytotoxicity is advantageous in the treatment of cancer. For example, a replication-defective HSV vector (OncoVEX^{GM-CSF}) has shown positive results in clinical trials for metastatic melanoma.[105,106]

Recombinant adenoviral vectors are the most efficient gene transfer system in a broad range of tissues, infect both dividing and non-dividing cells and are retained as extra-chromosomal episomes. The main draw-back of these vectors is their potent immuno-genicity which caused the death of Jesse Gelsinger during a clinical trial for the treatment of ornithine transcarbamylase in 1999.[107] To reduce vector-related immuno-genicity, progressively more viral genes were deleted[108] and a helper-dependent adenoviral vector, with all viral genes deleted, was developed.[109]

Adeno-associated viral vectors, on the other hand, have little immunogenicity and the host immune response is limited to the production of neutralizing antibodies. However, AAV vectors have the smallest packaging capacity and although they generally persist as episomal DNA in human cells, random integration and insertional mutagenesis have been reported in some studies using liver-directed AAV2.[110,111]

Currently, eleven different serotypes (designated AAV1 to AAV11) with diverse tissue tropism have been identified.[112] Pseudotyping with receptor-binding proteins from other strains has become common practice partly to avoid wide-spread pre-existing humoral immunity to wild-type AAV2 in the human population.[113] Although relative late-comers, AAV vectors have demonstrated their safety in a number of clinical trials for diverse diseases such as haemophilia B, cystic fibrosis and Parkinson's.[114]

1.4.1.2 Non-viral vectors

Non-viral vectors have long been discounted as a viable alternative to viral vectors, mainly because of low transfection efficiency and transient transgene expression. On the plus side, non-viral vectors are rather cheap to produce even in large quantities, their immunogenicity is minimal and recent advances have improved transduction efficiency and persistence of transgene expression. Historically, non-viral gene therapy has concentrated on the use of naked DNA,[82] usually plasmids, though polymerase chain reaction (PCR) products have also been used.

However, unlike viral vectors, naked DNA is not protected from the extra- and intracellular environments and does not have a specific tissue tropism or means of cell entry. Thus, transfection efficiency is lower than with viral vectors, but specialized gene delivery methods can be used to improve transduction. These methods will be discussed separately in the next section (section 1.4.2).

Transgene expression from transduced DNA is short-lived, mostly because plasmids are retained in the nucleus in an extra-chromosomal form and thus lost during the breakdown of the nuclear envelope at mitosis. Two different approaches aim to address this problem. In the first approach, a plasmid containing an SV40-ori sequence and the scaffold/matrix attachment region (S/MAR) from the human β-interferon gene cluster[115] is tethered to the nuclear matrix.[116] This ensures replication and episomal maintenance as well as preventing epigenetic silencing of the cytomegalovirus immediate-early (CMV) promoter.[117]

In the second approach, loss of the vector DNA is prevented by integration into the host genome. This approach mainly uses transposons, mobile genetic elements that can re-locate within the genome via a cut-and-paste mechanism mediated by an encoded transposase. The transposase gene can be replaced by up to 10 kb of exogenous DNA and the transposase is then provided either *in cis* on the same plasmid or *in trans* as protein or additional plasmid.[118]

Long-term transgene expression has been reported in mice[119,120] and embryonic stem cells[121] after integration mediated by the Sleeping Beauty transposon. This transposon was reconstituted from an inactive Tc1-like transposon recovered from fish[122] and is currently the most active transposon available.[120] As with viral vectors, transposon-mediated integration carries the risk of insertional mutagenesis, but this could potentially be avoided by using a site-specific transposase that directs integration into defined DNA sequences.[123]

1.4.2 Gene delivery

Cellular uptake of naked DNA is extremely inefficient due to electrostatic repulsion between the negatively charged cell surface and phosphate backbone of DNA. Other barriers to efficient DNA delivery are rapid hydrolysis of the phosphodiester links in the DNA backbone by endo- and exonucleases and the need for DNA entering the cell via endosomal uptake to escape from these vesicles (Figure 1-4). Also, transfer of the therapeutic DNA into the nucleus is poor, but required for therapeutic effect in most cases.[124]

To alleviate these problems, physical and chemical gene delivery methods have been developed. Physical methods, including mechanical and electrical techniques, require direct exposure of the target tissue and as such are largely limited to topical applications *in vivo*. A wide variety of chemical methods are available which can be roughly divided into cationic polymers and liposomes.[124]

In general, the chemical delivery reagents are highly efficient *in vitro*, but there seems to be an association between transfection efficiency and cytotoxicity, suggesting that a certain level of cell damage is necessary for successful transduction. However, *in vivo* transfection efficiency using chemical gene delivery is poor compared to viral vectors, because systemic delivery of cationic polymers and liposomes lacks target cell specificity. Other barriers to efficient *in vivo* transfection include rapid clearance from the blood circulation, coating by plasma proteins and subsequent phagocytosis by macrophages, and difficulties with crossing vascular endothelial barriers.[125]

Figure 1-4 Barriers to efficient gene delivery

Gene delivery vehicles must protect the therapeutic DNA against the extra-cellular environment, bind to the target cells, be internalized efficiently (by endocytosis in the case of cationic polymers or membrane fusion for some liposomes), escape the degradative pathway and facilitate unpackaging and nuclear delivery of the DNA. Figure taken from Pack *et al.*[126]

1.4.2.1 Physical methods for gene delivery

Physical gene delivery techniques include particle bombardment, electroporation, ultrasound, microinjection and hydrodynamic injection.[127] Bombardment with DNA-coated gold particles, also known as "gene gun", ensures direct penetration of the cell and nuclear membranes avoiding the endosomal pathway altogether. However, this method requires direct physical access to the target tissue and only mediates shallow tissue penetration. Therefore, the gene gun has mainly been used in DNA vaccination and immunotherapy.[128]

Electroporation uses a combination of short high-voltage pulses followed by a series of longer, low-voltage pulses. The high-voltage pulses create transient hydrophilic pores in the cell membrane of target cells through which the DNA can enter, assisted by the electrophoretic effect of the low-voltage pulses.

Electroporation was originally developed for *in vitro* transfection[129] and is routinely used for DNA transfer into cultured cells, but has now found application *in vivo.*[130] Skeletal muscle is the preferred target tissue due to its ready access and ability to sustain long-tem expression of the episomal plasmid, but the electroporation effect is limited to the area between the electrodes.

Ultrasound is less invasive than the gene gun or electroporation and has already demonstrated its safety during clinical application for diagnostic imaging and targeted therapeutic purposes. However, it appears to be less effective than electroporation.[130] Microinjection, on the other hand, is extremely effective as it directly injects the DNA into the cell nucleus, but is very labour-intensive because cells have to be treated one at a time. Thus, it has been used mainly in the creation of transgenic mice[131] but recent progress in automating the process[132] should encourage more wide-spread application.

Rapid, high volume injection of DNA, also known as hydrodynamic injection, into mice tail veins causes transient hyper-permeability of hepatocytes cell membranes due to the rise in venous pressure and subsequent enlargement of the fenestrae in the liver sinusoids.[133] Although not directly applicable to humans, hydrodynamic injection could be used in conjunction with isolated limb perfusion to enhance limb musculature transduction[134] or the isolated segment of inferior vena cava approach for total liver transfection.[135]

1.4.2.2 Chemical facilitation of gene delivery

One of the cheapest *in vitro* methods for DNA transfection is calcium phosphate precipitation, but multiple parameters such as DNA, calcium and phosphate concentrations, temperature, pH and reaction time affect the formation of the DNA-calcium phosphate co-precipitate and thus the reproducibility and efficiency of transfection.[136] Thus, calcium phosphate precipitation is mainly used in repeat or large-scale transfections with little inter-experimental variation, such as viral vector production.

Cationic polymers, on the other hand, easily complex with the anionic DNA due to electrostatic interactions and form positively charged polyplexes that can then interact with the negatively charged cell surface.[124] Polyethylenimine (PEI), chitosan and atelocollagen are examples of such cationic polymers. In PEI every third atom is an amino nitrogen that can bind a DNA phosphate and PEI is therefore highly efficient at compacting DNA.[137] The partial nitrogen protonation at physiological pH acts as a "proton sponge" that causes osmotic swelling and subsequent rupture of endosomes, thus efficiently releasing the polyplexes into the cytoplasm.[138] Linear (LPEI; commercially available as jetPEI) and branched forms of PEI are available. Also, efforts to improve transfection efficiency and reduce toxicity have resulted in PEI modifications such as pegylated PEI[139] or fully deacylated LPEI.[140]

Chitosan, a natural linear amino-polysaccharide, is bio-compatible, bio-degradable and non-toxic.[141] Like PEI, chitosan and its derivatives form positively charged polyplexes with DNA that can interact with the cell membrane and mediate cellular uptake. A chitosan derivative, trimethylated oligomeric chitosan (TMO), demonstrated very low cytotoxicity compared to lipoplexes and improved transfection efficiency in some cell lines.[142] Particularly interesting is the possibility of using TMO with antennary galactose residues for specific targeting to HepG2 cells.[143]

Atelocollagen is another bio-compatible and bio-degradable polymer with low toxicity.[144] It is derived from collagen which has been used extensively for medical and cosmetic applications. In atelocollagen, the telopeptides that are the main antigenic sites of collagen are removed, giving it an even better safety profile.[144] Remarkably, atelocollagen mediated high levels of oligonucleotide-mediated gene editing in HepG2 cells (up to 11%),[145] although this efficiency has yet to be duplicated by other researchers.

Another class of chemical delivery agents are cationic liposomes. Amphipathic phospholipids spontaneously form lipid bilayers in aqueous solutions and encapsulate any DNA present in the solution due to electrostatic interactions. Generally, cationic liposomes contain a mixture of cationic and zwitterionic lipids, with the former mediating DNA condensation and cellular association while the latter perturb the cell membrane and thus help with membrane fusion.

Commonly used cationic lipids include 1,2-dioleoyl-3-trimethylammonium propane (DOTAP) and N-[1-(2,3-dioleyloxy) propyl]-N,N,N-trimethylammonium chloride (DOTMA), while 1,2-dioleoyl-phosphatidylethanolamine (DOPE) is an example of a zwitterionic lipid.[124] A large variety of cationic lipids with proprietary formulations are commercially available, e.g. Lipofectamine 2000,[146] FuGene and Trojene.[147] Trojene is a mixture of N-1-cholesteryloxycarbonyl-3,7-diazanononane-1,9-diamine (CDAN) with DOPE and has been shown to be significantly less cytotoxic than Lipofectamine 2000.[148] Lipofection is one of the most commonly used *in vitro* transfection methods, but has also been used in 7.6% of clinical trials to date.[82]

1.4.3 Gene modulation

Unlike gene addition or gene targeting, gene modulation does not interact directly with the target gene, but exerts indirect influence on the expression of an endogenous gene at the transcriptional or translational level. As such it is mostly limited to the treatment of abnormal endogenous gene expression or prevention of gene expression from exogenous viral DNA with the exception of correcting abnormal pre-RNA splicing.[149]

Gene modulation is mostly achieved by short RNA or DNA segments. Since RNA and DNA sequences are subject to endo- and exonucleases in the extra- and intracellular environment, various chemical backbone modifications have been developed to improve stability. Typically used backbone modifications for RNA and DNA oligonucleotides include phosphorothioate (PTO), locked nucleic acid (LNA) and morpholino backbone modifications (Figure 1-5), as well as the use of 2'-modified pyrimidines.

Phosphorothioate modified nucleotides contain a sulphur instead of an oxygen anion as one of the non-bridging ligands in the linking phosphate groups.[150] In LNA nucleotides the ribose is locked in the 3'-endo structural conformation by an extra bridge connecting the 2' and 4' carbons,[151] whereas in morpholino oligonucleotides the riboses have been replaced by morpholine rings that are linked through phosphorodiamidate groups instead of phosphates.[152] Pyrimidines can be modified at the 2' position with fluoro, amino or methyl groups.[153]

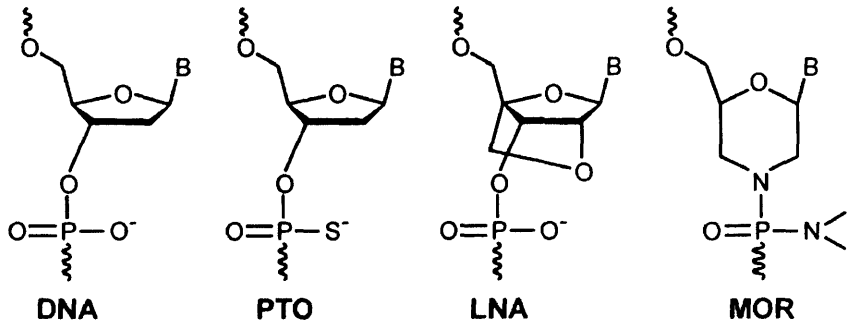


Figure 1-5 Chemical structure of oligonucleotide backbone modifications
In phosphorthioate (PTO) modifications, sulphur replaces an oxygen anion in the linking phosphate group. Locked nucleic acids (LNA) contain bridges connecting the 2' and 4' carbons of the 2-deoxyribose that lock the ribose into the 3'-endo structural conformation. In morpholino (MOR) oligonucleotides the riboses have been replaced by morpholine rings that are linked through phosphorodiamidate groups instead of phosphates.

Small interfering RNA (siRNA), short double-stranded RNA sequences that are complementary to the mRNA of the target protein, can induce degradation of the mRNA as well as suppressing protein synthesis and transcription of the target gene by appropriating the cellular RNA interference (RNAi) machinery (Figure 1-6). It has been suggested that RNAi is an ancient genome immune system against exogenous nucleic acids[154] and this could explain some of the difficulties with sustaining transgene expression. SiRNA is probably the best-known and most widely applied gene modulation method *in vitro*, because of its easy and convenient application. However, off-target effects can mediate down-regulation of sequence-related genes.[155] Nevertheless, several siRNA therapeutics are in clinical trials.[155]

Antisense oligonucleotides that bind to regions of a messenger RNA (mRNA) due to sequence complementarity can interfere with ribosomal initiation and thus prevent translation.[124] Successful examples of this mode of action include fomivirsen sodium (Vitravene), an inhibitor of the immediate early region 2 of human CMV, which has been approved for treatment of CMV retinitis in acquired immune deficiency syndrome patients[156] and AP 12009 which targets the mRNA of human transforming growth factor β2 and has shown significant benefits in patients with high grade gliomas.[157]

Figure 1-6 The RNAi machinery

The Dicer and RISC complexes are central to the protection against invading viral double-stranded RNA (①), the elimination of transcripts from transposons (②), the blockage of protein synthesis mediated by cellular miRNAs (③), and the RNAi-mediated suppression of transcription (④). Exogenous siRNA (⑤) can usurp the RNAi machinery to inhibit the activity of specific genes. Figure taken from Daneholt.[158]

Antisense oligonucleotides may also interfere with pre-mRNA processing by preventing the binding of small nuclear ribonucleoproteins at the intron borders. This process has been exploited to correct or even induce aberrant mRNA splicing in order to restore the correct reading frame or exclude exons containing mutations. A clinical trial using a morpholino oligonucleotide to skip exon 51 of the human dystrophin gene with the purpose of treating Duchenne muscular dystrophy (AVI-4658) has been initiated.[159]

Aptamers are another therapeutic application of oligonucleotides. They are nucleic acid sequences that have been selected over multiple rounds to bind specifically to targets such as proteins, nucleic acids or small molecules.[153] Unlike antibodies, aptamers can be produced by chemical synthesis, are more sensitive to conformational changes[160] and may bind to sites that are sterically inaccessible to antibodies.[161] The first therapeutic aptamer approved for clinical use is pegaptanib (brand name Macugen), a vascular endothelial growth factor inhibitor for treatment of age-related macular degeneration.[83]

Ribozymes are catalytic RNA molecules that are capable of sequence-specific cleavage of mRNA. Several RNAzymes such as HERzyme which targets human epithelial growth factor 2 or ANGIOzyme, a vascular endothelial growth factor receptor 1 inhibitor, are in clinical development.[162]

1.4.4 Gene targeting

Ultimately, however, gene therapy aspires to permanently repair the faulty gene *in situ* without the need for additional sequences or genome modifications. The gene would then remain under control of its endogenous regulatory elements and the natural expression pattern would be unchanged. Thus, gene expression in inappropriate cells and subsequent activation of host immune responses, as seen with gene addition strategies,[91] would be avoided.

The term gene targeting usually refers to the targeted exchange between an exogenous donor and a homologous endogenous acceptor DNA sequence which is mediated by homologous recombination (HR; see section 1.4.4.1). However, conventional gene targeting suffers from low efficiency, as typically only one cell in 10^5-10^7 treated cells is successfully altered[163] and unspecific integration of the donor DNA via non-homologous end joining (NHEJ; see section 1.4.4.1) occurs in one per 10^2-10^4 cells.[163] Thus, additional non-homologous sequences providing the necessary selection markers must be included in the donor DNA and this may result in epigenetic down-regulation.

To address these problems, a number of alternative gene targeting strategies have been developed (Figure 1-7). These are small fragment homologous replacement (SFHR), AAV-mediated gene targeting, and stimulation of HR by zinc finger nucleases (ZFN). Other approaches use triplex-forming oligonucleotides (TFO), chimeric RNA/DNA oligonucleotides (chimeraplasts) or single-stranded oligonucleotides (ssODN). At present it is unclear if the latter three approaches are mediated by HR or other cellular DNA repair pathways. Also, they aim to repair or edit the disease-associated mutation(s) only and do not replace extended parts of the endogenous gene sequence. Thus, these approaches are considered a subclass of gene targeting and herein are termed 'gene editing'; they will be discussed in section 1.4.5.

Figure 1-7 Overview of gene targeting strategies

The size of the homologous (blue) and non-homologous (open blue) donor DNA as well as the endogenous DNA (red) sequences for each strategy are indicated by the size bars. During gene targeting, the genetic information of the donor DNA is transferred to the endogenous DNA. Binding of triplex-forming oligonucleotide (TFO; orange) to the target DNA stimulates recombination whereas zinc finger nucleases (ZFN; green) create sequence-specific double-strand breaks (DSB) to increase homologous recombination (HR). Figure taken from Olsen.[164]

1.4.4.1 DNA damage response pathways relevant for gene editing

A short explanation of HR, NHEJ and other relevant DNA damage response pathways is provided here to aid understanding of the issues involved in therapeutic gene targeting. Mammalian cells are constantly exposed to a wide range of DNA damaging external and internal sources, such as reactive oxygen species or certain types of radiation. To maintain genome stability, cells have evolved a variety of DNA damage response pathways; among them mismatch repair (MMR), nucleotide excision repair (NER) and the double-strand break repair pathways of HR and NHEJ.

Base mismatches and small insertion/deletion loops in the DNA are removed by the MMR system. In mammalian cells, various heterodimers of MSH2, MSH3 and MSH6 recognize specific mismatch or loop lesions.[165] Thus, the MutSα complex (MSH2-MSH6) is responsible for detection of base-base mismatches and small loops, while MutSβ (MSH2-MSH3) identifies larger loops. The subsequent steps remain to be elucidated, but the MutLα complex (MLH1 and post-meiotic segregation protein 2), exonuclease I, proliferating cell nuclear antigen, replication factor C, DNA polymerase δ and DNA ligase I are all implicated.[165]

DNA lesions affecting only one DNA strand are repaired by the NER system which is subdivided into global genome and transcription-coupled repair. The latter is initiated by stalling of the RNA polymerase II and depends on the ERCC6 and ERCC8 proteins. However, proteins further down-stream in both pathways are thought to be identical. The helicases ERCC2 and ERCC3 unwind the DNA around the lesion while the endonucleases ERCC1-ERCC4 and ERCC5 make incisions on the 5' and 3' side of the lesion, respectively. Thus, the lesion is released as an oligonucleotide and repair is completed by DNA synthesis and ligation.[166]

Double-strand breaks occur in the cell as part of the recombination between homologous chromosomes during meiosis, during the V(D)J recombination of lymphocyte maturation, or as a result of collapsed replication forks.[167] DSB can be repaired by HR, a template-dependent and therefore error-free pathway that is most active during the late S phase and G2 phase when the cell contains two DNA copies.[168]

In HR, DSB are first processed by the MRN complex (MRE11-RAD50-NBS1) which generates single-stranded 3' overhangs. The ssDNA is then coated with replication protein A and RAD52, which assist in loading RAD51 and RAD54. Other proteins implicated in this process are BRCA1 and 2, as well as the RAD51 paralogs XRCC2 and 3.[169] The resulting nucleoprotein filament invades the undamaged homologous dsDNA and forms Holliday junctions. DNA polymerase extends the ssDNA ends using the homologous sequence as template, followed by branch migration, Holliday junction resolution and DNA ligation.[169]

In contrast to HR, NHEJ is template-independent and thus very error-prone. This repair pathway is initiated by binding of the Ku70/86 (also known as Ku70/80) heterodimer to the damaged DNA ends and subsequent recruitment of DNA-PKcs (DNA-dependent protein kinase catalytic subunit). The resulting DNA-PK holoenzyme activates Artemis, an endonuclease that can process unsuitable DNA ends. Lastly, DNA ligase IV and XRCC4 catalyze ligation of the DNA ends.[169]

Cell cycle checkpoints are triggered by the presence of DSB or ssDNA in the cell and thus progression through the cell cycle is dependent on the success of DNA repair systems in restoring genome integrity.[170] The protein kinases, ATM and ATR, play central roles in the signalling pathway between DNA damage and cell cycle regulation as they phosphorylate and thus, activate CHK1, CHK2 and p53 among others. In turn, CHK1 and CHK2 phosphorylate and hence inactivate the CDC25A phosphatase which ultimately prevents replication initiation and triggers cell cycle arrest, while p53 is responsible for maintaining the arrest.[167] Eventually, cell cycle will resume or cells will undergo programmed cell death.

1.4.4.2 Conventional gene targeting

Conventional gene targeting is now routinely used in the production of transgenic animals for functional studies of genes or as disease models.[171] As such, the technology has had an immense influence on biological and medical research and this was recognized by awarding Mario Capecchi, Martin Evans and Oliver Smithies the Nobel prize in Physiology or Medicine 2007 for their pioneering work in this area.[172]

However, conventional gene targeting relies on specially designed targeting vectors that contain the gene of interest with the intended modifications together with selectable marker genes and restriction endonuclease sites for screening.[173] In addition to the already mentioned problem with down-regulation of bacterial and viral sequences, vector production and screening of integrants is labour-intensive.[174]

1.4.4.3 Small fragment homologous replacement

In contrast to conventional gene targeting, the donor DNA in SFHR is relatively short (usually 400-1000 bp[175,176]) and only contains homologous DNA sequences with the exception of the intended sequence alterations. Thus, production of donor DNA is simplified. Although first applied to the targeted correction of mutations in the cystic fibrosis trans-membrane conductance regulator gene,[175,176] SFHR has been shown to mediate gene targeting in the dystrophin,[177] β-globulin[178] and survival of motor neuron 1 genes.[179] Evidence showing integration of the donor DNA into an episomal target sequence supports a homologous recombination mechanism.[180]

Chromosomal gene targeting efficiencies of up to 20% have been described in some studies,[175,177] but another laboratory demonstrated that their own results showing similar efficiencies were in fact due to false positive results caused by a PCR artefact.[181] In this study, one of the PCR primers was chosen inside the region of homology between the exogenous and endogenous DNA sequences, as had been done in the studies showing high efficiency.[175,177] Also, phenotypic evaluation of chromosomal targeting in recombinant reporter gene cell lines yielded much lower efficiencies (≤0.02%).[182-185]

1.4.4.4 Adeno-associated virus vectors

In an effort to exploit the efficiency of viral gene entry into cells, retroviral,[186] adenoviral[187] and AAV[188] vectors have been evaluated for use in gene targeting. Only AAV vectors showed significantly improved efficiency (0.1-1%[189]) compared to conventional transfection approaches. Although AAV-mediated gene targeting mostly introduced only the intended nucleotide alterations at the target site, the ratio of random integration to gene targeting was found to be around 10:1.[190] Thus, non-homologous sequences coding for selectable markers and employing promoter trap strategies have been added to the donor DNA.

On the other hand, AAV-mediated gene targeting has been confirmed by Southern blot analysis in mammalian cells[189,191] and in two different mouse models, though *in vivo* efficiency was below one per 10^4 cells.[192] The primary mechanism of AAV-mediated gene targeting appears to be HR, as RNAi-mediated down-regulation of RAD54L, RAD54B and XRCC3 significantly reduced or abolished long-term expression.[193] It also seems that random integration and gene targeting are distinct mechanisms,[194] so perhaps AAV can be engineered to avoid the former and enhance the latter.

1.4.4.5 Zinc-finger nucleases

In 1994 it was shown that the homing endonuclease SceI from *Saccharomyces cerevisiae* induced DSB in mammalian cells at specific sites containing its 18 bp long recognition site[195] and that this stimulated HR[196] and gene targeting.[197] However, most mammalian genes do not contain homing endonuclease recognition sites. At the same time, chimeric restriction endonucleases consisting of 3-4 DNA-binding zinc finger motifs and a FokI cleavage domain were created.[198,199] Since the FokI nuclease requires dimerization of the cleavage domain, two ZFN subunits must bind on opposite strands of the DNA in close proximity, creating an 18-24 bp recognition sequence.[200]

When mammalian cells were transfected with plasmids encoding engineered ZFN and donor DNA, ZFN-mediated gene targeting resulted in up to 10% in a mutated EGFP system.[201] Targeting of the interleukin 2 receptor γ gene with optimized 4-finger ZNF yielded 18% corrected cells in K562 and 5% in primary human T cells without any selection.[202] However, ZNF-mediated gene targeting suffers from high levels of cytotoxicity, likely caused by excessive and non-specific DNA cleavage due to imperfect target site recognition or dimerization of the ZFN.[203]

Also, DSB induced by endonucleases are not always repaired using only the intended homologous donor DNA only, as parts of the plasmid ZNF vector, especially of the CMV promoter, were also used as donor.[204] Furthermore, simultaneous delivery of two different ZFN subunits and a donor DNA is a major challenge for current gene delivery systems. Thus, further optimization of ZFN and homing endonuclease design as well as of delivery systems is necessary before this technique can be applied to the clinic.

1.4.5 Gene editing

In contrast to gene targeting, which generally replaces large parts of the target gene, gene editing aims to replace or repair the sequences immediately surrounding the disease-associated mutation only and is thus limited to modifying 1-4 nucleotides in close proximity. Gene editing is mediated by oligonucleotides of various structures such as TFO, chimeraplasts and ssODN. In contrast to plasmids, oligonucleotides can freely diffuse into the nucleus of mammalian cells due to their small size, enabling high *in vitro* transfection efficiencies. It is not yet clear if the oligonucleotide-based gene editing methods described here are mediated by HR.

1.4.5.1 Triplex-forming oligonucleotides

TFO can bind to the major groove of double-stranded homopurine-rich DNA in a sequence-specific manner resulting in a triple helix structure. These reagents have been used to inhibit gene transcription,[205] induce mutagenesis[206] or stimulate HR with the TFO alone[207] or coupled to a mutagen such as psoralen.[208] Also, a bifunctional TFO containing an oligonucleotide as donor DNA was shown to mediate up to 0.1% gene editing in a recombinant single copy reporter gene cell line.[209]

A recent paper reported that LNA-modified TFO linked to orthophenanthroline, a DNA cleaving molecule, induced sequence-specific DSB and that in the presence of homologous donor DNA more than 1.5% of targeted cells demonstrated gene targeting.[210] So far, application of TFO-mediated gene targeting has been limited by the target sequence requirement of 15-30 bp long pure polypurine stretches and by their binding inefficiency at physiological pH, but progress has been made in addressing both issues.[211]

1.4.5.2 Chimeraplasts

Chimeraplasts are synthetic RNA/DNA chimeric oligonucleotides, typically 68 nucleotides long that can form a double hairpin with its target sequence as shown in Figure 1-8. The first paper using chimeraplasts for gene-editing was published in 1996 and provided evidence for up to 30% editing efficiency in an episomal assay.[212] A second paper, published in the same year, presented data from a PCR restriction fragment length polymorphism (PCR-RFLP) analysis suggesting up to 50% gene editing efficiency when targeting the β-globin gene in lymphoblastoid cells from a sickle cell anaemia patient.[213]

However, both papers were soon criticized for their lack of controls to exclude PCR artefacts, the absence of data showing gene editing in isogenic cell lines and, in the second paper, inconsistencies in the sequencing data, suggesting cell culture contamination as the real cause of the observed nucleotide alterations.[214,215] Indeed, the senior investigator on both papers, Eric Kmiec, himself admitted in an author reply that contamination with wild-type cells was a likely explanation for the sequencing data irregularities.[215]

Figure 1-8 Design of chimeraplasts

An example chimeraplast designed to edit the wild-type ApoA-I and resulting in the ApoA-I_M sequence (which differs from ApoA-I by C→T; Arg173→Cys). The bottom all-DNA sequence is complementary to the *APOA-I* genomic sequence with the exception of a central mismatched T (underlined). The top chimeric strand of the double-stranded molecule contains 2'-O-methylated protected RNA residues (lower case), which provide strong RNA-DNA hybridization, and hence efficient targeting, as well as resistance to RNase H-mediated degradation. Its complementary A (underlined) to the T mismatch in the all-DNA strand is in the middle of a 5-base-DNA region between the 10-base RNA sequences. The structure is stabilized by a pentameric G/C clamp, while polyT hairpin caps provide flexibility.

Papers from another laboratory described the use of chimeraplasts for targeting the alkaline phosphatase gene in human hepatoma cells, with up to 43% initial gene editing efficiency as determined by colony lift hybridization analysis[216] and targeting the wild-type factor IX gene in rats, resulting in up to 40% editing efficiency and a subsequent 50% reduction in blood coagulation.[217] However, another group attempting chimeraplast-mediated gene editing, using a wide variety of experimental conditions was unable to isolate clones containing the intended sequence alterations in experiments where PCR-RFLP or allele-specific PCR had detected 2-10% gene editing.[218]

Without showing any supporting data, they concluded that their results from the PCR-based analysis must have been due to PCR artefacts and suggested, therefore, that PCR-based analysis of gene editing was suspect in general.[218] They further suggested that the reduction in blood coagulation seen by Kren *et al.*[217] was caused by liver toxicity of the transfection reagent or antisense effects of the oligonucleotide.[218]

Others also questioned if 40% gene editing efficiency was realistically achievable *in vivo* considering the problems with liver gene delivery and an assumed maximal efficiency of 50% based on the proposed MMR-mediated mechanism.[219] On the other hand, further proof of gene editing was obtained by targeting a mutant tyrosinase gene resulting in black pigmentation in albino mice melanocytes[220] and the skin of albino mice.[221] In both studies, analysis by PCR-RFLP three months after the chimeraplast treatment confirmed single allele conversion. Thus, it is highly unlikely that these results were caused by PCR-based artefacts. Also, several independent groups from around the world showed chimeraplast-mediated gene editing in the dystrophin,[222,223] tyrosinase[224] and β-globin[225] genes.

As reports of failures have kept coming over the past years,[226-233] it has become clear that chimeraplast-mediated gene editing suffers from poor reproducibility among different research groups and even within the same group as seen in our own laboratory. For example, we showed that the dysfunctional apoE2 isoform was efficiently converted to wild-type apoE3 in recombinant CHO cells as demonstrated by genomic sequencing and iso-electric focusing gel-electrophoresis of secreted protein.[234] However, attempts to extend the work and correct the human ε4 allele,[235] which causes a dominant hyperlipidaemia or introduce the E2 and E4 mutations into wild-type cells,[236] were disappointing.

Conversions were noted, but were unstable and cloned cells could not be isolated.[235,236] On the other hand, we could show that the PCR-RFLP results were not artefactual, as spiking experiments with chimeraplast or even degraded chimeraplast did not yield any false-positive PCR products.[235] Similarly, the *APOA-I* gene was successfully targeted to generate ApoA-I_M, but the result could not be reproduced in subsequent experiments.[237] These problems were attributed to a poorer quality of chimeraplast, consistent with changes to production methods of these difficult-to-synthesize reagents.[237]

1.4.5.3 Single-stranded oligonucleotides (ssODNs)

Introduction

In 1989, Campbell *et al.*[238] described the use of ssODN for gene editing of episomal targets in mammalian cells, but the technique did not gain much attention until experiments by Gamper *et al.*[239] revealed that the all-DNA strand was the functionally active domain in chimeraplasts and that ssODN containing PTO linkages at both ends were up to 3.7 times more effective in mediating gene editing in cell-free extracts than chimeraplasts. Another paper also showed that ssODN with 2'-O-methyl RNA nucleotides at both ends could correct a mutant lacZ gene in up to 0.1% of recombinant CHO cells, whereas chimeraplasts did not.[240]

Two other comparisons have shown that ssODN can reliably edit chromosomal DNA in mammalian cells, while chimeraplasts resulted in little or no gene editing even in episomal assays.[183,241] Since then, gene editing research has shifted from using unreliable chimeraplasts as the functional reagent to the much cheaper and easier to produce ssODN and several independent groups have reported consistent and reproducible gene editing efficiencies.[184,241-248] In response to the potential of possible false-positive results from PCR-based detection, cell lines expressing mutated forms of EGFP have been developed (e.g. CHO-mEGFP cells; see also chapter 4).[228,249,250]

However, the best chromosomal gene editing frequencies described in such systems are less than five percent in cell lines with multiple copy integration[251] and around 0.1-0.2% in cells with low copy number integrations of the target gene.[241] These results are nowhere near the gene editing efficiencies previously reported with chimeraplast-mediated gene editing and PCR-based detection and thus reinforce the belief that the high efficiencies previously seen were due to PCR artefacts.

Sequencing results demonstrating the presence of specific base changes in pools of treated cells were similarly suspected of being caused by PCR artefacts.[218] Furthermore, it was inferred that gene editing introduced transient changes only and that genomic conversions were not heritable;[252] this was seemingly endorsed by the observation that gene-edited green fluorescent cells were quickly lost from the total population of targeted CHO-mEGFP cells.[242]

In response to these and other criticisms of chimeraplast-mediated gene editing, Albuquerque-Silva *et al.*[227] proposed four experimental criteria for rigorous validation. First, the targeting oligonucleotide must introduce a rare mutant genotype to rule out spontaneous reversion events. Second, this genotype should not exist in cells within the laboratory to prevent cross-contamination artefacts. Third, mutated clones must be studied individually to exclude PCR and sequencing artefacts caused by large molar amounts of targeting oligonucleotides. Fourth, editing efficiency must be confirmed at the protein level in pooled cells at time points that eliminate artefacts of selection and expansion of rare spontaneous mutations.

These criteria are equally applicable to ssODN mediated gene editing but, so far, only one paper has studied individual gene-edited clones.[241] Unfortunately, this study used a mEGFP containing a premature stop codon and the gene-edited genotype did not allow exclusion of spontaneous reversion events, though a down-stream sequence tag excluded contamination with a wild-type EGFP.[241] Thus, unequivocal confirmation of genotype alteration is still lacking.

Oligonucleotide design

Design of the targeting ssODN has critical effects on the resulting gene editing efficiency and the viability of edited cells. Typically, targeting oligonucleotides are 25-30 nucleotides long, as longer ssODN have only shown slightly higher gene editing efficiencies,[240,243] and not in all cases.[183,246,253] Moreover, longer oligonucleotides are more expensive and require purification because of the increased percentage of -1 product.

In general, ssODN contain 3-6 PTO linkages or 2'-O-methyl nucleotides at each end to protect against exonuclease digestion,[239,240] though in a direct comparison of the two chemistries the PTO backbone modifications mediated more episomal gene editing in CHO cells.[245] Other oligonucleotide modifications such as LNA residues have also been used, but there is disagreement as to whether these are more efficient than PTO-modified ssODN.[184,246] A recent paper showed that 5' acridine-conjugated oligonucleotides were significantly more effective than either LNA-modified, psoralen-conjugated or peptide-nucleic acid containing oligonucleotides.[254]

However, modifications may influence viability of gene-edited cells, as PTO-modified ssODN were found to induce more cell death than unmodified ssODN and, in contrast to longer or modified ssODNs, 13- or 19-mer unmodified ssODNs did not seem to induce any growth inhibition.[241] Also, targeting of mutS homolog 2 (MSH2) deficient embryonic stem cells containing a mutated neomycin gene with unmodified ssODN yielded 2-7 times more G418 resistant colonies.[243]

Several papers have reported that the optimal location for the mismatch to the target sequence is in the centre or slightly towards the 3' end of the oligonucleotide as shifts to the 3' end were less disruptive than shifts towards the 5' end and in some cases even increased efficiency.[246,255-258] Curiously, blocking the 3' end with octyl or 3' hydrogen groups reduced editing efficiency significantly, but only in the presence of PTO linkages, whereas 5' octyl or phosphate groups had little influence.[251,259]

It has also been shown that 2-4 nucleotide mismatches between ssODN and target sequence are corrected together if they are in close proximity, whereas this conversion linkage decreases with increased separation.[243,244,247,255] Some publications even found that increasing the number of mismatches in close proximity increased efficiency.[242,260,261] The two papers from the te Riele laboratory attributed this phenomenon to decreased recognition by the MMR.[260,261]

Generally, the sequence of the targeting ssODN is complementary to the non-transcribed strand, i.e. it has the same sequence as the antisense strand, as numerous papers have reported a strand bias favouring this orientation.[183,240,246,251,262,263] However, in some reports no strand bias[184,241,243] or opposite strand bias[253,264] was observed.

Mechanism of ssODN-mediated gene editing

The detailed mechanisms involved in oligonucleotide-mediated gene editing are thus far not fully understood, but recent evidence points to some degree of physical sequence substitution with the genomic sequence as the initial step (Figure 1-9). For example, the fact that blocking the 3' end results in massive reductions in gene editing efficiency, while a 5' block does not, suggests that elongation at the 3' end is necessary.[251,259]

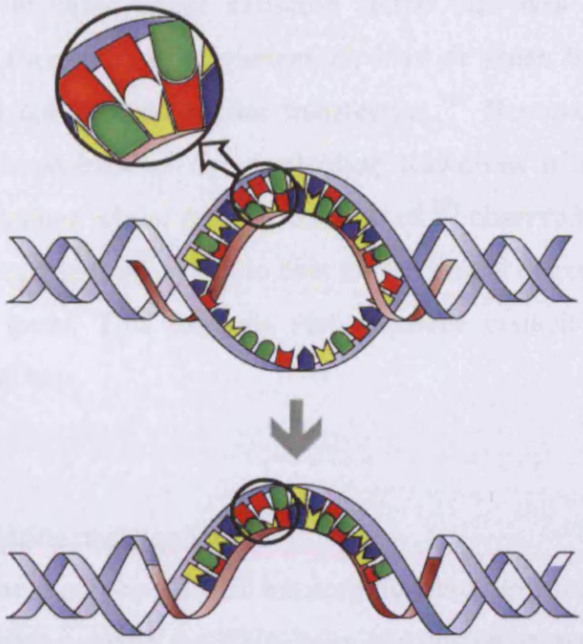


Figure 1-9 Mechanism of ssODN-mediated gene editing

The targeting oligonucleotide (brown) is thought to access the double-strand helix (grey) during transcription and/or replication. The ssODN then binds to its complementary sequence (with the exception of the intended mismatch) and is integrated in an as yet undefined process. The mismatch with the endogenous sequence is resolved in a second step.

Also, in an elegant experiment, Radecke *et al.*[259] captured genomic DNA fragments isolated from gene-edited cells that had been treated with a biotin-labelled ssODN using streptavidin-beads. PCR-amplification from these DNA fragments using target gene specific primers resulted in a defined PCR product, demonstrating that the biotin-labelled ssODN had been integrated into the genomic sequence.

Olsen[164] speculated that if the ssODN was physically integrated into the homologous genomic sequence, then an antisense ssODN incorporated into the transcribed strand would lead to expression of the gene-edited phenotype before the mismatch between ssODN and genomic sequence was resolved. Integration of sense ssODN, on the other hand, would necessitate repair of the mismatch prior to expression of the edited phenotype. Thus, gene editing using sense ssODN could simply take longer and hence variation in time to analysis could explain strand bias.

He established that in cases where antisense strand bias was found, analysis was performed 24-48 h after treatment, whereas no bias or sense bias was found when analysis was carried out 7-10 days after transfection.[164] However, Sorensen *et al.*[253] observed opposite strand bias for two nucleotide alterations in the same gene in an episomal CHO cell system, whilst Andrieu-Soler *et al.*[184] observed antisense strand bias for a blue fluorescent protein gene, but no bias for the retinal degeneration gene 1 at the same chromosomal locus. This suggests that sequence context may play a role in oligonucleotide strand bias.

Influence of Transcription and Replication

Strand bias has also been associated with transcription and replication of the target gene. In theory, both processes open up the DNA helix and enable access of the ssODN to the homologous endogenous sequence without the need for double-strand invasion. During active transcription access to the transcribed strand is blocked by the transcription complex while the non-transcribed strand is accessible.[265] Indeed, Liu *at al.*[262] showed that in *E.coli* T7 phage polymerase displaced sense ssODN from the transcribed strand of a plasmid, but the antisense ssODN remained hybridized to non-transcribed strand.

Several papers have also observed an association between gene editing efficiencies and transcription in *S. cerevisiae*,[262] *E. coli*,[266] mammalian episomal targets[267] and in individual cell lines derived from the same parent population, but containing disparate copy numbers of the target gene.[263,268] On the other hand, Olsen *et al.*[242] did not observe gene editing in G1 arrested cells suggesting that transcription without replication may not lead to successful gene editing.

Assuming physical integration of the ssODN into the transcribed strand as discussed on the previous page, this observation could suggest that the mismatch between integrated ssODN and endogenous sequence is not easily resolved and that only replication results in a completely altered sequence in one of the daughter cells. This was seemingly confirmed by data showing that in CHO-mEGFP cells only one daughter cell of dividing gene-edited cells remained strongly green fluorescent.[242]

In *E.coli* λ-Red recombination systems oligonucleotide-mediated gene editing appears to be replication dependent and strand bias is for ssODN hybridizing to the template strand for lagging strand synthesis.[269] However, in mammalian episomal[267] and chromosomal[251] systems strand bias is not dependent on chromosomal position or lagging strand synthesis. Yet, there is a clear correlation between replication and gene editing, as cells synchronized in the S phase showed higher gene editing efficiencies[242,268,270] and treatment with aphidicolin, a polymerase inhibitor, significantly reduced rates.[251]

Also, slowing replication fork movement by incubation with thymidine[271] or 2'3'-dideoxycytidine[272] increases gene editing, while completely stopping replication with 1-β-D-arabinofuranosylcytosine does not.[272] Other DNA damaging agents such as hydroxylurea,[270] etoposide,[270] camptothecin,[273] bleomycin[274] and methyl methanesulphonate[275] increase gene editing as well. These findings support a central role for HR-mediated DSB-break repair which is mainly active in the S phase[276] and can be induced by introducing DSB into the genome.[197]

Influence of DNA damage response pathways

Notwithstanding, a recent paper demonstrated that SceI-induced HR DSB-repair was completely inhibited by caffeine and pentoxifylline whereas ssODN-mediated gene editing was only partially repressed.[277] Both reagents are inhibitors of the ataxia telangiectasia mutated (ATM)/ataxia telangiectasia and Rad3 related (ATR) damage response pathway.[278] Furthermore, Wang *et al.*[277] showed that although thymidine addition increased gene editing, it actually reduced DSB-repair. These findings indicate that in contrast to DSB-induced HR repair, gene editing is not dependent on activation of the ATM/ATR pathway.

On the other hand, data presented by Ferrara *et al.*[279] hinted that the checkpoint kinases CHK1 and CHK2 which are down-stream substrates for ATM/ATR are only phosphorylated and hence activated in gene-edited cells. Also, chimeraplast-mediated gene editing was increased 2-fold compared to wild-type in cells over-expressing RAD51, a central protein in HR,[249] while low editing levels in CHO cells deficient for the RAD51 paralog XRCC2 could be rescued by complementation with XRCC2 cDNA.[250]

In *S. cerevisiae*, over-expression of RAD51 raised editing efficiency 3-fold and a mutated RAD51 with higher affinity for its partner RAD54 amplified editing levels 100-fold.[280] Moreover, co-injection of recombinant RAD51 and/or RAD54 with ssODN produced recombinant mouse embryos, whereas ssODN alone did not.[248] A similar finding was reported in zebra fish, where only ssODN annealed with a RAD52 mutant protein resulted in recombinant embryos.[258] However, suppression of NHEJ by co-injection with Ku70/86 antibodies generated the highest number of mouse embryos with altered sequence.[248]

Conversely, the MMR pathway seems to have an inhibitory effect on oligonucleotide-mediated gene editing. Dekker *et al.* could show successful editing in MSH2[243] or MSH3[260] deficient mouse embryonic stem cells but not in wild-type cells which had also shown little or no repair in previous publications.[183,245] Also, expression of mutL homolog 1 (MLH1) in HEK293T cells known to be deficient in this MMR protein,[281] reduced editing levels by 50-70% (P. Olsen, unpublished data). Transient down-regulation of MSH2 in mouse embryonic stem cells using RNAi also increased gene editing significantly, but the increase was dependent on the number of mismatches.[261]

This antagonistic effect of MMR proteins on oligonucleotide-mediated gene editing is generally attributed to their role in preventing heteroduplex formation between homologous sequences with sequence discrepancies.[282] However, siRNA-mediated suppression of MSH2 in CHO-mEGFP cells did not affect ssODN-induced G2 arrest 16 h after treatment, though it did result in a higher proportion of cells resuming a normal cell cycle 40 h after transfection (I. Papaioannou, unpublished data). Thus, MMR might instead play a role in the second step of gene editing, when the mismatch between integrated ssODN and endogenous sequence has to be resolved.

Recently, Igoucheva *et al.*[283] have shown that CHO cells deficient in the excision repair cross-complementation group (ERCC) 1, 2, 5 or 6 proteins exhibited reduced episomal gene editing efficiency compared to the parental cell line. The ERCC1 and 4 genes encode a 5' endonuclease, ERCC5 a 3' endonuclease and ERCC6, the Cockayne syndrome B protein, which couples excision repair and transcription. Lack of two helicases (ERCC2 and 3), which are involved in unwinding DNA near a lesion, did not affect gene editing. Thus, the nucleotide excision repair pathway (NER) may also play a role in oligonucleotide-mediated gene editing.

1.5 Aims of thesis

Gene therapy strategies for the introduction of sequence specific modifications in the genome of mammalian cells hold great promise for lasting therapeutic treatment of many diseases. In this study I explore the use of oligonucleotide-mediated gene editing to introduce the anti-atherogenic point mutation ApoA-I_M into the genomic sequence of ApoA-I. Additionally, I investigate various aspects of oligonucleotide-mediated gene editing using mEGFP model systems.

In particular, I will

1. Demonstrate successful gene editing *in vitro* using either CHO-AI or HepG2 cells by PCR-RFLP and sequencing.
2. Confirm that the RFLP results are not artefacts.
3. Evaluate the influence of experimental conditions such as transfection reagent and oligonucleotide design on gene editing efficiency and viability of edited cells using a mEGFP model system.
4. Investigate ways to elevate the gene editing efficiency or select for edited cells independent of a marker gene.
5. Validate oligonucleotide-mediated gene editing using the experimental criteria set out by Albuquerque-Silva.
6. Show that the observed loss of green fluorescent cells from the population of edited cell clones is due to phenotypic instability in the mEGFP model system and is not specific to the gene-edited allele.
7. Verify that gene editing efficiencies in three mEGFP systems were significantly underestimated due to the phenotypic instability.
8. Exploit the phenotypic instability to show a positive association between target gene expression and gene editing efficiency in three mEGFP cell lines.
9. Assess an ApoA-I_M homodimer specific antibody for use in flow cytometry to isolate gene-edited cell lines expressing ApoA-I_M.

CHAPTER 2: MATERIALS AND METHODS

2.1 Materials

2.1.1 Cell Biology

2.1.1.1 Cell lines, plasmids and antibodies from non-commercial sources

CHO-AI and CHO-AIM are adherent cell lines derived from CHO^{dhfr-} cells by stable transfection with the plasmids p7055.AI and p7055.AIM, which express ApoA-I and ApoA-I_M, respectively.[237] This expression vector also contains the *DHFR* gene, thus allowing selection by incubation in hypoxanthine and thymidine lacking medium. Clones were isolated by limiting dilution and screened for protein secretion.[237]

The CHO-mEGFP and HEK293T-mEGFP cell lines as well as the pmEGFP plasmid were kindly provided by Prof. Stefan Krauss (Cellular and Genetic Therapy Research Group, Institute of Microbiology, Rikshospitalet, Oslo, Norway). The CHO-mEGFP cells have been described previously.[242,251] In brief, a Tyr66 (TAC) to Ser66 (TCC) missense mutation was introduced in the pEGFP-C1 plasmid (Clontech, Takara Bio) to give non-fluorescent, full-length EGFP (mEGFP). CHO-K1 (ATCC CCL-61) cells were stably transfected with this plasmid and individual clones were selected with 400 µg/ml G418. CHO-mEGFP cells contain 38 copies of the mEGFP transgene.[251]

To produce the HEK293T-mEGFP cell line, HEK293T cells (ATCC CRL-11268) were transfected with a pS66-IRES-puro construct made by cloning the EGFP-S66 fragment into the pIRES-puro3 vector (BD Biosciences/Clontech) which contains the internal ribosome entry site (IRES) of the encephalomyocarditis virus. After selection in 500 µg/ml puromycin, individual clones were isolated. The mEGFP transgene copy number has not been determined in these cells.

HepG2 cells were acquired from the European Collection of Cell Cultures (ECACC 85011430). The red fluorescent protein (RFP) expressing plasmid was kindly provided by Dr. Takis Athanasopoulous (Royal Holloway, University of London).

The monoclonal mouse anti-human ApoA-I_M antibody (clone 17F3) was kindly made available by Dr. Paul L. Barclay at Pfizer Global Research & Development, Groton, USA. It is of the IgG2a subclass with Kappa light chains.

2.1.1.2 Cell culture reagents

Alpha minimal essential medium (Alpha-MEM Invitrogen, Paisley, UK)

Caltag Fix&Perm solution (Invitrogen)

Dimethylsulphoxide (DMSO; Sigma, Gillingham, UK)

Dulbecco's modified Eagle's medium (DMEM; Invitrogen)

Draq5 (Biostatus, Shepshed, UK)

Fetal bovine serum (FBS, heat-inactivated for 30 min at 55 °C; Invitrogen)

Hoechst 33342 (Invitrogen)

Normal goat serum (Abcam, Cambridge, UK)

G-418 sulphate (Invitrogen)

Glutamax (Invitrogen)

Ham F12 (Invitrogen)

HEPES buffer (Invitrogen)

HT supplement (Invitrogen)

Iscove's DMEM (Sigma)

L-glutamine (Sigma)

Non-essential amino acids (Invitrogen)

Opti-MEM (Invitrogen)

Propidium iodide (PI; Sigma)

Rabbit polyclonal anti-EGFP (ab6556 from Abcam)

Rabbit polyclonal anti-human C-reactive protein (A-0073 from Dako, Ely, UK)

Sodium butyrate (Sigma)

Streptomycin (10 mg/ml) and penicillin (10,000 units/ml) solution (Sigma)

Trypan blue (Sigma)

Trypsin-EDTA (Sigma; 0.25% (w/v) trypsin and 0.02 % (w/v) EDTA)

Zenon Alexa Fluor 647 Rabbit IgG Labelling kit (Invitrogen)

Zenon Alexa Fluor 488 Mouse IgG2a Labelling kit (Invitrogen)

5-aza-2'-deoxycytidine (Aza-dC; Sigma)

2.1.1.3 Transfection reagents

jetPEI (Autogen Bioclear, Calne, UK)

Lipofectamine 2000 (Invitrogen)

Deacylated LPEI (kindly provided by Dr. Maya Thanou, Imperial College London)

CDAN/DOPE (also provided by Dr. Maya Thanou)

Atelocollagen (Koken, Tokyo, Japan)

2.1.2 Molecular Biology

2.1.2.1 Molecular biology kits and reagents

Agarose (Biogene, Kimbolton, UK)

[α-32P] dCTP (GE Healthcare, Amersham, UK)

Ampicillin sodium salt (Sigma)

Bovine serum albumin (BSA; Sigma)

DH5α chemically competent *E.coli* (Invitrogen)

DNeasy Tissue Kit (Qiagen, Crawley, UK)

dNTP set (Invitrogen)

Ethidium bromide (Sigma)

Hybond XL nitrocellulose membrane (GE Healthcare)

Illustra ProbeQuant G-50 microcolumns (GE Healthcare)

Isopropanol (Sigma)

IPTG (Sigma)

Kanamycin (Sigma)

Kodak Biomax-MS autoradiography film (GE Healthcare)

Metaphor high resolution agarose (Cambrex Bio Science, Wokingham, UK)

Microcon YM-100 centrifugal filter units (Millipore, Watford, UK)

NuPAGE Novex 20% Tris-borate polyacrylamide gels (Invitrogen)

Pfu Turbo DNA polymerase (Stratagene, Amsterdam, The Netherlands)

pGEM-T Easy vector (Promega)

Proteinase K (Sigma)

QIAquick gel extraction kit (Qiagen)

QIAquick spin miniprep kit (Qiagen)

Rapid Hybridization buffer (GE Healthcare)

Ready-to-go DNA labelling beads (GE Healthcare)

RNase A (Sigma)

Restriction endonucleases (New England BioLabs, Ipswich, MA, USA)

Taq DNA polymerase (Promega, Southampton, UK)

Tween 20 (Sigma)

T4 DNA Ligase (New England BioLabs)

X-gal (Sigma)

100/G Genomic tips (Qiagen)

2.1.2.2 Buffers and solutions

Table 2-1 Qiagen buffers for DNA extraction and purification

C1	1.28 M sucrose, 40 mM Tris-HCl [pH 7.5], 20 mM $MgCl_2$, 4% Triton X-100
G2	800 mM guanidine HCl, 30 mM Tris-HCl [pH 8], 30 mM EDTA [pH 8], 5% Tween-20, 0.5% Triton X-100
QBT	750 mM NaCl, 50 mM MOPS [pH 7], 15% isopropanol, 0.15% Triton X-100
QC	1 M NaCl, 50 mM MOPS [pH 7], 15% isopropanol
QF	1.25 M NaCl, 50 mM Tris-HCl [pH 8.5], 15% isopropanol
P1	50 mM Tris-HCl [pH 8], 10 mM EDTA, 100 μg/ml RNase A
P2	200m M NaOH, 1% SDS (w/v)

Table 2-2 Electrophoresis buffers

10x TBE running buffer	108 g/L Tris base (trizma), 55 g/L boric acid, 40 ml 0.5 M EDTA [pH 8.0]
Novex TBE Hi-Density Sample Buffer (Invitrogen)	18 mM Tris base, 18 mM boric acid, 0.4 mM EDTA, 3% Ficoll, 0.02% bromophenol blue, 0.02% Xylene Cyanol
UltraPure 10x TAE running buffer (Invitrogen)	400 mM Tris-acetate, 10 mM EDTA
10x Bluejuice Gel Loading Buffer (Invitrogen)	65% sucrose, 10 mM Tris-HCl, 10 mM EDTA, 0.3% bromophenol blue

Table 2-3 Southern blot buffers and solutions

20x SSC	175 g/L NaCl, 88 g/L $Na_3citrate{\cdot}2H_2O$, adjust pH to 7.0 with 1 M HCl
Alkaline Transfer Buffer	0.4 M NaOH, 1 M NaCl
SB wash Buffer I	2x SSC, 0.1% (w/v) SDS
SB wash Buffer II	0.1x SSC, 0.1% (w/v) SDS

Table 2-4 Bacterial growth media

Lysogeny broth (LB; Sigma)	10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl
LB agar (Sigma)	10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, 15 g/L agar
Super optimal catabolite repression broth (S.O.C; Invitrogen)	20 g/L tryptone, 5 g/L yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM $MgSO_4$, 20 mM glucose

Table 2-5 Other buffers

Electroporation buffer	120 mM Na_2HPO_4 [pH 7.2], 40 mM mannitol, 15 mM NaCl, 5 mM KCl
Phosphate-buffered saline (PBS)	138 mM NaCl, 10 mM Na_2HPO_4, 1.75 mM KH_2PO_4, 7.2 mM KCl [pH 7.4]
0.5 M EDTA [pH 8.0]	186.1 g/L $Na_2EDTA{\cdot}2H_2O$, adjust pH to 8.0 with ~ 50 ml 10 M NaOH

2.1.2.3 DNA ladders

All ladders were purchased from Invitrogen.

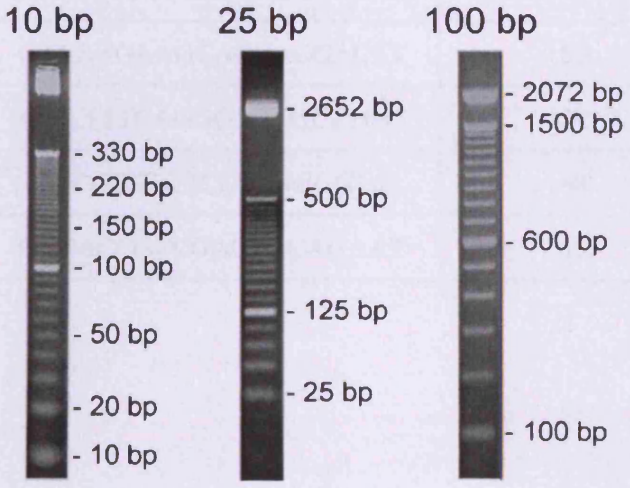


Figure 2-1 DNA ladders

2.1.2.4 PCR primers

Table 2-6 ApoA-I primers

Primer	Sequence 5' to 3'	G/C content in %	Tm in °C
AI2F	GGTGGAGCCGCTGCGCGC	83	79
AI4R	GCCAGTCTGGCGCCGCCG	83	79
ApoAIouterF	GAGGCAAGAGATGAGCAAGG	55	64
ApoAIouterF2	GAAGAAGTGGCAGGAGGAGA	55	63
ApoAIinnerF	AAGAGAAGCTGAGCCCACTG	55	64
ApoAIinnerR	CGCCGTTCTCCTTGAGAG	61	64
ApoAIinnerR 2	GGCCTTGGCGTGGTACTC	67	66
ApoAIouterR	CAGGAAGCTGACCTTGAAGC	55	64

Table 2-7 pEGFP primers

Primer	Sequence 5' to 3'	G/C content in %	Tm in °C
pEGFP-241for	AGTACGCCCCCTATTGACG	58	64
pEGFP-603for	CGCCACCATGGTGAGCAAGGG	67	76
pEGFP-844for	CACATGAAGCAGCACGACTT	50	61
pEGFP-754rev	GAATTTCAGGGTCAGCTTGC	50	64
pEGFP-1261rev	CGCGCTTCTCGTTGGGGTC	68	72
pEGFP-1380rev	GGTACCGTCGACTGCAGAAT	55	64

2.1.2.5 Gene editing oligonucleotides

Phosphorothioate links are denoted with a *, LNA modifications with a +, correcting nucleotides are in bold and underscored.

Table 2-8 ApoA-I to ApoA-I_M editing oligonucleotides

ssODN	Sequence 5' to 3'
huApoA-I_M-Cy5-49NT-PTO	Cy5-G*A*G*CCTCAAGGCGCGCGGCCAAGC**A**CTGGCGC AGCTCGTCGCTGTA*G*G*G
huApoA-I_M-49NT	GAGCCTCAAGGCGCGCGGCCAAGC**A**CTGGCGC AGCTCGTCGCTGTAGGG
huApoA-I_M-49NT-PTO	G*A*G*CCTCAAGGCGCGCGGCCAAGC**A**CTGGCGC AGCTCGTCGCTGTA*G*G*G
huApoA-I_M-49T-PTO	C*C*C*TACAGCGACGAGCTGCGCCAG**T**GCTTGGC CGCGCGCCTTGAGG*C*T*C

Table 2-9 mEGFP to edEGFP editing oligonucleotides

ssODN	Sequence 5' to 3'
mEGFP-2A-NT	GAAGCACTGCACGCC**AT**AGGTCAGGGT
mEGFP-2A-NT-PTO	G*A*A*GCACTGCACGCC**AT**AGGTCAG*G*G*T
mEGFP-2A-NT-LNA	+GAAGCACTGCACGCC**AT**AGGTCAGGG+T
mEGFP-Cy3-2A-NT-PTO	Cy3-G*A*A*GCACTGCACGCC**AT**AGGTCAG*G*G*T
mEGFP-3A-NT	GAAGCACTGCACGCC**AT**A**T**GTCAGGGT
mEGFP-3A-NT-PTO	G*A*A*GCACTGCACGCC**AT**A**T**GTCAG*G*G*T
mEGFP-Cy5-3A-NT-PTO	Cy5-G*A*A*GCACTGCACGCC**AT**A**T**GTCAG*G*G*T
mEGFP-3A-NT-LNA	+GAAGCACTGCACGCC**AT**A**T**GTCAGGG+T

Table 2-10 mEGFP to edEGFP control oligonucleotides

ssODN	Sequence 5' to 3'
mEGFP-Ctr-NT	GAAGCACTGCACGCCGGAGGTCAGGGT
mEGFP-Ctr-NT-PTO	G*A*A*GCACTGCACGCCGGAGGTCAG*G*G*T
mEGFP-Ctr-NT-LNA	+GAAGCACTGCACGCCGGAGGTCAGGG+T

2.1.3 Equipment

BioRad iCycler (Bio-Rad, Hemel Hempstead, UK)

Class II microbiological safety cabinet (Envair Ltd., Lancashire, UK)

Compact X4 Automatic X-ray film processor (X-ograph Healthcare, Tetbury, UK)

Cytopulse PA-4000 electroporator (Labtech International, Ringmer, UK)

FACScan flow cytometer with CellQuest v3.1 software (BD Biosciences, Oxford, UK)

FACSCalibur flow cytometer with CellQuest Pro software (BD Biosciences)

Heraeus Megafuge 1.0R (Fisher Scientific, Loughborough, UK)

Heraeus Biofuge Pico (Fisher Scientific)

Horizon gel apparatus (Life Technologies, Invitrogen)

Inverted phase-contrast microscope (Nikon TMS, Jencons-PLC)

Inverted fluorescence microscope (Nikon Eclipse TE200) with LUCIA G image analysis software

Nalgene Cryo 1°C Freezing Container (Fisher Scientific)

NanoDrop ND-1000 UV/Vis Spectrophotometer (Labtech, Ringmer, UK)

Partec PAS flow cytometer (Partec, Münster, Germany)

UVP Epi Chem II Darkroom (Jencons-PLC) with Labworks UVP image analysis software

Xcell SureLock Mini-Cell gel apparatus (Invitrogen)

37 °C incubator (Jencons Millenium CO2 incubator, Jencons-PLC, Leighton Buzzard, UK)

2.2 Methods

2.2.1 Cell Biology

2.2.1.1 Culture maintenance

Cell culture work was performed in a class II microbiological safety cabinet using disposable sterile plastic ware. All cell lines are adherent and were cultured in a 37 °C incubator with a humidified atmosphere of 5% CO_2 with addition of streptomycin (100 µg/ml) and penicillin (100 U/ml) to all media. Cells were split by trypsinisation when semi-confluent. After a quick wash with warm PBS, cells were incubated with 1 ml of 0.25% (v/v) trypsin-EDTA per 25 cm^2 of growth area, generally for 2 - 3 min at room temperature (RT). An equal volume or more of pre-warmed fresh growth medium containing FBS was added to neutralise the reaction and cells were split at the ratios indicated for each cell line.

CHO^{dhfr-} cells were maintained in Iscove's DMEM containing 10% (v/v) dialysed FBS, 1% (v/v) HT supplement (final concentration 16 µM thymidine, 100 µM sodium hypoxanthine) and 1% (v/v) non-essential amino acids. Cells were split 1:4 to 1:6 twice a week.

CHO-AI and CHO-AIM were maintained in Iscove's DMEM supplemented with 10% (v/v) FBS, 2 mM L-glutamine and 1% (v/v) non-essential amino acids. Cells were split 1:4 to 1:6 twice a week. CHO-mEGFP and CHO-K1 cells were cultured in Ham's F12 medium containing 2 mM L-glutamine and 5% (v/v) FBS. Cells were split twice a week at 1:6 to 1:8 ratios.

HEK293T and HEK293T-mEGFP cells were grown in DMEM with 10% (v/v) FBS, 2 mM L-glutamine and 4.5 mg/ml glucose, but without sodium pyruvate. Cells were split every 4-6 days at 1:4 to 1:6 ratios.

HepG2 cells were grown in DMEM augmented with 10% (v/v) FBS and 2 mM L-glutamine or, where indicated, in "Super HepG2 medium" (alpha-MEM containing nucleosides, 10% (v/v) FBS, 3.5 iu/ml insulin, 0.3 mg/ml Glutamax, 50 mg/ml BSA/linoleic acid, 10^{-7} M sodium selenite, 10^{-6} M hydrocortisone, 10^{-6} M thyrotropin-releasing hormone and 1.25 mg/ml D-glucose). Cells were split 1:10 to 1:20 every 4 - 6 days by trypsinisation (5 min at 37 °C). After trypsinisation HepG2 cells were squeezed 3 - 5 times through a 21G needle to break up cell clumps and ensure homogeneous plating. HepG2-mEGFP cells were cultured in Super-HepG2 medium.

2.2.1.2 Cryopreservation and Reconstitution

For cryopreservation, 10^6 to 10^7 cells were pelleted at 300 *g* for 5 min and re-suspended in 1 ml 10% (v/v) DMSO in FBS (15% for HepG2 cells). DMSO lowers the freezing point and encourages intracellular dehydration, thus allowing cells to be slowly cooled without damage. The cell suspension was pipetted into 1.8 ml cryovials and these were placed in a Nalgene Cryo 1 °C Freezing Container at -80 °C overnight. These special containers limit temperature reduction to 1 °C per min when stored at -80 °C and thus avoid damage to cells caused by latent heat of fusion. The cryovials were then transferred to liquid nitrogen for long-term storage.

For reconstitution, cryovials were removed from liquid nitrogen and rapidly warmed in a 37 °C water bath. As soon as the cell suspension was thawed, it was diluted 1:10 in fresh pre-warmed culture medium and transferred to culture flasks. Cells were allowed to adhere overnight before the medium was replaced with fresh medium.

2.2.1.3 Cell counting and viability

When cell counts were necessary, prior to plating for transfection or to assess cell viability, Trypan blue was used. This dye cannot pass through the membrane of live cells as it is a bis-azo dye and negatively charged at physiological pH, but it can traverse the membrane of dead cells. It appears as a blue stain under the phase-contrast microscope and thus allows simultaneous determination of cell count and viability. The haemocytometer chambers were prepared by firmly attaching a cover-slide with a drop of oil. Cells were trypsinised and resuspended to give a cell concentration of $>1x10^6$/ml.

A mix of 50 µl Trypan blue, 40 µl PBS and 10 µl of this cell suspension was pipetted into the haemocytometer chambers and observed under an inverted phase-contrast microscope. The number of viable cells was counted in eight 1 mm^2 squares with a total volume of 10^{-4} cm^3. The average number of cells per square was then multiplied by the dilution factor and 10^4 to give the number of cells per ml. To calculate the total number of cells, the number of cells per ml was multiplied by the volume (in ml) of the cell suspension. Generally the concentration of the cell suspension was adjusted to $1x10^6$ for pre-plating by subtracting the volume (in ml) of the cell suspension from the total number of cells (expressed as $x10^6$) and then adding the resulting number in ml of fresh medium to the cell suspension.

2.2.1.4 Transient transfection

Transfections were either performed on cells plated the previous day or using "reverse transfection",[284] where cells are trypsinised, counted and plated immediately before addition of the transfection complexes, rather than pre-plated. Transfection optimisation experiments showed that reverse transfection did not improve transfection efficiency of CHO cell lines, but instead caused increased toxicity. However, transfection efficiency was enhanced moderately in HEK293T cells and significantly in HepG2 cells if cell clumps were broken up by repeated passage through a 21G needle. This increase was probably due to the disruption of the dense growth pattern and tight cell-to-cell contacts in HepG2 cells that would normally preclude complete cell surface exposure to transfection agents, as happens with pre-plated cells. Slight membrane damage caused by the homogenisation might also make cells more permissible.

Transfections using Lipofectamine 2000, a cationic lipid, were carried out as follows: typically, DNA (4-6 µg of oligonucleotide or 1-4 µg of plasmid) and Lipofectamine 2000 at a w/v ratio of 1:2 were diluted in 250 µl Opti-MEM each in separate microfuge tubes. After 5 min incubation the solutions were mixed together gently and incubated at room temperature for 20 min to allow complex formation. During the incubation step, the normal growth medium was removed from the pre-plated cells and replaced with 1 ml of growth medium without antibiotics to avoid cytotoxicity associated with concomitant up-take of antibiotics.

For reverse transfection, cells were trypsinised as usual and an aliquot was taken for counting while cells were pelleted by 5 min centrifugation at 500 *g*. Cells were then resuspended in culture medium without antibiotics at a concentration of $1x10^6$/ml and 1 ml was added per well of a 6-well plate. Then 500 µl of the DNA-Lipofectamine 2000 solution was added drop-wise to the cells in each well and the cultures were incubated at 37 °C. The number of cells that were transfected and incubation times are detailed in Table 2-11. After incubation, cells were washed with 2 ml pre-warmed PBS and fresh culture medium containing antibiotics was added.

Table 2-11 Cell numbers and incubation times for transfection

Cell line	Number of cells per 6-well		Incubation time
	pre-plated	reverse transfected	
CHO-AI	$2x10^5$		2 h
CHO^{dhfr-}	$2x10^5$		2 h
CHO-AIM	$2x10^5$		2 h
CHO-K1	$2x10^5$		2 h
CHO-mEGFP	$2x10^5$		2 h
HEK293T		$5x10^5$	overnight
HEK293T-mEGFP		$5x10^5$	overnight
HepG2		$1x10^6$	overnight
HepG2-mEGFP		$1x10^6$	overnight

Transfections using jetPEI were carried out as follows: typically, 4-12 µg of oligonucleotide and jetPEI at a N/P ratio of 5:1 were each diluted in 100 µl NaCl. After gently mixing each separately, jetPEI was added to ssODN and the two solutions were mixed together. Following 15 min incubation at room temperature the combined transfection complexes were added to the cells.

For transfection with CDAN/DOPE, a cationic lipid similar to Lipofectamine 2000, but supposedly less cytotoxic, the reagent was diluted in 250 µl PBS and incubated 5 min, before mixing with the ssODN in volume to weight ratios of 1:1 and 2:1. The ssODN had also been diluted in 250 µl PBS. The solution was incubated 5-10 min at room temperature and then added to the wells. All transfected wells were washed into normal medium 3 h later.

2.2.1.5 Cell synchronization

CHO-mEGFP cells were synchronized using 48 h serum starvation and 100 µM mimosine as described in Olsen *et al.*[242] Cells were seeded at a density of $1x10^5$ cells per 6-well and cultured in full growth medium containing 5% FBS for 24 h. Then the medium was changed to reduced serum medium (0.2% FBS) and cells were cultivated for a further 48 h. Subsequently, the medium was changed back to full growth medium and 100 µM mimosine was added. Cells were transfected 24 h later.

2.2.1.6 Fluorescence resonance energy transfer (FRET)

Fluorescence energy transfer (FRET) occurs when one fluorochrome (the donor) excites another fluorochrome (the acceptor) through direct resonance energy transfer. This can only happen if the emission spectrum of the donor overlaps significantly with the excitation spectrum of the acceptor, both partners are in close physical proximity (<10 nm or 100 Å) and the emission and acception dipoles are not perpendicular to each other. In conjunction with the ability to tag proteins of interest with fluorescent proteins such as EGFP or RFP, FRET allows determination of the interactions of two tagged proteins in living cells with a higher certainty than can be reached by microscopic colocalization alone.

2.2.1.7 Loading cells with Cy3-dCTP

When using electroporation, $1x10^6$ cells in 100 µl electroporation buffer containing 2 µl of Cy3-dCTP were subjected to 2 pulses at 250 V for 0.4 s, and 6 pulses at 125 V for 0.075 s, with 0.5 s intervals. Following electroporation, 200 µl of complete medium was added directly into the cuvette and, after thorough mixing, the cell suspension was transferred into a 6-well containing 2 ml growth medium.

CHO-mEGFP cells were also loaded with Cy3-dCTP using the Influx Pinocytic Cell-loading reagent from Invitrogen. This reagent utilizes a combination of hypertonic and hypotonic extra-cellular conditions to induce cells to take up small molecules. Hypertonic extra-cellular conditions create a salt concentration gradient that results in water efflux from cells due to osmotic pressure, whereas hypotonic extra-cellular conditions result in influx of water to the cells.

Hypertonic loading medium was prepared by melting the waxy substance in one of the supplied tubes at 80 °C and adding 4.7 ml of Ham F12 medium without serum while vortexing for 5 min. Then 250 µl serum and 100 µl of 0.5 M HEPES buffer [pH 7.2] were added, the solution was vortexed again and filter sterilized. For the hypotonic lysis solution, Ham F12 medium was mixed with de-ionized water at a 6:4 ratio.

Per sample, $1x10^6$ cells were pelleted by centrifugation for 1 min at 300 *g*. Cells were resuspended in 20 µl hypertonic loading medium containing 2 µl of Cy3-dCTP and incubated 10 min at 37 °C. Then 1 ml of hypotonic lysis medium was added and 0.5 ml of the suspension were immediately transferred into 2 ml microfuge tubes containing 1 ml hypotonic lysis medium. After incubation for 1.5 min at 37 °C, cells were pelleted by centrifugation for 1 min at 300 *g*. The supernatant was removed and cells were resuspended in 1.5 ml of full growth medium before plating in 6-well plates.

2.2.1.8 Flow cytometry

Principles of flow cytometry

Flow cytometry is a technique for analysing and sorting microscopic particles (e.g. mammalian cells) in a hydro-dynamically focused stream of fluid. A single wavelength beam of light, usually from a laser, is aimed at the stream and excites fluorescent chemicals or proteins contained or attached to the particles passing in the stream, provided the wavelength matches the excitation spectrum of the fluorophore. The excited fluorophore then emits light at a longer wavelength which is measured by detectors aimed at the point where the stream passes through the light beam.

Typically, one detector is in line with the light beam (Forward Scatter or FSC) and several are perpendicular to it to measure Side Scatter (SSC) and the emission wavelengths of one or more fluorophores. Forward Scatter correlates with cell volumes and SSC depends on the inner complexity of the particle (i.e. shape of the nucleus, the amount and type of cytoplasmic granules or the membrane roughness). Flow cytometers can analyse several thousand particles per second and quantify the intensity of fluorescent emission from each particle accurately.

The resulting data can then be plotted either as a single dimensional (histogram) or two dimensionsal plot. Often, two dimensional plots are modified to represent a certain number of cells with the same fluorescent intensities by defined colours (density or contour plots). Selected subsets of events can be analysed separately by creating "gates" around the population of interest.

Generally, FSC vs. SSC plots and DNA cycle plots have linear scales as the signal differences are small, whereas plots showing fluorescent intensities are logarithmic as differences can span several magnitudes. Flow cytometers also employ complex arrangements of di-chroic splitting and band pass filters to isolate the emission spectra of different fluorophores. Excitation wavelengths and band pass filters for the dyes used here are summarized in Table 2-12.

Nevertheless, the emission spectra of different fluorescent dyes can overlap, so signals at the detectors have to be compensated. This can be done electronically by changing acquisition settings on the flow cytometer or computationally after data has been aquired. The latter is preferred as it avoids operator prejudice.

Table 2-12 Excitation wavelengths and band pass filters (in nm)

Excitation	Filter	Dyes
350	485/30	Hoechst 33342
488	530/30	EGFP, Alexa Fluor 488
	585/42	PI, RFP
633	660/16	Alexa Fluor 647

General flow cytometry

Flow cytometric analysis was performed on either a FACScan equipped with a 488 nm argon laser or a FACSCalibur with a 488 nm argon laser and a 633 nm red helium-neon laser. Cells were trypsinized, washed once with PBS and resuspended in at least 200 µl PBS at an approximate concentration of $1x10^6$ cells/ml.

Before analysis, 2 µg/ml propidium iodide (PI) was added for dead cell discrimination unless otherwise stated. Data were acquired with CellQuest 3.1 (FACScan) or CellQuest Pro (FACSCalibur) software without compensation and analysed using WinMDI 2.8 or WinList 6.0 software. Cells were always gated based on FSC versus SSC as well as PI to exclude cell debris and dead cells from the analysis. Where necessary, compensation was applied using the automatic compensation function of WinList 6.0.

Cell cycle analysis

Cell cycle studies were performed on a Partec PAS flow cytometer equipped with a 488 nm argon laser and a HBO lamp for UV excitation. Cells were washed once in PBS and stained with 5 µg/ml Hoechst 33342 for 30 min at 37 °C before analysis. Typically, $1x10^4$ events were acquired and flow rate was kept below 1000 events per second to minimize the coefficient of variation of the DNA histogram peaks.

Intracellular staining

Cells were trypsinized, washed in cold PBS and re-suspended in PBS at a concentration of $1x10^6$/100 µl. An equal volume of Caltag Fix&Perm solution A was added and cells were incubated for 15 min. Cells were washed once with PBS containing 5% normal goat serum and re-suspended in 100 µl of Caltag Fix&Perm solution B.

Per sample 250 ng of rabbit polyclonal anti-GFP antibody (or 0.5 µl rabbit polyclonal anti-human C reactive protein antibody) labelled with the Zenon Alexa Fluor 647 Rabbit IgG kit was added. Cells were incubated in the dark for 20 min, washed with PBS containing 5% normal goat serum and analyzed immediately. Data were acquired without compensation on a FACSCalibur flow cytometer. CHO-K1, HEK293T and HepG2 cells were used as negative controls to determine expression levels in CHO-, HEK293T- and HepG2-mEGFP cells. Typically, $1\text{-}8x10^6$ events were acquired.

2.2.1.9 Fluorescence-activated cell sorting (FACS)

Fluorescence-activated cell sorting (FACS) is the application of flow cytometry to physically separate cells based on specific light scattering and fluorescent characteristics. As with flow cytometry, cells are hydro-dynamically focused in a fluid stream, but this stream is then broken up into individual droplets by vibration. Excitation of fluorophores and analysis of resulting emission are also done the same way. However, cell sorters can apply electrical charges to individual droplets containing cells of interest. The charged droplets can then be directed into separate collection containers based upon their charge.

All cell sorting was performed at the Flow Core laboratory of the Institute for Cancer Research. Single green fluorescent gene-edited cells were isolated using a FACSAria with DiVa software and a 488 nm laser. Cells were gated based on FSC/SSC, Live/Dead (DAPI) and doublet discrimination (fluorescent signal area/width). Cells were cultured for at least four weeks before further analysis.

2.2.1.10 Production of HepG2-mEGFP cells

Attempts to isolate HepG2-mEGFP cell clones grown in DMEM or conditioned DMEM containing 20% FBS had failed, but following a recommendation by Dr. Demetra Mavri, HepG2 cells were grown and selected in Super HepG2 medium. Cells were transfected with 2 µg pmEGFP using Lipofectamine 2000 as described in section 2.2.1.4. Fluorescence microscope examination of a control transfection with pEGFP showed that the transfection efficiency was around 50%.

Antibiotic selection pressure was applied 2 d after transfection by addition of 500 µg/ml G418 and cells were plated in 6 cm plates at varying densities. In a mock-transfected control sample treated the same way, complete cell death was reached 10 d after the start of incubation. At this time, over 100 colonies were picked from the transfected cell sample and transferred into the wells of a 96-well plate. Selection pressure was maintained for a further 3 weeks until the clones had proliferated enough to grow in flasks. Clones were tested for stable integration of the transgene by transfection with 4 µg of the editing oligonucleotide mEGFP-3A-NT-PTO and flow cytometric quantification of the resulting green fluorescence cells.

2.2.1.11 Fluorescence microscopy

Fluorescence was examined on an inverted fluorescence microscope (Nikon Eclipse TE2000) equipped with a mercury arc lamp using the filter sets detailed in Table 2-13. In EGFP editing experiments, autofluorescence was controlled for by visual examination of cells transfected with control oligonucleotides. For other experiments, mock-transfected control samples were used. Phase contrast and fluorescence pictures of the same field were acquired and, if required, overlayed with Lucia G software.

Table 2-13 Fluorescence microscope filters

Excitation	Emission	Dyes
480/20	510/20	EGFP
546/10	580/30	RFP
540/25	605/55	Cy3
625/40	700/40	Cy5

2.2.1.12 Confocal microscopy

CHO-mEGFP cells were grown on cover-slides and transfected with Cy3-mEGFP-2A-PTO using Lipofectamine 2000. Cell nuclei were counterstained by adding 5 μM Draq5 to culture medium and incubating for 10 min at 37 °C. The cover slides were briefly rinsed in PBS and mounted onto slides using Citifluor. Confocal pictures were taken with a Zeiss LSM 510 Meta system.

2.2.2 Molecular Biology

2.2.2.1 Extraction and purification of DNA from mammalian cells

Cells ($1x10^6$ for small scale or $2x10^7$ for large scale extraction) were trypsinized, pelleted by 5 min centrifugation at 500 *g* and washed once in PBS. The supernatant was taken off and DNA extraction was either performed directly or cell pellets were stored at -20 °C.

For small scale extraction, cell pellets were resuspended in 200 µl PBS and extraction was performed using the DNeasy tissue kit according to the manufacturer's protocol. This kit uses spin columns for 1.5 ml microfuge tubes that contain silica membranes. DNA binds to silica in the presence of high chaotropic salt concentrations, but can be eluted easily by low salt conditions.

To digest proteins and thus protect DNA from nucleases during cell lysis, 20 µl proteinase K (600 mAU/ml) were added to the suspension. Contaminating RNA was removed by incubation with 4 µl RNase A (100mg/ml) for 2 min before 200 µl lysis buffer AL (containing guanidine hydrochloride) was added and samples were incubated for 10 min at 56 °C. After addition of 200 µl ethanol (96-100%) to precipitate the DNA, the samples were vortexed vigorously, pipetted into a spin column and centrifuged for 1 min at >6000 *g*. Samples were washed sequentially with 500 µl AW1 and AW2 to remove contaminating proteins. Flow-through and collection tubes were discarded after each wash. After the last wash, samples were centrifuged for 3 min at 20,000 *g* to remove residual ethanol from the membrane. Genomic DNA was eluted from the purification column with 200 µl 10 mM Tris-HCl [pH 8.5].

Large scale DNA extraction was performed using Qiagen 100/G Genomic tips according to the manufacturer's protocol. This kit uses anion-exchange columns that bind DNA under low-salt and pH conditions. Genomic DNA is eluted in high-salt buffer and then concentrated and desalted by isopropanol precipitation.

Pellets of $2x10^7$ cells were resuspended in 2 ml PBS, mixed with 2 ml ice-cold buffer C1 and 6 ml ice-cold distilled water and then incubated 10 min on ice to lyse cells, but preserve nuclei. After centrifugation for 15 min at 1300 *g* and 4 °C, the supernatant was discarded and samples were washed with 1 ml buffer C1 and 3 ml distilled water to remove all contaminating cell debris. The pellet was then resuspended in 5 ml buffer G2 to lyse the nuclei and denature proteins, 95 µl proteinase K was added and the suspension was incubated for 45 min at 56 °C to digest the denatured proteins.

Meanwhile, the required number of Genomic tip 100/G columns was equilibrated with 4 ml of buffer QBT. The samples were added to the columns and allowed to enter the resin by gravity flow. Columns were then washed twice with 7.5 ml buffer QC and genomic DNA was eluted in 5 ml buffer QF. DNA was precipitated with 3.5 ml isopropanol, recovered with glass hooks and transferred into 600 µl 10 mM Tris-HCl [pH 8.5]. To dissolve the genomic DNA, samples were placed on a rotational shaker overnight.

2.2.2.2 Determination of DNA concentration

The concentration of DNA solutions was determined by measuring the absorbance at 260 nm (A_{260}) with a Nanodrop ND-1000 spectrophotometer and applying Beer's law. The measurements were blanked with the solution buffer and samples were mixed before 1 µl was pipetted onto the light path. The ratio of sample absorbance at 260 nm and 280 nm was used to assess the purity of DNA solutions (optimal 1.8 for DNA). The nucleic acid concentration was automatically given as ng/µl, but is derived from the following formula:

$$Conc.\ (ng/\mu l) = (A_{260} \times extinction\ coefficient\ for\ DNA\ in\ ng\text{-}cm/\mu l) / path\ length\ in\ cm$$

The extinction coefficient for DNA is 50 ng-cm/µl and the path length is 0.1 cm on a Nanodrop.

2.2.2.3 Polymerase chain reaction (PCR)

Principles of PCR

A polymerase chain reaction is a process by which a section of DNA can be amplified in a sequence-specific manner. The selected region of a genome can be amplified a billion-fold with the use of short oligonucleotide primers of 15-30 bases in length, which are complementary to the target sequence. The primers lie on opposite strands of DNA and flank the region to be amplified.

Each cycle of the reaction requires a heat denaturation step to separate the double-stranded DNA and this is achieved by heating the target DNA to 94 °C or more. To allow the oligonucleotide primers to anneal specifically to their complementary sequences the reaction is then cooled to the annealing temperature which is typically chosen 5 °C below the melting temperature (T_M) of the primers. The reaction is then heated to the optimal working temperature for the thermo-stable DNA polymerase (usually 72 °C) and the primers are extended with the complementary sequence as template.

Thus, the extension reaction creates two double-stranded hybrids between a strand of target DNA and a shorter PCR product that begins with a primer sequence. After a second denaturation step, the strands starting with a primer sequence can function as templates for annealing of the other primer. Hybridization and extension then produces double-stranded molecules that comprise the stretch of target region flanked by the two primers.

Repeated cycles of heat denaturation, primer annealing and extension exponentially amplify the target fragment of DNA. After 20-40 such cycles a final 'polishing off' extension step at 72 °C for 5 min completes the PCR.

Generally, PCR is performed in 20–100 µl reaction volumes which contain the following components: a reaction buffer providing optimised conditions for the DNA polymerase, 1-4 mM $MgCl_2$ (low concentrations yield less product but higher concentrations increase non-specific product), 200 µM dNTPs (an equimolar mixture of dATP, dCTP, dGTP and dTTP), 0.1–1 µM primers, 1.5 U/50 µl DNA polymerase and 0.01-1 ng plasmid DNA or 0.1-1 µg genomic DNA template. Reactions are made up to required volumes with nuclease-free water. Adjuvants such as 5-10% DMSO may be added to the reaction of GC-rich templates as it facilitates strand separation by disrupting base pairing and thus increases efficiency.

Primer3, a primer design tool (*http://frodo.wi.mit.edu*) was used to design primers. This tool provides essential information such as melting temperature (T_m), and potential secondary structure and primer-dimer formation. Melting temperatures are dependent on GC content as well as length of the primers and should be similar within a pair. Secondary structure and formation of primer-dimers will reduce the quantity of PCR product. Primers should be around 18-22 bases long, though in some cases they might be longer to accommodate non-complementary nucleotides.

Thermo-stable polymerases are purified from thermophilic bacteria (for example Taq polymerase from *Thermus aquaticus*) or archaea (Pfu from *Pyrococcus furiosus*) and remain stable at high temperatures. Unlike Taq polymerase, Pfu possesses 3'-5' exonuclease proof-reading activity.

All PCRs were pipetted on ice, where possible as master mixes, and sample tubes were directly transferred to a Bio-Rad iCycler already at 94 °C. This machine has a heated lid to prevent loss of sample through evaporation upon heating. A PCR control sample without DNA template was always included and only PCRs with a clean control, as determined by gel electrophoresis, were used for subsequent analysis. In order to minimise contamination, gloves were worn at all times, filter-protected pipette tips were used and the reactions were prepared in a designated PCR room.

Table 2-14 ApoA-I/ApoA-I_M with primers AI2F and AI4R

Reaction component	
10x PCR buffer	1x
$MgCl_2$	2 mM
dNTPs	400 µM
AI2F primer	0.5 µM
AI4R primer	0.5 µM
Taq DNA polymerase	1.5 U
Genomic DNA	5 µl
Total volume	50 µl

Temperature	Time	Cycles
95 °C	5 min	
94 °C	1 min	35x
62 °C	30 s	
72 °C	30 s	
72 °C	5 min	
4 °C	∞	

Table 2-15 ApoA-I/ApoA-I_M with primers ApoAIinnerF and ApoAIinnerR2

Reaction component	
10x PCR buffer	1x
$MgCl_2$	2 mM
DMSO	5% (v/v)
dNTPs	200 µM
ApoAIinnerF primer	0.3 µM
ApoAIinnerR2 primer	0.3 µM
Taq DNA polymerase	3 U
Genomic DNA	5 µl
Total volume	30 µl

Temperature	Time	Cycles
95 °C	5 min	
95 °C	1 min	35x
68 °C	30 s	
72 °C	30 s	
72 °C	5 min	
4 °C	∞	

Table 2-16 Southern blot probes from pmEGFP

Reaction component	
10x PCR buffer	1x
$MgCl_2$	2 mM
dNTPs	200 µM
Forward primer	0.5 µM
Reverse primer	0.5 µM
Taq DNA polymerase	6 U
pmEGFP DNA	1 ng
Total volume	300 µl

Temperature	Time	Cycles
95 °C	5 min	
95 °C	1 min	30x
58 °C	30 s	
72 °C	1 min	
72 °C	5 min	
4 °C	∞	

Table 2-17 Amplification of EGFP alleles from the genomic DNA of edited clones

Reaction component	
10x PCR buffer	1x
$MgCl_2$	2.5 mM
dNTPs	200 µM
pEGFP-241for	0.5 µM
pEGFP-1380rev	0.5 µM
Pfu Turbo DNA polymerase	2 U
Genomic DNA	100 ng
Total volume	300 µl

Temperature	Time	Cycles
95 °C	5 min	
95 °C	1 min	15x
58 °C	1 min	
72 °C	1 min	
72 °C	5 min	
4 °C	∞	

The reaction was diluted 1:200 and then used as template in a second amplification step.

Reaction component	
10x PCR buffer	1x
$MgCl_2$	2.5 mM
dNTPs	200 µM
pEGFP-241for	0.5 µM
pEGFP-1380rev	0.5 µM
Pfu Turbo DNA polymerase	12 U
Diluted PCR product	10 µl
Total volume	300 µl

Temperature	Time	Cycles
95 °C	5 min	
95 °C	1 min	25x
58 °C	1 min	
72 °C	1 min	
72 °C	5 min	
4 °C	∞	

Table 2-18 Amplification of EGFP alleles from plasmid DNA

Reaction component	
10x PCR buffer	1x
$MgCl_2$	2 mM
dNTPs	200 µM
pEGFP-241for	0.5 µM
pEGFP-1380rev	0.5 µM
Taq DNA polymerase	2 U
Pooled denatured bacteria	2 µl
Total volume	20 µl

Temperature	Time	Cycles
95 °C	5 min	
95 °C	1 min	20x
58 °C	1 min	
72 °C	1 min	
72 °C	5 min	
4 °C	∞	

2.2.2.4 Restriction endonuclease digestion of DNA

Restriction endonucleases are enzymes purified from bacteria that recognise and cleave specific and short nucleotide sequences within double-stranded DNA. Restriction sites typically comprise palindromes of 4, 5, 6 or more bp with an axis of rotational symmetry. Cleavage by a restriction enzyme produces either cohesive (having either a 5′ or 3′ single-stranded protrusion) or blunt-ended (no single-stranded protrusion) fragments. Usually, these ends have 5'-phosphate and 3'-hydroxyl ends, although exceptions are known. Cohesive fragments can be subsequently ligated to each other if their single-stranded protrusions or 'overhangs' are compatible. All blunt-ended fragments can be ligated to each other.

Restriction endonuclease reactions contain optimal buffer conditions for the specific enzyme (varying concentrations of Tris-HCl [pH 7.5], NaCl, dithiothreitol and $MgCl_2$), 100 µg/ml bovine serum albumin (BSA) if required by the enzyme, the target DNA, nuclease-free water and the enzyme itself. One unit of enzyme is defined as the amount needed to digest 1 µg of λ-DNA per hour, but for optimal use the manufacturer often recommends 10x or 20x over-digestion. However, enzymes are supplied in 50% glycerol and the concentration of glycerol in the reaction should not exceed 5% as higher glycerol concentrations can inhibit enzyme activity or induce "star" activity (relaxed specificity). The reagents were mixed and the reaction was incubated at 37 °C for an hour or where indicated overnight.

The enzymes used in this study and their required conditions are detailed in Table 2-19. When DNA had to be digested with two restriction enzymes, a double digest was performed and the buffer that resulted in the most activity for both enzymes was selected. For Southern blot, samples were digested with the amounts of enzyme recommended by the manufacturer, taking into account the suggested over-digestion. Total digestion volumes were adjusted with H_2O to keep glycerol concentrations below 5%. Reactions were incubated overnight at 37 °C with 5% (v/v) enzyme and another 5% was added for 2 h in the morning. Reactions were stopped by heat inactivation for 20 min at 65 °C.

Table 2-19 Restriction endonuclease cleavage sites and reaction conditions

Restriction enzyme	Cleavage site	Reaction conditions for optimal activity
HaeII (20 U/µl)	R GCGC/Y* Y/CGCG R	NEBuffer 4 at 37 °C, + BSA; 10x over-digestion
NdeI (20 U/µl)	CA/TA TG GT AT/AC	NEBuffer 4 (or Buffer 2) at 37 °C; 20x over-digestion
MfeI (10 U/µl)	C/AATT G G TTAA/C	NEBuffer 4 at 37 °C; 20x over-digestion
DraI (20 U/µl)	TTT/AAA AAA/TTT	NEBuffer 4 at 37 °C; 20x over-digestion
XbaI (20 U/µl)	T/CTAG A A GATC/T	NEBuffer 2 at 37 °C, + BSA; 20x over-digestion

*R = A or G; Y = C or T

2.2.2.5 Polymerase chain reaction restriction fragment length polymorphism

Polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP) analysis was employed in this study to detect point mutations introduced by oligonucleotide-mediated gene editing The nucleotide alterations destroy the HaeII restriction site previously present in the wild-type sequence and thus result in the presence of undigested DNA fragments which can be visualized with gel electrophoresis.

PCR products gained as explained in section 2.2.2.3 were digested with 5% (v/v) of restriction endonuclease HaeII as described above. Reactions were incubated overnight at 37 °C and another 5% (v/v) of enzyme was added for 2 h in the morning. Results are representative of at least two independent experiments, each analyzed at least twice to ensure that PCR-RFLP bands did not arise from contamination during PCR.

2.2.2.6 Concentrating DNA samples

The large sample volumes resulting from digestion of genomic DNA for Southern Blot had to be concentrated in order to fit into the wells of the agarose gel. This was accomplished by pipetting the samples into Microcon YM-100 centrifugal filter units seated in 1.5 ml microfuge tubes followed by centrifugation for 30-45 min at 500 *g*. When the sample volumes were reduced to approximately 30 µl, the filter units were placed upside down in new tubes and centrifuged for 3 min at 1000 *g* to recover the concentrated samples.

2.2.2.7 Agarose gel electrophoresis

Generally, PCR products or restriction endonuclease digests were analysed using TBE-buffered agarose (1%) gel electrophoresis to facilitate quick resolution of 500 bp-3 kb fragments. Double-stranded DNA molecules migrate through the gel matrix from the cathode to the anode and separate according to their electrophoretic mobility which is inversely proportional to the $\log_{10}$ of the number of base pairs. The length of the DNA molecules can be ascertained by comparison to molecular weight standards of known molecule sizes.

The percentage of agarose in gels may be adjusted to accommodate optimal resolution of fragments larger than 10 kb (0.5%) or smaller than 400 bp (2%). Ethidium bromide, when added to the gel, intercalates between adjacent base pairs of DNA and allows visualisation of distinct DNA bands when viewed under ultraviolet (UV) light.

A minigel electrophoresis apparatus was set up as recommended by the manufacturer. For a 1% agarose gel, 0.5 g of agarose was dissolved in 50 ml of 1x TBE buffer by heating in a microwave. Ethidium bromide was added at a concentration of 0.5 µg/ml, the gel solution was poured into the gel cassette and a comb with the required number and size of wells was added. When the gel had set, the comb was removed and a sufficient volume of 1x TBE buffer was added to cover the surface of the gel. The DNA samples were mixed 9:1 with 10x Bluejuice Gel Loading Buffer and 10-20 µl of sample, or 1 µg DNA ladder, were loaded per well. Gels were run at 15 V/cm until the bromophenol blue front had migrated a sufficient distance and visualised on a UV trans-illuminator connected to a computer-operated camera with image analysis software.

MetaPhor high resolution agarose was used when DNA fragments of less than 200 bp needed to be extracted from the gel. Gels were prepared according to the manufacturer's instructions. For Southern blotting, gels were prepared as described above with 1x TAE buffer substituted for 1x TBE and with an electrical field strength of 3.5 V/cm instead of 15 V/cm.

2.2.2.8 Polyacrylamide gel electrophoresis

Pre-cast 20% polyacrylamide TBE gels were used for the routine detection of PCR and restriction endonuclease digestion products of less than 250 bp as these gels offer particularly good resolution of small DNA fragments. Gels were inserted into the XCell SureLock Mini-Cell gel electrophoresis apparatus and 200 ml of 1x TBE running buffer was added to the upper reservoir and 600 ml to the lower. DNA samples were mixed with 5x Novex Hi-Density Sample Buffer (final concentration 1x) and loaded into the respective wells along with an appropriate DNA ladder.

Gels were run at 200 V for 45-90 min depending on the size of the expected bands, removed and stained with 0.5 µg/ml ethidium bromide for 10 min on a shaker. DNA bands were visualised on a UV transilluminator connected to a computer-operated camera with image analysis software.

2.2.2.9 Extraction and purification of DNA from gels or enzymatic reactions

The QIAquick Gel Extraction Kit was used for the purification of DNA from agarose gel or directly from a PCR and restriction enzyme digest reaction. Gels intended for extraction of DNA bands were poured and run without ethidium bromide. Bands were cut out with a scalpel after a short incubation in ethidium bromide solution and exposure to UV light was kept to a minimum to avoid DNA degradation.

The gel pieces were weighed, buffer QG was added to the gel slice at a ratio of 3:1 (e.g. 300 µl QG to 100 mg gel) and the mixture placed at 50 °C to dissolve the gel. The buffer contains the high concentration of chaotropic salt necessary for binding of DNA to a silica-membrane and also a pH indicator as binding only occurs below pH 7.5. One gel volume of room temperature isopropanol was then added to the solution. If DNA was purified directly from a PCR or restriction enzyme digest, an additional 3 volumes of QG solution and 1 volume of isopropanol were added to the reaction mix.

The solution was then transferred to a QIAquick spin column and centrifuged at 16,000 *g* for 1 min to allow binding of DNA to the silica-membrane. The flow-through was discarded and the column was washed with 0.75 ml of buffer PE to remove excess salt. The supernatant was removed and the column was centrifuged for an additional 2 min to remove any remaining traces of ethanol. The DNA was eluted with 30 µl 10 mM Tris-HCl, [pH 8.5].

2.2.2.10 Southern Blot (alkaline transfer)

Genomic DNA was extracted from cells as in section 2.2.2.1, quantified using a Nanodrop ND-1000 (section 2.2.2.2) and double digested with the restriction endonucleases NdeI and MfeI, NdeI and DraI, or NdeI and XbaI (section 2.2.2.4). The endonuclease digestion reactions were concentrated with Microcon YM-100 centrifugal filter units (section 2.2.2.6) and 15 µg per lane, as well as 2 µg of 100 bp DNA ladder, were loaded onto a 1.5% TAE-agarose gel. Gels were prepared and run as described in section 2.2.2.7 and photographed on the UV-transilluminator together with a fluorescent ruler. The ruler was necessary to establish the size of bands on the blot in comparison with the size of marker bands on the gel.

Extraneous parts of the gel were removed with a scalpel and the DNA on the gel was prepared for transfer by denaturation in alkaline transfer buffer with gentle agitation for 2x 15 min. As the expected fragment sizes were well under 10 kb depurination was deemed unnecessary. During the denaturation step the blotting apparatus for capillary transfer of the DNA to Hybond XL, a positively charged nylon membrane, was set up as detailed in Figure 2-2 with alkaline transfer buffer in the reservoir.

Air bubbles were avoided by gently rolling a sterile 5 ml pipette over each layer. Alkaline transfer was chosen because of increased sensitivity and because the membrane does not require baking or UV-crosslinking for DNA fixation. After overnight transfer, the apparatus was carefully taken apart and the position of wells, corners and DNA side was labelled with a pencil on the membrane. The membrane was then soaked in 2x SSC for 15 min and pre-hybridised in Rapid-hybridization buffer (100 µl/cm^2) at 65 °C for 30 min in a rotational hybridization oven.

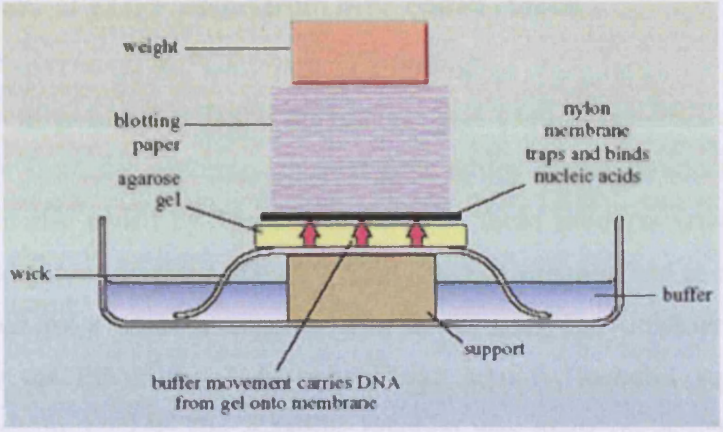


Figure 2-2 Blotting apparatus for Southern blot

Figure from the Open University, Milton Keynes, UK

Southern blot probes were prepared by PCR with primer pairs pEGFP-241for/pEGFP-754rev (Probe A) and pEGFP-844for/pEGFP-1380rev (Probe B) from pmEGFP plasmid (section 2.2.2.3) and purified as described in section 2.2.2.9. For each blot 50 ng of probe was diluted in 45 µl of nuclease-free water. The sample was denatured for 2 min at 100 °C, placed on ice for another 2 min and spun down.

Behind a β-radiation shield, the sample was added to a tube containing a Ready-to-go DNA random-prime labelling bead and 5 µl [α-^{32}P] dCTP (3000 Ci/mmol). The reaction was incubated 30 min at 37 °C and then stopped by the addition of 5 µl 0.2 M EDTA [pH 8.0]. Unincorporated nucleotides were removed by centrifugation through an Illustra ProbeQuant G-50 size exclusion microcolumn.

The probe was denatured again, as described above, and added to the pre-hybridization solution. After 2 h hybridization at 65 °C, unbound probe was removed by washing in SB wash buffer I at room temperature followed by two washes in SB wash buffer II at 65 °C. Pre-flashed Kodak Biomax-MS film was exposed to the membrane for 2 - 24 h at -70 °C. The size of bands on the blot was established by determining the distance marker bands had migrated from the wells on the photographed gel with the help of the fluorescent ruler.

2.2.2.11 Recovery of EGFP alleles from gene-edited clones

Addition of 3'-A overhangs to PCR product (A-tailing)

Because the proofreading Pfu Turbo polymerase was used in the PCR amplifying the EGFP sequence from the genomic DNA of gene-edited clones, products lacked 3'-A overhangs ordinarily added by Taq polymerase. As these products were to be ligated into the T-tailed plasmid vector pGEM-T Easy, such overhangs had to be added. This can be achieved by a short incubation with dATP and Taq polymerase. To avoid interference by the Pfu Turbo 3'-5' exonuclease activity, samples were purified as described in section 2.2.2.9. The A-tailing reaction was made up as detailed in Table 2-20 and incubated 30 min at 37 °C. As an unspecific band was visible on a test agarose gel, samples were then purified as described in section 2.2.2.9.

Table 2-20 Components of the A-tailing reaction

PCR product	1350 µg
Taq DNA polymerase	25 U
PCR buffer	1x
dATP	200 µM
$MgCl_2$	1.5 mM
Total volume	50 µl

Ligation

To insert a desired stretch of DNA or PCR product into a vector, ATP-dependent formation of a phosphodiester bond between the 3'hydroxyl end of a double-stranded DNA fragment and the 5'phosphate end of the same or another DNA fragment is necessary. This reaction is catalyzed by T4 DNA ligase and can take place between fragments that possess complementary cohesive ends or blunt ends to produce circular recombinant molecules.

A molar ratio of 3:1 (insert:vector) is recommended for ligation of DNA fragments with complementary cohesive ends such as T-tailed vector and A-tailed PCR product. The following equation was used to determine the required amount of plasmid and insert for given fragment sizes:

(ng of vector × kb size of insert/ kb size of vector) × molar ratio (insert:vector) = ng of insert

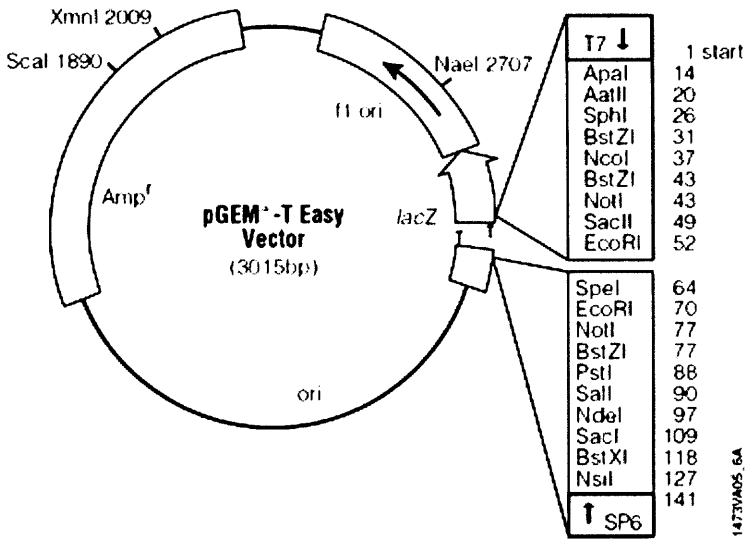


Figure 2-3 Vector map of pGEM-T Easy

The pGEM-T Easy vector is 3015 bp (Figure 2-3) and the EGFP PCR product 1139 bp long. Thus, 50 ng EGFP PCR product and 50 ng pGEM-T Easy have a 3:1 molar ratio. The reaction components detailed in Table 2-21 were mixed and incubated overnight at room temperature. A control reaction mix containing every component apart from the insert DNA was also prepared to assess the background level of re-circularised plasmid.

Table 2-21 Components of the ligation

PCR product	50 ng
pGEM-T Easy vector	50 ng
Rapid ligation buffer	1x
T4 DNA Ligase	1.5 mM
Total volume	10 µl

Transformation of plasmid DNA into competent E. coli

After ligating the DNA of interest into a vector, the resulting plasmid is inserted into competent bacterial cells for easy amplification. Typically, competent bacteria are made permeable to DNA by incubation in ice-cold $CaCl_2$ solution and are kept frozen at -80 °C as they are extremely fragile in this state. Thus, the DH5α chemically competent *E.coli* were thawed on ice and 5 µl of the ligation reaction were gently added to 100 µl of bacteria.

After incubation on ice for 30 min to permit diffusion of the plasmid into the cells, samples were heat-shocked for 90 sec in a 42 °C water bath to reverse the permeabilization. The reactions were placed on ice for 2 min, then 900 µl S.O.C medium was added and samples were incubated for 1 h at 37 °C in a shaking incubator to allow time for expression of the antibiotic resistant gene encoded by the plasmid.

Usually, 50 µl of the transformation mix is then spread onto LB agar plates containing 100 µg/ml ampicillin. The ampicillin exerts selection pressure, so only bacteria that contain vectors carrying an ampicillin resistance gene survive. However, self-ligated vectors also contain a functional ampicillin resistance gene. To distinguish further between plasmids carrying the intended insert or self-ligated vectors, α-complementation, also known as blue/white screening, was employed.

α-Complementation (blue/white screening)

The pGEM-T Easy vector contains a multiple cloning region within the α-peptide coding region for the enzyme β-galactosidase *(LacZ* gene). Inactivation of the α-peptide by the insert leads to formation of non-functional β-galactosidase enzyme. Thus, vectors without an insert express functional enzyme that converts substrates such as X-gal to a blue coloured product, whereas vectors with inserts express non-functional enzyme that cannot convert X-gal.

Here, 40 µl of X-gal (20 mg/ml) and 6 µl of IPTG (100 mM) were added directly to pre-made LB agar plates supplemented with 100 µg/ml ampicillin and spread over the entire surface. The plates were then incubated 30 min at 37 °C for the solution to be absorbed. Then the plates were inoculated with the transformed bacteria and placed in the 37 °C incubator overnight.

PCR screening

Single white colonies were used to inoculate the wells of a 96-well plate containing 100 µl LB medium supplemented with 100 µg/ml ampicillin. The plate was covered with a plastic film and incubated overnight. The next morning, 1 µl of cultures in the same row or column were pooled and diluted in 100 µl H_2O.

After heat denaturation for 10 min at 95 °C, the EGFP alleles contained in each colony were amplified by PCR with primers pEGFP-241for and pEGFP-1380rev (section 2.2.2.3). PCR products were digested with NdeI (section 2.2.2.4) and examined by electrophoresis. Correlation of positive row samples with positive column samples identified bacterial colonies carrying the edited EGFP allele (edEGFP).

Extraction and purification of plasmid DNA

DNA purification with the QIAquick spin miniprep kit is based on alkaline lysis of bacterial cells and subsequent binding of DNA to a silica membrane under high salt conditions. Colonies carrying the edEGFP or control mEGFP alleles were streaked out on LB agar plates supplemented with 100 μg/ml ampicillin and incubated overnight at 37 °C. Single colonies from each bacterial clone were picked and used to inoculate 5 ml LB medium containing 100 μg/ml ampicillin.

Cultures were grown overnight, but no more than 12-16 hours to prevent ampicillin depletion (hydrolysis) in the culture. Bacteria were pelleted by 10 min centrifugation at 5,000 *g* and resuspended in 250 μl ice-cold Buffer P1 containing RNase A. Samples were mixed vigorously until no cell clumps were visible and 250 μl Buffer P2 was added to denature proteins and to solubilize the phospholipids in the cell membrane. The tubes were inverted gently and incubated for exactly 5 min.

Chromosomal DNA and cell debris were precipitated with 350 μl Buffer N3, which also stops the lysis and adjusts the salt conditions. The solution was mixed thoroughly and centrifuged 10 min at 16,000 *g* before the supernatants were loaded onto the supplied spin columns. Columns were centrifuged for 30 sec and washed with 750 μl of Buffer PE. The flow-through was discarded and the columns were centrifuged for an additional 1 min to remove residual ethanol. Plasmid DNA was eluted with 50 μl 10 mM Tris-HCl, [pH 8.5].

2.2.2.12 Sequencing

Analysis of samples was carried out by the UCL Wolfson Institute for Biomedical Research sequencing service using a Beckman Coulter CEQ 8000 sequencer.

2.3 Picture editing

Gel images, scanned X-Ray films and microscope pictures were edited with Adobe Photoshop CS2. Typically, images were cropped to show only the area of interest and levels were adjusted equally across the whole image to optimize contrast between bands and background. No bands were added or removed.

2.4 Statistical analysis

Statistical analysis was performed using GraphPad Prism 5. One-way ANOVA was chosen when comparing more than two values, as it is more accurate than Student's t-test for multiple comparisons. Values in text, tables and figures were expressed as the mean±standard deviation (SD) and reflected either duplicates, triplicates or two independent experiments each with triplicates as stated in the text.

Repeat experiments were compared to each other using two-way ANOVA and if significant difference was found, both experiments were normalized to account for experimental variation by setting the highest gene editing efficiency to 100%. The resulting values were then analysed by one-way ANOVA. Two-way ANOVA was also used to analyse samples differing in two independent conditions, such as time to measurement and oligonucleotide backbone modification. Bonferroni post-tests were included in every ANOVA statistical analysis to determine p values.

Where stated, data were analysed using contingency tables. This method numerically states the results of an experiment in which the outcome is a categorical variable (results are grouped into mutually exclusive categories). Fisher's exact test rather than chi-square test was applied to each contingency table to determine P values as it gives the exact value.

Generally, a $P<0.05$ was considered to be significant. For easy visualization, P values are expressed as * ($P<0.05$), ** ($P<0.01$) and *** ($P<0.001$) in Tables and Figures.

CHAPTER 3: EVALUATION OF APOA-I GENE EDITING WITH PCR-RFLP

INTRODUCTION

As I discussed in the general introduction, ApoA-I_M is more atheroprotective than ApoA-I and gene therapy would be more cost efficient than administering recombinant ApoA-I_M/POPC complexes as a treatment. I also examined the advantages and disadvantages of the different types of gene therapy and argued that gene editing would be the ideal solution in this case, if problems with reproducibility of chimeraplast-mediated gene editing could be solved. Indeed, previous experiments in this laboratory demonstrated low and inconsistent levels of ApoA-I to ApoA-I_M gene conversion with chimeraplasts in HepG2 cells.[237] However, recent reports have shown that the DNA and not the RNA strand of chimeraplasts is the active domain which mediates gene editing,[239,240] suggesting that consistent and repeatable conversions might be achieved if single-stranded oligonucleotides (ssODNs) were used instead of chimeraplasts.

Closely following the conditions described in Manzano *et al.*[237] I evaluated the ability of ssODNs to induce the ApoA-I to ApoA-IM mutation in recombinant CHO-AI (see section 2.1.1.1) and HepG2 cells. Unfortunately, in my hands the primers used by Manzano *et al.*[237] caused unspecific bands that interfered with the analysis of the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP). Therefore, I designed new primers and optimised PCR conditions to avoid unspecific bands.

Using the newly optimised PCR-RFLP and sequencing of diagnostic bands I showed that oligonucleotide-mediated gene editing induces at least a temporary change in the genomic DNA of targeted CHO-AI and HepG2 cells. I further excluded the possibility that PCR-based artefacts gave rise to the diagnostic bands by various spiking experiments.

3.1 Optimisation of PCR-RFLP analysis for the detection of ApoA-I to ApoA-I_M gene conversion by ssODNs

3.1.1 Results

CHO-AI cells expressing human ApoA-I were transfected with 6 or 12 μg of huApoA-I_M-49T-PTO and huApoA-I_M-49NT-PTO oligonucleotide using LPEI as described in Section (2.2.1.4). Cells were harvested 24 h after transfection, genomic DNA was extracted and samples were analysed using PCR-RFLP.

Figure 3-1 explains the ApoA-I/ApoA-I_M PCR-RFLP in more detail. The 2385 bp genomic DNA of ApoA-I consists of four exons, which encode an 893 bp coding sequence generating a 243 amino acid long protein. PCR utilizing the primers AI2F and AI4R amplifies a 216 bp product from the ApoA-I coding sequence which has two HaeII restriction endonuclease recognition sites (5'-RGCGC/Y-3'). Digestion with HaeII yields three fragments of 166 bp, 41 bp and 9 bp length.

If the amplified PCR product has the C to T mutation associated with the ApoA-I_M genotype, the first of these HaeII recognition sites is changed to 5'-RGCGT/Y-3' and consequently cannot be cut by the enzyme; this results in fragments of 207 bp and 9 bp. However the 207 bp band could also arise from incomplete digestion and it is therefore important to have a pure ApoA-I sample (i.e. mock-transfected CHO-AI cells) as a digestion control and to sequence this band before successful conversion can be claimed. The Milano mutation changes an amino acid from Arginine (R) to Cysteine (C) at position 173, so this particular PCR-RFLP is called R173C PCR-RFLP.

In this experiment the 216 bp PCR fragment was amplified with the primers AI2F and AI4R (as described in Manzano *et al.*[237]) and an annealing temperature of 65 °C.

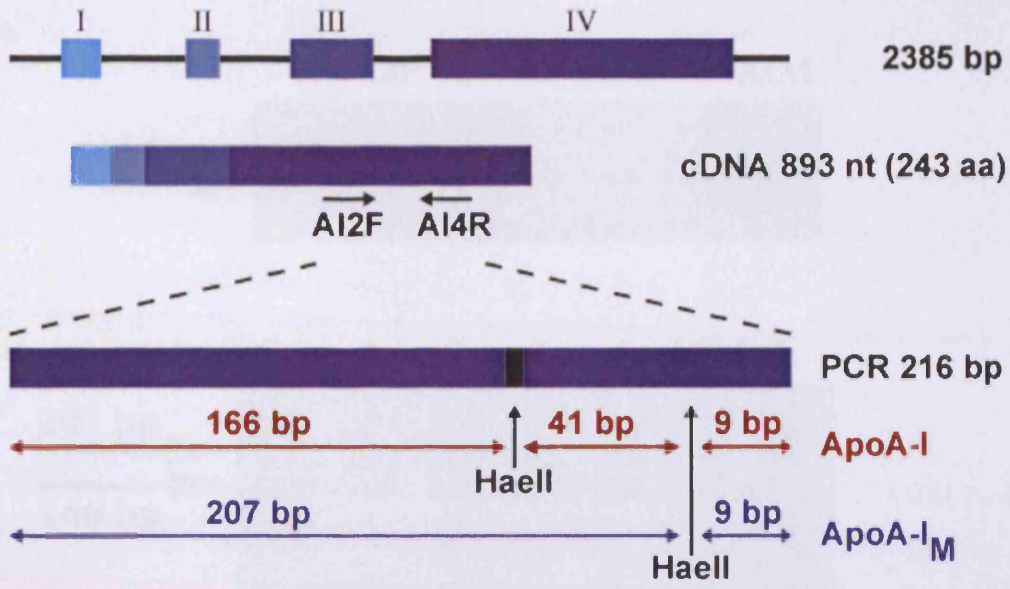


Figure 3-1 Schematic of R173C PCR-RFLP

The ApoA-I gene is represented above with boxes for its four exons. The primers AI2F and AI4R amplify a 216 bp product, which yields HaeII digestion fragments of 166, 41 and 9 bp for wild-type ApoA-I or 207 and 9 bp for ApoA-I_M.

Figure 3-2a shows that the resulting PCR product was 216 bp, exactly as expected from the position of the primer binding sites in the sequence (see Figure 3-3). The PCR products were then digested with the restriction endonuclease HaeII as described in Section (2.2.2.4) and the resulting fragments were resolved on a 20% TBE gel (Figure 3-2b)

The expected bands at 166 bp and 41 bp signifying the presence of the ApoA-I sequence in the genomic DNA were present in all samples except AIM and were easily identifiable. Identification of the 207 bp band was complicated by a strong unspecific band in close proximity, which was estimated to be ~195 bp. The 207 bp band denoting the presence of the Milano mutation could be seen in all samples except the mock-transfected cells and therefore this band was most likely not due to incomplete digestion. Thus, it seemed that successfully converted cells were present in the treated cells from which the genomic DNA samples were extracted.

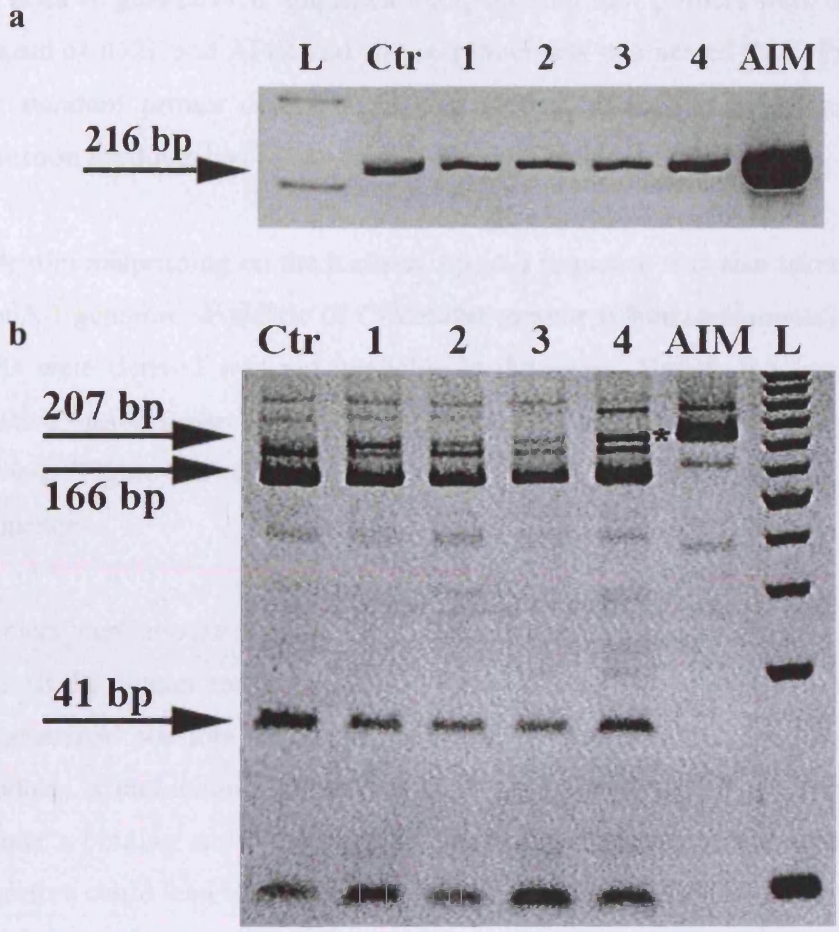


Figure 3-2 R173C PCR-RFLP with primers AI2F and AI4R

a) PCR product from genomic DNA of treated CHO-AI cells resolved on a 1% agarose gel. L: 100 bp ladder. Ctr: mock-transfected CHO-AI cells. Lanes 1-4: Cells transfected with 6 µg, 12 µg huApoA-I_M-49T-PTO, 6 µg, 12 µg huApoA-I_M-49NT-PTO. AIM: pcDNA3.1-ApoA-I_M.

b) 20% TBE gel resolving the HaeII restriction fragments. Ctr: mock-transfected CHO-AI cells. Lanes 1-4: Cells transfected with 6 µg, 12 µg huApoA-I_M-49T-PTO, 6 µg, 12 µg huApoA-I_M-49NT-PTO; AIM: pcDNA3.1-ApoA-I_M. L: 25 bp ladder. An asterisk marks the diagnostic 207 bp band in lanes 4 and AIM.

However, the presence of the unspecific bands cast doubt on the data obtained in this PCR-RFLP. To obtain clear results the unspecific bands had to be eliminated. An investigation into the causes of the unspecific bands revealed that both the AI2F and AI4R primer form self-annealing dimers with a ΔG = -11 kcal/mol as well as duplexes with a ΔG = -5 kcal/mol. Furthermore, their melting temperatures were elevated at 79°C due to a G/C content of 83% (section 2.1.2.4).

To achieve greater PCR amplification specificity new primers were designed to be used instead of AI2F and AI4R and also as primer sets in a nested PCR. Primers had to fulfil the standard primer design rules (see section 2.2.2.3) and ensuing PCR and HaeII digestion products had to result in distinct gel bands.

Potential mispriming on the hamster ApoA-I sequence was also taken into account. The ApoA-I genomic sequence of *Cricetulus griseus* (Chinese Hamster) from which CHO cells were derived was not available in databases. Hence, the sequence of the close relative *Mesocricetus auratus* (Syrian Hamster) was substituted in a comparison with *H. sapiens* (Figure 3-3) to assess the potential cross-species mispriming of selected primer sequences.

Primers were chosen with the help of the primer3 design program[285] and were screened against the human and rodent mispriming libraries available in this program. Table 3-1 summarizes possible combinations of PCR primers and resulting HaeII digestion products. Combinations including the primer ApoAIinnerR were disregarded as this primer's binding site overlaps with the target sequence of the correcting ssODN and therefore could lead to mispriming events during the PCR.

Primers ApoAIouterF2 and ApoAIouterR were selected to avoid inter-species mispriming; however, the HaeII digestion bands distinguishing ApoA-I from ApoA-I_M, 205 and 246 bp respectively would be difficult to separate on either agarose or TBE gels. On agarose gels, these fragment sizes are at the lower detection limit and a high amount of fragment would be required for visualisation; whereas on TBE gels the fragment sizes are at the upper detection limit and gels would need to be run for extended times to get good resolution, causing smears, distortions and brittleness in the gels.

Therefore, the ApoAIouterF2 and ApoAIouterR primers were intended as outer primer set in nested PCRs only. ApoAIinnerF and ApoAIinnerR2 were selected as inner primers for nested PCRs, because the expected HaeII restriction fragment sizes would be easily resolved on TBE gels.

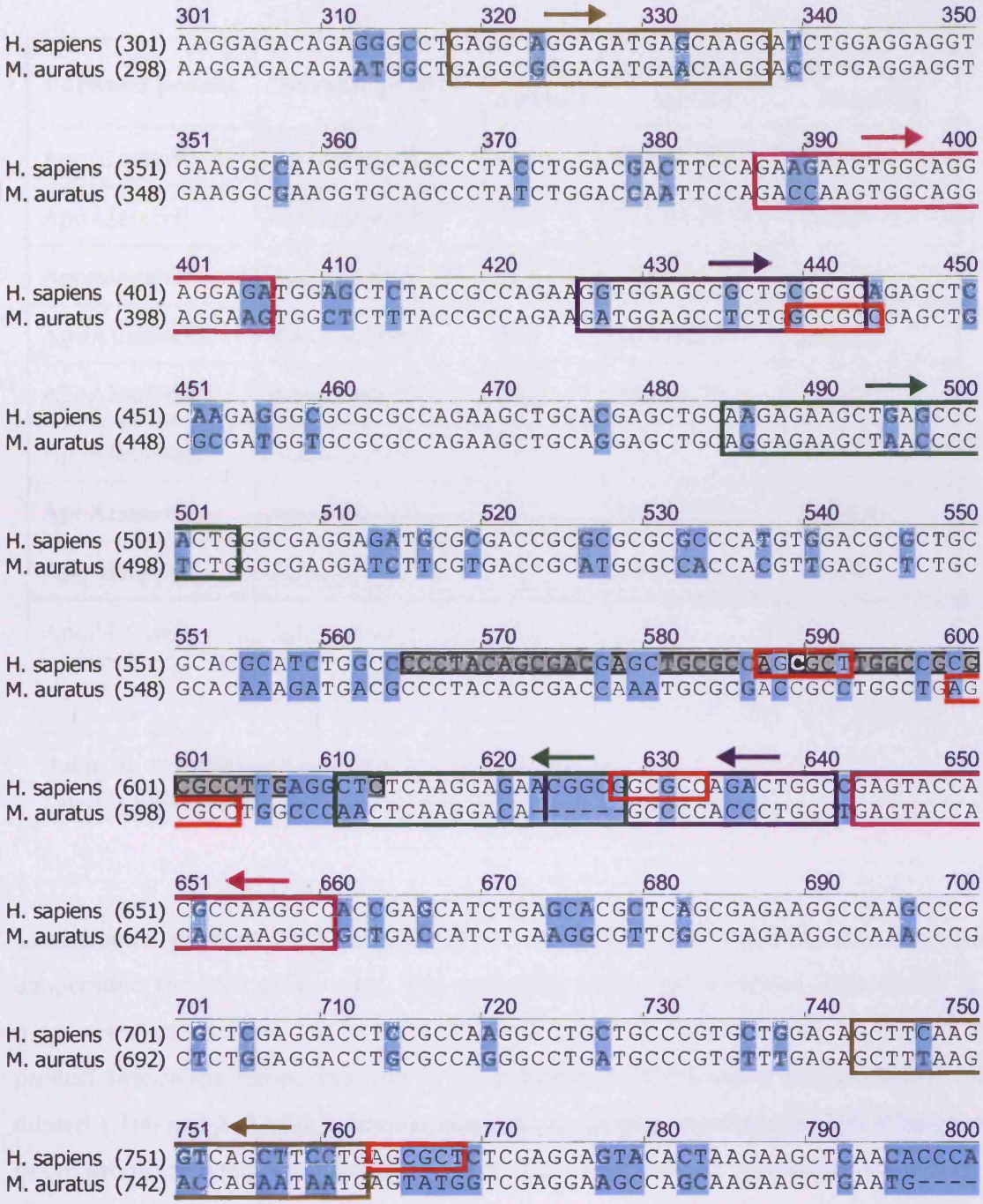


Figure 3-3 ApoA-I sequence comparison between *H. sapiens* and *M. auratus*

Blue background: Sequence differences between species. Black background: Site of Milano mutation C→T. Grey background: binding site of the editing oligonucleotide. Red boxes: HaeII restriction endonuclease recognition sites. Blue boxes: AI2F and AI4R primer binding sites; Green: ApoAIinnerR and ApoAIinnerF; Purple: ApoAIouterF2 and ApoAIinnerR2; Brown: ApoAIouterF and ApoAIouterR.

Forward primer	Reverse primer	PCR product	HaeII digest ApoA-I	HaeII digest ApoA-I_M
ApoAIouterF	ApoAIouterR	444	273,41,130	314,130
ApoAIouterF	ApoAIinnerR2	342	273,41,28	314,28
ApoAIouterF	ApoAIinnerR	310	273,41	314
ApoAIouterF2	**ApoAIouterR**	**376**	**205,41,130**	**246,130**
ApoAIouterF2	ApoAIinnerR2	274	205,41,28	246,28
ApoAIouterF2	ApoAIinnerR	246	205,41	246
ApoAIinnerF	ApoAIouterR	277	106,41,130	147,130
ApoAIinnerF	**ApoAIinnerR2**	**175**	**106,41,28**	**147,28**
ApoAIinnerF	ApoAIinnerR	143	106,41	147

Table 3-1 Possible combinations of ApoA-I PCR primers

Selected primer pairs in bold; primer pairs including ApoAIinnerR are in grey.

Temperature gradient PCRs (Figure 3-4) were utilised to optimise the annealing temperature for both primer sets. The annealing temperatures ranged from 55-68 °C, more specifically, from left to right 55/56/57.5/59.8/63.1/65.5/67.1/68 °C. The PCR product (annealing temperature 65.5 °C) indicated with an arrow (Figure 3-4a) was diluted 1:100 and 2 µl of this dilution was used as template for each 20 µl PCR reaction in the second temperature gradient using the ApoAIinnerF and ApoAIinnerR2 primers (Figure 3-4b).

As demonstrated in Figure 3-4a an unspecific band was present even when annealing temperatures higher than the calculated T_m of the primer set were used (see section 2.1.2.4). Notably, a band at around ~210 bp is present in equal measure in all samples of the second gradient PCR (Figure 3-4b).

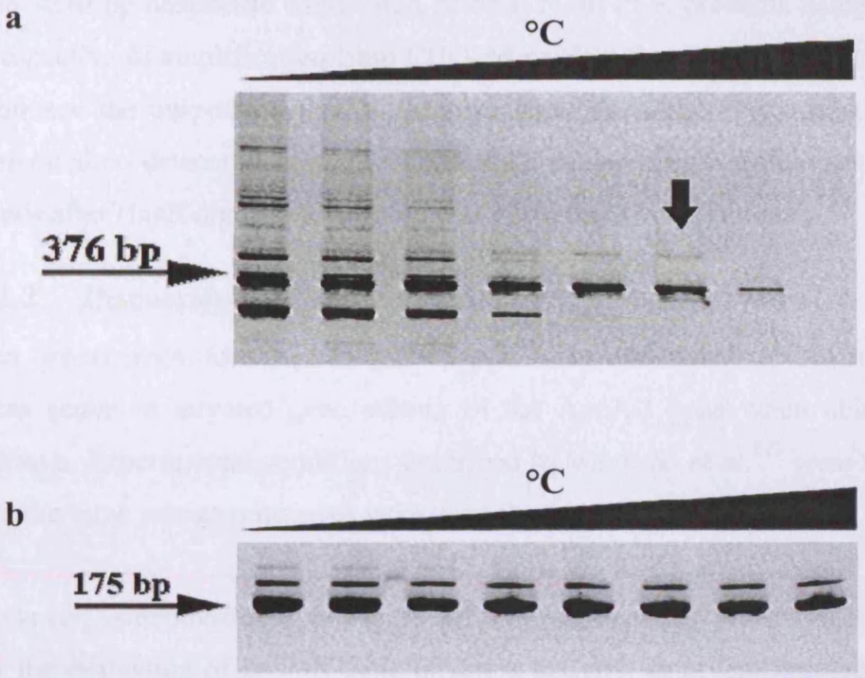


Figure 3-4 Temperature gradient PCRs with new primers

a) Primers ApoAIouterF2 and ApoAIouterR. Annealing temperature 55-68 °C from left to right. Template was genomic DNA from mock-transfected CHO-AI cells. The arrow indicates the sample used as template in the second temperature gradient PCR which is shown in **b)**.

b) Second temperature gradient PCR with primers ApoAIinnerF and ApoAIinnerR2. Annealing temperature 55-68 °C from left to right. An unspecific band of around 210 bp is present in all samples.

To determine if this ~210 bp unspecific band arose from the faint unspecific band at around 440 bp which can be seen in the sample indicated with an arrow (Figure 3-4a), a temperature gradient PCR using the primers ApoAIinnerF and ApoAIinnerR2 for amplification directly from the genomic sequence was carried out. This PCR yielded bands similar to those seen in Figure 3-4b but the 210 bp unspecific band was less prominent (data not shown). Therefore, further experiments utilized the primers ApoAIinnerF and ApoAIinnerR2 for amplification directly from genomic DNA.

The ~210 bp unspecific band could be seen in all PCR products using this primer set, irrespective of amplification from CHO-AI or HepG2 cells (Figure 3-11a). Attempts to sequence the unspecific band with either the ApoAIinnerF or ApoAIinnerR2 primer generated no detectable sequence trace. As the unspecific band did not result in visible bands after HaeII digestion it was ignored in further investigations.

3.1.2 Discussion

The experiments described in this chapter were undertaken to determine if ssODNs were active in targeted gene editing of the ApoA-I gene when chimeraplasts were inactive. Experimental conditions described in Manzano *et al.*[237] were followed closely and the same primer sequences were used for this PCR-RFLP analysis.

However, as demonstrated in Figure 3-2, the primers AI2F and AI4R proved unsuitable for the evaluation of ssODN gene targeting activity, as several unspecific bands could be seen after the restriction endonuclease digest. In particular, two equally strong bands close together with sizes around 200 bp made it difficult to identify which of them was the expected 207 bp band. Careful comparison with the 207 bp band in the AIM control sample strongly suggested that the upper of these two bands was equivalent to 207 bp. The lower band was estimated to be around 195 bp.

False identification of these bands would completely change the outcome of the experiment, as the lower band was still present in the digestion control sample as well as the treated samples. If the lower band was the correct 207 bp band then the bands in the treated samples would be due to incomplete digest and would not indicate successful gene targeting. Thus the presence of these unspecific bands cast doubt on the data obtained in this PCR-RFLP. To obtain clear results the unspecific bands had to be eliminated.

It is unclear why these unspecific bands were present on this analytic gel but not the gels shown by Manzano *et al.*[237] even though the conditions and reagents reflected the published ones. A possible explanation is the use of different Taq polymerases, which could influence enzyme-DNA affinity or mispriming. Another possibility is the percentage of DMSO in the PCR mixture. Based on personal experience, I added 5% DMSO to the PCR in this analysis to help with the strand separation of genomic DNA during the denaturing step. Different PCR machines or PCR tubes may also have influenced the PCR conditions.

The biggest problem however seemed to be the primers itself. A check of the AI2F and AI4R primers revealed that both primers form self-annealing dimers and duplexes and have melting temperatures of 79 °C. I reasoned that the selection of better primers was the quickest way of eliminating unspecific bands from PCR-RFLP analytic gels.
Choosing the right primer sets and set combinations proved rather difficult as I had to take a number of different requirements into account. Following standard primer design rules, such as similar melting temperatures and no primer dimer or duplex formation, were less of a problem, but it proved difficult to find primers with only ~50% G/C content. The 500 bp sequence stretch surrounding the HaeII site of interest has an average G/C content of 67% and this severely limited the choice of primers.

A further limiting factor was the sequence homology between *H. sapiens* and *M. auratus*. Primers specific for the human sequence should ideally have at least 3 mismatches with the hamster sequence, preferably as close to the 3' end as possible. Primer sequences were also screened against the human and rodent mispriming libraries in primer3, again reducing the number of possible primers. Furthermore, HaeII digestion fragment sizes for ApoA-I and ApoA-I_M had to be easily distinguishable on either agarose or TBE gels.

Fulfilling all requirements seemed impossible with only one primer set, so a nested PCR with outer and inner primers sets was designed. However, experimental data showed that a PCR with the inner primer set alone resulted in clean HaeII digestion bands, so all further analysis was done in this way.

The remaining ~210 bp unspecific band that was seen in the PCR product but not after HaeII digestion could not be identified. It could not have been due to mispriming on the hamster sequence, as it was also present in PCRs using genomic DNA from the human HepG2 cell line as template. I considered possible amplification of an apolipoprotein family member as an explanation since this family of genes is closely related, but the mispriming function in primer3 was supposed to exclude this. In the end, this unspecific band did not interfere with evaluation of the PCR-RFLP, so I deemed further investigation into the causes of this band unnecessary.

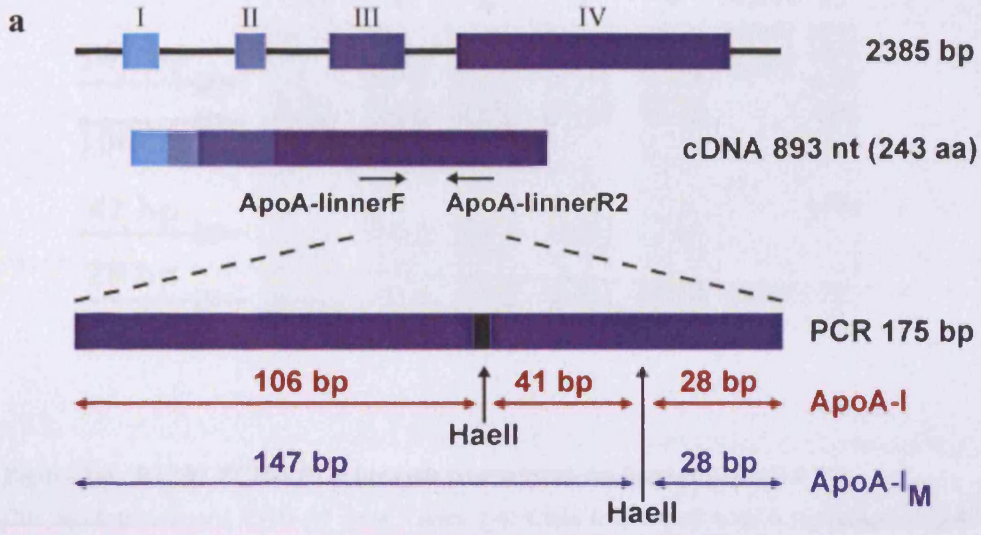
3.2 Conversion of ApoA-I to ApoA-I_M in CHO-AI cells and persistence of gene edited cells in the population

3.2.1 Results

CHO-AI cells expressing human ApoA-I were transfected with 6 or 12 μg of huApoA-I_M-49T-PTO and huApoA-I_M-49NT-PTO oligonucleotide using LPEI. Cells were trypsinised 24 h after transfection and one-third of the cells were re-plated while genomic DNA was extracted from the other two-thirds. Samples were analysed by a R173C PCR-RFLP assay using the primers ApoAIinnerF and ApoAIinnerR2 and 68 °C annealing temperature (Figure 3-5).

The predicted bands at 106 and 41 bp demonstrating the presence of the ApoA-I wild-type sequence were present in all samples except the AIM control (Figure 3-5). The 147 bp band denoting the presence of the Milano mutation was strongly visible in lanes 3, 4 and AIM, but barely detectable in the mock-transfected control sample (Ctr). Therefore, it was assumed that all samples were completely digested and that the 147 bp bands still visible were not due to incomplete digestion. Thus, targeted gene conversion was seen in cells transfected with 6 μg of huApoA-I_M-49NT-PTO although it seemed to be less pronounced in cells treated with 12 μg of the same ssODN. Cells transfected with huApoA-I_M-49T-PTO did not show conversion, as the extremely faint 147 bp band was less detectable than the Ctr sample. Results were consistent in two independent experiments and three separate PCR-RFLP analyses.

To establish the persistence of the introduced nucleotide change, samples for genomic DNA extraction were taken at every passage (days 1, 4, 7 and 11). The level of conversion diminished during passage 1-3, but was still evident in passage 4 (Figure 3-6).



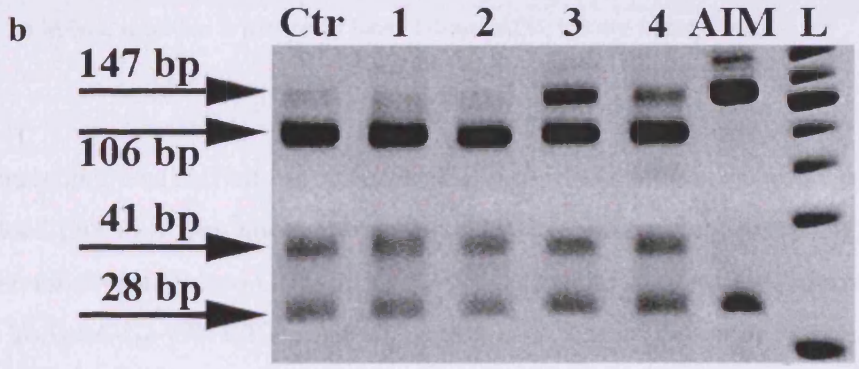


Figure 3-5 R173C PCR-RFLP with new primers

a) The ApoA-I gene is represented above with boxes for its four exons. The primers ApoAIinnerF and ApoAIinnerR2 amplify a 175 bp product, which yields HaeII digestion fragments of 106, 41 and 28 bp for wild-type ApoA-I or 147 and 28 bp for ApoA-I_M.

b) 20% TBE gel resolving the HaeII digestion fragments. Ctr: mock-transfected CHO-AI cells. Lanes 1-4: Cells transfected with 6 μg, 12 μg huApoA-I_M-49T-PTO, 6 μg, 12 μg huApoA-I_M-49NT-PTO; AIM: CHO-AIM cells. L: 25 bp ladder.

For the ApoA-I_M sequence positive control, genomic DNA was extracted directly from defrosted CHO-AIM cells without growing the cells in culture. Thus, any potential contamination of CHO-AI with CHO-AIM cells was avoided.

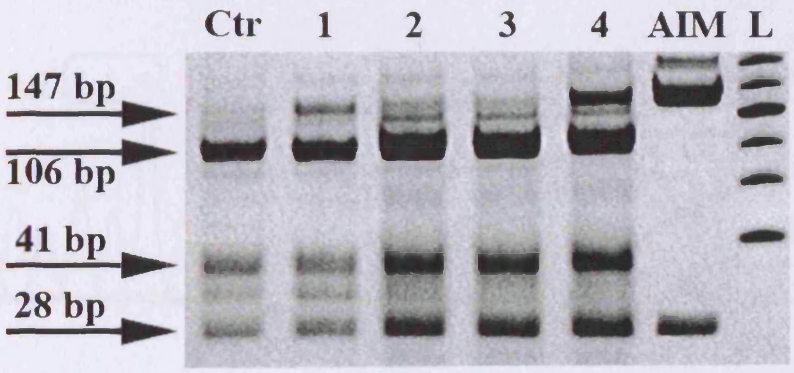


Figure 3-6 R173C PCR-RFLP for cells treated with huApoA-I_M-49NT-PTO

Ctr: mock-transfected CHO-AI cells. Lanes 1-4: Cells transfected with 6 μg huApoA-I_M-49NT-PTO, 1st, 2nd, 3rd and 4th passage. AIM: CHO-AIM cells. L: 25 bp ladder. The expected bands at 106 and 41 bp are present in all samples except AIM. The 147 bp band denoting the presence of the Milano mutation is present in lanes 1-4 and AIM, but not in the control.

Sequencing was carried out to determine if the 147 bp band in lane 4 of Figure 3-6 was indeed due to a true nucleotide change and not incomplete digestion. Genomic DNA from mock-transfected CHO-AI cells, cells from the 4th passage after transfection with 6 μg huApoA-I_M-49NT-PTO (same sample as in Figure 3-6 lane 4) and CHO-AIM cells was PCR amplified with the primers ApoAIinnerF and ApoAIinnerR2.

The PCR products from 4th passage huApoA-I_M-49NT-PTO and CHO-AIM cells were digested with HaeII and resolved on a 3.5% MetaPhor Agarose gel. The 147 bp band was excised, eluted with the QiaQuick Kit and used in sequencing (section 2.2.2.12). Sequencing from the mock-transfected cells was performed directly from the PCR product. HaeII digestion of this PCR product would result in termination of the sequence trace at nucleotide 591 in the coding sequence of ApoA-I, only two nucleotides downstream of the expected mutation (see Figure 3-3). Therefore, this PCR product was not digested with HaeII.

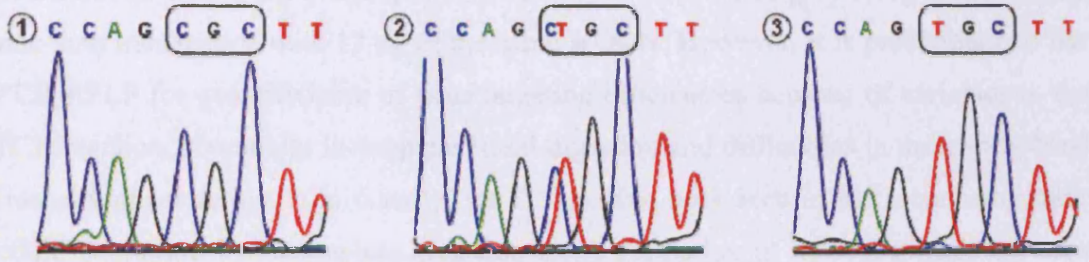


Figure 3-7 Sequencing of PCR-RFLP band

① Mock-transfected cells. ② 4th passage huApoA-I_M-49NT-PTO cells. ③ CHO-AIM cells. The nucleotides coding for amino acid Arg^{173} (CGC; ApoA-I) and Cys^{173} (TGC; ApoA-I_M) are ringed.

Figure 3-7 shows the sequencing traces from nucleotides 585 to 593 for all three samples. Genomic DNA from mock-transfected CHO-AI cells codes for CGC at nucleotides 589-591 translating into Arginine at amino acid position 173. CHO-AIM cells carried a TGC codon instead, which codes for Cysteine. Sequencing of the 147 bp band from the 4th passage huApoA-I_M-49NT-PTO cells showed a clear C/T double peak at nucleotide 589 and therefore established the presence of successfully gene targeted cells 11 days after transfection.

3.2.2 Discussion

Using the new primer set ApoAIinnerF and ApoAIinnerR2, gene editing could be demonstrated in CHO-AI cells transfected with huApoA-I_M-49NT-PTO. Cells transfected with the huApoA-I_M-49T-PTO ssODN did not show any detectable correction, but this was not surprising as strand bias has been described previously.[183,240] Strand bias seems to predominantly favour ssODNs complementary to the non-transcribed strand, possibly because of increased accessibility of this strand during transcription.[263] Replication orientation of the target gene has also been investigated as the possible cause for this phenomenon,[251,267] but Sorensen *et al.*[253] recently reported opposite strand bias for two mutations in the same gene.

An interesting finding is the apparent lack of dose-response in this experiment. Transfection with 6 μg of huApoA-I_M-49NT-PTO appears to give a higher correction rate than transfection with 12 μg of the same ssODN. However, it is problematic to use PCR-RFLP for quantification of gene targeting efficiencies because of variables in the PCR reaction, potentially incomplete HaeII digestion and difficulties in measuring band intensity at saturation. Additionally, the C/T double peak seen in the sequencing trace could be caused by incomplete digestion or by formation of heteroduplexes between ApoA-I and ApoA-I_M strands during PCR. Both possibilities would certainly cause a significant overestimation of gene editing efficiencies. Potentially, detection of the diagnostic bands in the experiments presented here would be impossible without this phenomenon.

One could speculate that heteroduplex formation would also take place in PCR-RFLP analyses based on the creation of restriction sites. These heteroduplex sequences would not be cut efficiently by restriction endonucleases and thus low gene editing efficiencies would be concealed in the band diagnostic for non-edited cells. On the other hand, heteroduplex formation can be avoided by primer excess during the later cycles of PCR, which is generally the case in nested PCRs or re-amplification of low cycle reactions.[286] Thus, differing PCR conditions could at least in part explain why some laboratories are unable to repeat published results.

Nevertheless, the purely qualitative difference stated here is supported by the consistency of the results in two experiments and at least two separate PCR-RFLP analyses each, which suggests that the difference between the 6 and 12 μg samples was not caused by an artefact. Furthermore, the band intensity of the smaller digestion fragments in both samples has not reached saturation and can therefore be used for comparison. One possible explanation for the lack of dose-response could be cytotoxicity caused by the higher concentration of ssODN on the cells and visual examination of the cells during the experiment supported this conclusion.

When genomic DNA was extracted from the treated cells at passages 2 and 3, a marked reduction in the correction efficiency was observed in comparison to the first passage. With such a faint 147 bp band in the 3rd passage sample, it seemed that corrected cells had been largely lost from the population by that time. However, when genomic DNA was extracted in the 4th passage, the conversion was still evident.

Originally, this puzzling result was attributed to incomplete digest of the 4th passage sample, but when repeated PCR-RFLP analysis resulted in similar band intensities, the 147 bp from this sample was sequenced. The sequence trace confirmed the presence of successfully corrected cells in the population 11 days after transfection. This was extremely encouraging, as previous experiments in our lab had shown that corrected cells are lost from the population over several passages and attempts to clone out corrected cell lines had been unsuccessful.[235]

I concluded that successful correction had taken place, but that the correction event itself somehow had inhibited cell survival and proliferation of the affected cells. Thus the normal proliferation of the vast number of non-corrected cells in the population simply diluted out the limited number of growth inhibited corrected cells. This conclusion is supported by three recently published papers. Olsen *et al.*[242] showed that corrected cells accumulated at the G2/M checkpoint and that the majority of these cells subsequently succumbed to apoptosis. Only ~5% of corrected cells had divided 24 h after transfection, but this percentage rose to 30% when cells were treated with the Chk1/Chk2 inhibitor Gö6976. Chk1 and Chk2 are downstream mediators in the ATM/ATR DNA damage response pathway and involved in the G2/M cell cycle checkpoint regulation.

Ferrara *et al.*[279] confirmed that Chk1 and Chk2 were highly activated in corrected cells, but not in uncorrected cells, and that corrected cells did not actively replicate their DNA 24 h after introduction of targeting oligonucleotide. Engstrom *et al.*[287] showed that this selective arrest was responsible for the observed reduction in correction efficiency over time and that replication was reactivated in these cells after several days. She also demonstrated that 48 h incubation with 2 mM thymidine after targeting suppressed replication in uncorrected cells and that the percentage of corrected cells in the population could thus be stabilised for up to 6 days.

In addition, the finding that successfully corrected cells are present in the population 11 days after treatment provides evidence that the diagnostic bands in the PCR-RFLP analysis are not due to PCR artefacts caused by the correcting oligonucleotide as previously suggested.[218] Phosphorothioate oligonucleotides are protected against exonuclease, but not against endonuclease attack, and half-life in the cell nucleus is ~ 24 h as determined with a fluorescently labelled ssODN (section 4.1).

3.3 Diagnostic bands in the R173C PCR-RFLP are not due to PCR artefacts

3.3.1 Results

It was previously reported that PCR-RFLP could lead to false positive results because the correcting chimeraplast was acting as a template in the PCR reaction.[214,218] Subsequent experiments in our lab did not produce such artefacts when ApoE4-to-ApoE3 chimeraplast was spiked into the genomic DNA from ApoE4 expressing CHO cells before PCR-RFLP.[235] Neither did artefacts arise when the parental cell line CHO-K1 was transfected with chimeraplast and extracted genomic DNA mixed with genomic DNA from CHO-ApoE4 cells at various ratios. Nevertheless, the set of experiments described on the following pages attempted to determine if ssODNs could be the cause of such misleading results in the R173C PCR-RFLP.

PCR reactions using genomic DNA from mock-transfected CHO-AI cells as a template were spiked with 300 ng, 3 ng and 30 pg of huApoA-I_M-49NT-PTO before amplification (Figure 3-8a). The amount of 300 ng was determined by dividing the maximum amount of ssODN transfected into cells (12 µg) by the volume genomic DNA was eluted in (200 µl), multiplied by 5, as 5 µl of genomic DNA were generally used per PCR reaction. Thus, 300 ng of ssODN in a PCR reaction should reflect the theoretical maximum amount available for mispriming under the conditions used.

This theoretical maximum amount of ssODN should vastly exceed the amount actually present, as losses incurred during transfection (non-transfected complexes are washed off), cell culture (ssODN will be digested by endonucleases) and purification of genomic DNA (fragments of less than 100 bp are not efficiently retained by the purification column) were not taken into account.

As shown in Figure 3-8a, a very prominent band at around 100 bp is present in the 300 ng sample as well as fainter bands at ~70 bp, ~400 bp and ~1 kb. The expected 175 bp and the ~210 bp unspecific band were present in all three lanes. When these PCR products were digested with HaeII (Figure 3-8b), a faint 147 bp band was seen in lane 1, but not in 2 or 3. In fact, lanes 2 and 3 displayed the typical restriction pattern associated with ApoA-I, whereas lane 1 showed a prominent ~80 bp band and an unusually strong 28 bp band as well as faint ~60 bp and ~26 bp bands.

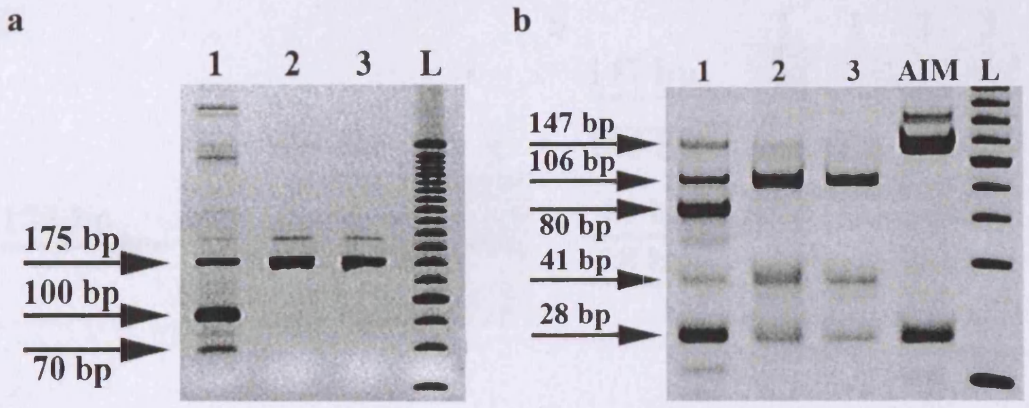


Figure 3-8 R173C PCR-RFLP spiked with huApoA-I_M-49NT-PTO before PCR

a) Lane 1-3: 300 ng, 3 ng and 30 pg of huApoA-I_M-49NT-PTO were spiked into PCR reactions containing 5 µl genomic DNA from mock-transfected CHO-AI cells. PCR products are shown. L: 25 bp ladder.

b) Lanes 1-3: HaeII digestion of samples from a). AIM: CHO-AIM cells. L: 25 bp ladder.

To determine if these unspecific bands seen in lane 1 were due to the presence of such a high amount of DNA in the PCR reaction or were indeed caused by the ssODN acting as a template, mock-transfected CHO-AI cells were harvested and spiked with 4 µg of huApoA-I_M-49NT-PTO or huApoA-I_M-49T-PTO before genomic DNA extraction. The PCR products were as expected (Figure 3-9a) but a faint 147 bp band was present in the huApoA-I_M-49NT-PTO sample after the HaeII digest (Figure 3-9b lane 2).

Another control experiment explored the possibility that PCR artefacts could be caused by cleavage products of ssODNs produced during incubation in cells. CHO^{dhfr-} cells, from the parental cell line for CHO-AI, were transfected with 4 µg of huApoA-I_M-49NT-PTO or huApoA-I_M-49T-PTO using Lipofectamine 2000. Transfection efficiency was controlled with a Cy5-tagged ssODN and determined to be ~95% (data not shown). Cells were harvested 24 h post transfection and genomic DNA was extracted. For the PCR reactions shown in lanes 3-5 of Figure 3-10 the genomic DNA from transfected CHO^{dhfr-} was mixed with genomic DNA from mock-transfected CHO-AI cells at various ratios (1:3, 1:1 and 3:1).

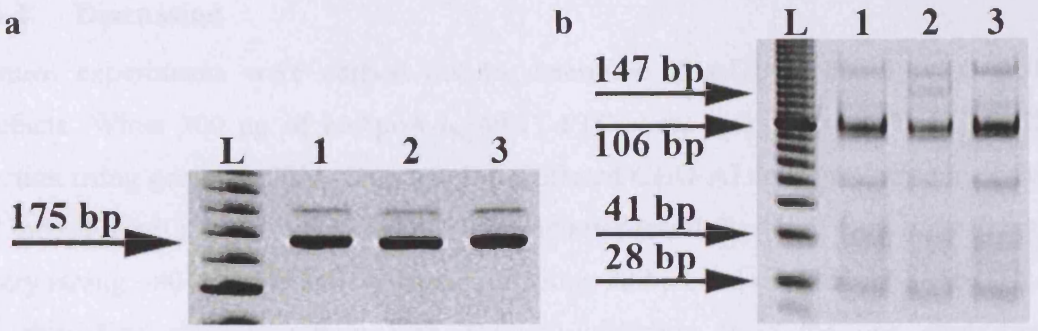


Figure 3-9 R173C PCR-RFLP spiked before genomic DNA extraction

a) Lane 1-3: mock-transfected CHO-AI cells, CHO-AI cells spiked with 4 µg huApoA-I_M-49NT-PTO, CHO-AI cells spiked with 4 µg huApoA-I_M-49T-PTO. L: 25 bp ladder.

b) Lanes 1-3: HaeII digestion of samples from a). L: 10 bp ladder.

A faint 175 bp PCR product was produced by CHO^{dhfr-} cells transfected with 4 µg huApoA-I_M-49NT-PTO or huApoA-I_M-49T-PTO, but no HaeII digestion fragments could be detected in these samples. There was a clear reduction in amount of PCR product when increasing amounts of CHO^{dhfr-} genomic DNA were mixed with decreasing amounts of CHO-AI genomic DNA. The HaeII digestion pattern of the mixed samples was as expected from ApoA-I DNA, with no evidence of the 147 bp band characteristic for ApoA-I_M.

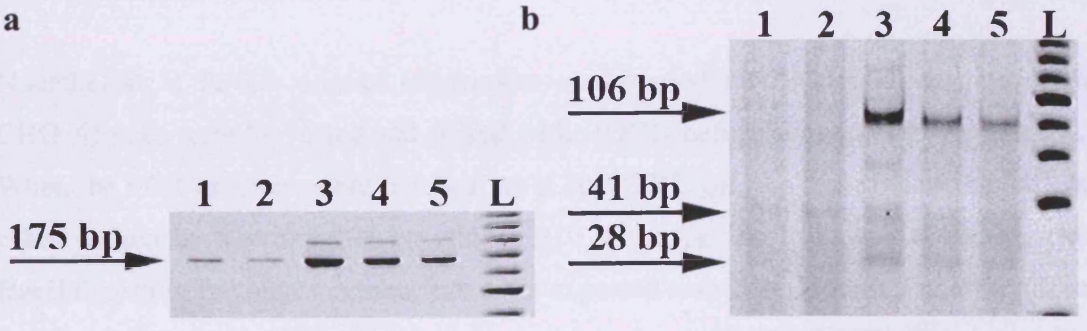


Figure 3-10 R173C PCR-RFLP of treated CHO^{dhfr-} cells mixed with CHO-AI cells

a) Lane 1, 2: CHO^{dhfr-} cells transfected with 4 µg huApoA-I_M-49NT-PTO or huApoA-I_M-49T-PTO. Lanes 3-5: CHO^{dhfr-} cells transfected with 4 µg huApoA-I_M-49NT-PTO mixed with CHO-AI cells at 1:3, 1:1 and 3:1 ratio. L: 25 bp ladder.

b) Lanes 1-5: HaeII digestion of samples from a). L: 25 bp ladder.

3.3.2 Discussion

Control experiments were carried out to determine if ssODNs could cause PCR artefacts. When 300 ng of huApoA-I_M-49NT-PTO were spiked directly into a PCR reaction using genomic DNA from mock-transfected CHO-AI cells, it seemed as if this did indeed result in a false-positive 147 bp band after HaeII digestion. Strikingly though a very strong ~80 bp band and an unusually strong 28 bp band could also be seen. Thus the restriction enzyme pattern was distinctly different from the one established previously for either ApoA-I or ApoA-I_M.

Moreover, when the PCR products of this spiking reaction were resolved on a TBE gel, several unspecific bands could be detected. In particular, a very strong ~100 bp band and a weak ~75 bp were observed. Again, this pattern did not agree with the one demonstrated in earlier experiments or by the other samples in the same analysis.

These data suggested that the high amount of ssODN present in the PCR reaction in this particular sample did interfere with the normal reaction parameters. Considering that 300 ng ssODN was equivalent to the theoretical maximum amount present in treated cells and that the resulting peculiar PCR products and HaeII digestion patterns did not appear in any of the experimental PCR-RFLP analyses, it seemed appropriate to use the sample spiked with 3 ng ssODN as the proper control. As the 3 ng PCR reactions did not result in a 147 bp band characteristic for ApoA-I_M, the possibility of false-positive PCR artefacts was rejected.

Nonetheless, a further control experiment was carried out, where mock-transfected CHO-AI cells were harvested and spiked with ssODN before genomic DNA extraction. When the PCR products were resolved on a 20% TBE gel, the typical banding pattern characterized by a strong 176 bp and a ~210 bp unspecific band was observed. The HaeII digestion fragments demonstrated the expected restriction pattern, but a very faint 147 bp could be seen in the sample spiked with huApoA-I_M-49NT-PTO. However, very faint ~65 bp and 176 bp bands were also present, suggesting incomplete digest even after extended HaeII digestion.

Another control experiment explored the possibility that PCR artefacts could be caused by cleavage products of ssODNs produced by exo- or endonucleases in the cells. Genomic DNA was extracted from CHO^{dhfr-} cells that had been transfected with huApoA-I_M-49NT-PTO or huApoA-I_M-49T-PTO and mixed with genomic DNA from mock-transfected CHO-AI cells before a PCR-RFLP analysis was carried out.

A weak 176 bp PCR product could be seen in the CHO^{dhfr-} cells transfected with ssODN, even though these cells did not contain the human ApoA-I sequence. This was likely due to amplification of the hamster ApoA-I in lieu of the human sequence. When these PCR products were digested with HaeII no bands were visible on the analytical gel, probably because the amount of digestion fragments derived from such a small quantity of PCR product was below the detection limit.

It is unclear if this small amount of cross-species mispriming would also occur in the presence of the human ApoA-I sequence. Considering that the CHO-AI cell line has multiple copies of the human ApoA-I sequence stably integrated into its genome compared to the two hamster alleles, and taking into account the greater sequence homology of the primers with the human sequence, this seems unlikely. Even if cross-species mispriming was not influenced by the presence of the human ApoA-I sequence in the PCR reaction, the digestion fragments produced from the misprimed PCR products would not interfere with the evaluation of the PCR-RFLP analysis, as demonstrated by the lack of visible bands after HaeII digestion.

In the mixed genomic DNA samples a clear dose-response to the presence of CHO-AI genomic DNA could be seen: the more CHO-AI genomic DNA in the reaction, the more PCR product was produced. There was no evidence of the 147 bp band characteristic for ApoA-I_M after HaeII digestion. Thus, fragments of ssODN produced by nuclease digestion in the cells did not cause any PCR artefacts.

In summary, the evidence presented here strongly suggests that the diagnostic bands derived in the PCR-RFLP analyses shown in chapter (section 3.2) were not due to artefacts.

3.4 Conversion of ApoA-I to ApoA-I_M in HepG2 cells

3.4.1 Results

In addition to the recombinant CHO-AI cells, I targeted the human hepatoblastoma cell line HepG2 to study gene editing in a cell line with endogenous ApoA-I expression. HepG2 cells were "reverse transfected", i.e. the transfection complexes were added directly after trypsinisation and suspension of cells.[284] This "reverse transfection" was particularly successful for HepG2 cells as this cell line tends to grow in cell clumps rather than a monolayer and prefers tight cell to cell contact. Therefore, the cell surface area accessible to transfection complexes is limited if cells are pre-plated, but maximised when cells are in suspension during transfection.

One million HepG2 cells per well of a 6-well plate were transfected with 4 μg of huApoA-I_M-49NT, huApoA-I_M-49T-PTO, huApoA-I_M-49NT-PTO or huApoA-I_M-Cy5-49NT-PTO using a ssODN to Lipofectamine 2000 ratio of 1:2 (w/v). Transfection efficiency was monitored and estimated to be 40% in pEGFP and ~80% in Cy5-ssODN transfected cells (data not shown). To maximise the percentage of converted cells, transfection was repeated 3 and 5 days after the initial transfection.

However, cells treated with huApoA-I_M-49NT-PTO showed a high level of cell death after the second transfection and therefore were not transfected for a third time. In comparison to huApoA-I_M-49NT-PTO the unmodified ssODN showed very little cytotoxicity, even after repeated administration. As previously discussed (section 3.1.2), the unspecific ~210 bp band was also present in PCR samples from human cells (Figure 3-11a).

Unfortunately, as seen in Figure 3-11b repeat administration did not translate into a higher conversion rate when cells were analysed 24 h after the third transfection, i.e. on day 6 after the first transfection. The expected bands at 106 and 41 bp are present in all samples except AIM. The 147 bp band denoting the presence of the Milano mutation is clearly present in lanes 3, 4 and AIM, but not in the control. A faint 147 bp band over the background control (Ctr) may be present in lane 2. When the 147 bp band from Figure 3-11b lane 3 was sequenced, again a clear C/T double peak could be seen (Figure 3-11c ②), indicating successful gene targeting. However, when cells from this sample where subjected to limiting dilution and clonal expansion, cell lines carrying the Milano mutation could not be established (data not shown).

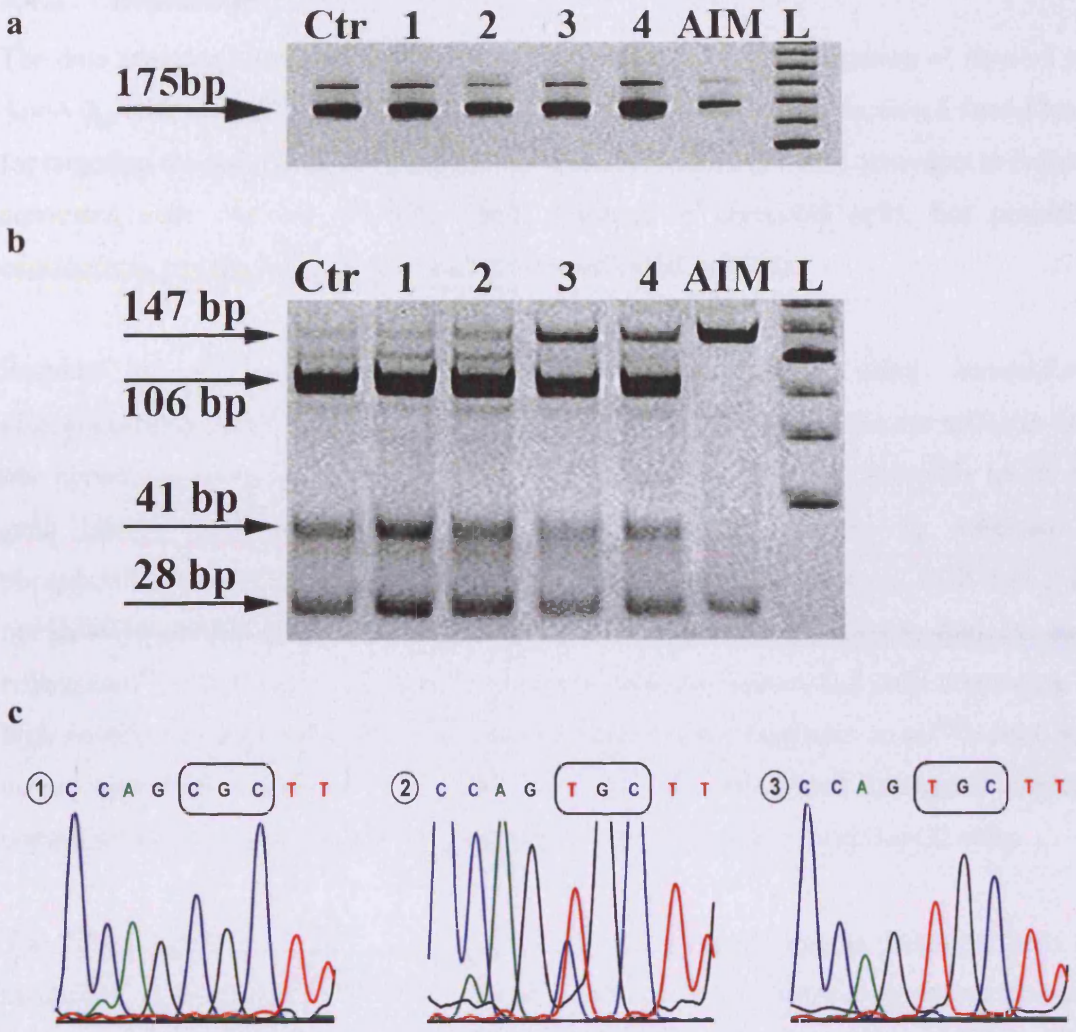


Figure 3-11 R173C ApoA-I PCR-RFLP and sequencing in HepG2 cells

a) PCR products. Ctr: mock-transfected HepG2 cells. Lanes 1-4: Cells transfected with 4 μg of huApoA-I_M-49NT, huApoA-I_M-49T-PTO, huApoA-I_M-49NT-PTO and huApoA-I_M-Cy5-49NT-PTO. AIM: CHO-AIM cells. L: 25 bp ladder.

b) HaeII digestion fragments. Ctr: mock-transfected HepG2 cells. Lanes 1-4: Cells transfected with 4 μg of huApoA-I_M-49NT, huApoA-I_M-49T-PTO, huApoA-I_M-49NT-PTO and huApoA-I_M-Cy5-49NT-PTO. AIM: CHO-AIM cells. L: 25 bp ladder.

c) ① Sequencing of mock-transfected cells. ② HepG2 cells transfected with 4 μg huApoA-I_M-49NT-PTO. ③ CHO-AIM. The bases corresponding to residues Arg^{173} (CGC; ApoA-I) and Cys^{173} (TGC; ApoA-I_M) are ringed.

3.4.2 Discussion

The data presented in this chapter demonstrates successful gene targeting of ApoA-I to ApoA-I_M with ssODN in HepG2 cells. As with CHO-AI cells, a pronounced strand bias for targeting the non-transcribed strand was observed (section 3.2.2). Attempts to isolate converted cells did not produce viable colonies of corrected cells, but possible explanations for this have already been discussed (section 3.2.2).

Radecke *et al.*[241] presented evidence that gene editing using unmodified oligonucleotides led to stable genotype conversion when phosphorothioate ssODNs did not. However, in our hands an unmodified ssODN did not produce detectable levels of gene editing even after triple transfection (Figure 3-11b lane 1), whereas a phosphorothioate oligonucleotide of the same sequence did. Of course, PCR-RFLP is not sensitive enough to pick up correction rates as low as those reported by Radecke and colleagues[241] (<1%) with a direct flow cytometry-based readout and cells containing a high number of integrated transgene copies. Furthermore, Radecke *et al.*[241] could not detect gene conversion in cells that had only 1-2 integrated transgene copies, comparable to the situation with the endogenous ApoA-I sequence in HepG2 cells.

The difference in correction efficiency observed in the experiments described here is most likely due to the difference in endonuclease resistance between phosphorothioate and unmodified ODN. Multiple transfections did seem to increase the correction rate, but again problems with quantification of PCR-RFLP diagnostic bands made it impossible to determine absolute numbers. Especially worrying in this context was the repeatedly demonstrated contamination of the 147 bp band with ApoA-I/ApoA-I_M heteroduplexes seen in the sequencing results.

The high cytotoxicity associated with repeated Lipofectamine 2000 administration of phosphorothioate ssODNs limited the usefulness of this approach at least for HepG2 cells. Later experiments with CHO-mEGFP cells demonstrated that repeated targeting did not improve gene editing efficiencies (section 5.1).

3.5 Effects of co-transfection with reporter gene plasmid

3.5.1 Results

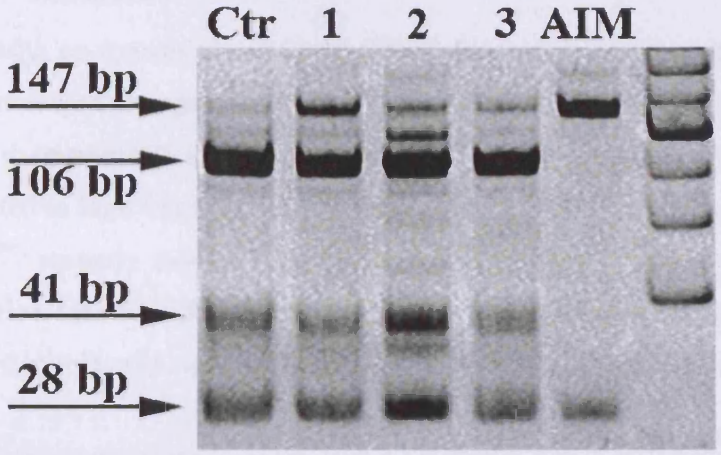
As established in the previous experiments, successful gene targeting can be shown in both CHO-AI and HepG2 cells several passages after transfection; however, isolation of converted cell clones proved difficult. Therefore, ways of selecting for successfully targeted cells were investigated.

Co-transfection has been successfully used with two plasmids where GFP expression from one plasmid was substituted as a marker to sort for cells expressing the gene of interest from the second plasmid.[288,289] Here, correcting ssODN was co-transfected with a pEGFP plasmid in the hope that reporter gene expression would mark cells in which plasmid, and potentially ssODN, was efficiently delivered to the nucleus. These cells should therefore have a higher percentage of gene-edited cells.

HepG2 cells were transfected with 4 μg ssODN and 1 μg pEGFP-C1 using Lipofectamine 2000. Cells were trypsinised 24 h post-transfection and cell sorted for EGFP positive and negative cell populations. DNA was extracted from the cells and a R173C PCR-RFLP analysis was carried out. However, due to the low number of cells in the EGFP positive population (~5,000 EGFP positive cells were recovered) the PCR product for this sample had to be re-amplified. For the re-amplification, 2 μl of the first PCR mixture was used as template in a second PCR of 20 cycles under the same conditions.

The PCR-RFLP analysis (Figure 3-12a) did show clear gene conversion in the unsorted population (lane 1), but also for both the EGFP positive (lane 2) and negative populations (lane 3). However, PCR-RFLP is not suitable for quantitative assessment of conversion efficiency and therefore it is difficult to draw firm conclusions from these results. Sequencing of the 147 bp HaeII digestion fragment from the unsorted population confirmed the nucleotide change from C to T (Figure 3-12b ②). This nucleotide change could also be seen in the EGFP positive and negative populations (data not shown). Corrected cells persisted for at least 10 d in the unsorted population as determined by PCR-RFLP and subsequent sequencing of the 147 bp diagnostic band (data not shown).

a



b

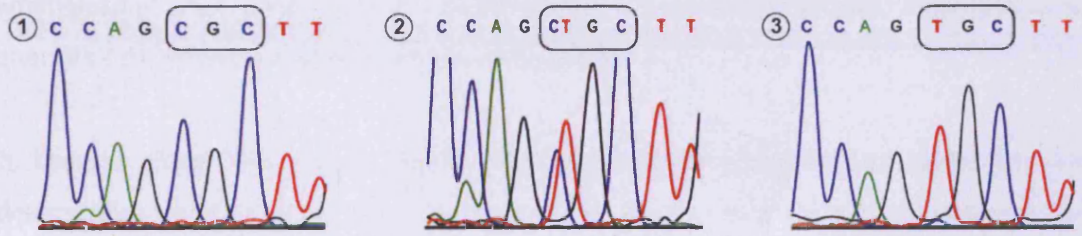


Figure 3-12 R173C PCR-RFLP for cells co-transfected with GFP plasmid

a) Ctr: mock-transfected HepG2 cells. Lane 1: Cells transfected with 4 µg of huApoA-I_M-49NT and 1 µg pEGFP-C1. Lane 2 and 3: Cells from lane 1 sorted for EGFP positive and negative cells. AIM: CHO-AIM cells. The expected bands at 106 and 41 bp are present in all samples except AIM. The 147 bp band denoting the presence of the Milano mutation is clearly present in lanes 1-3 and AIM, but not in the control.

b) ① Sequencing of mock-transfected cells. ② HepG2 cells transfected with 4 µg huApoA-I_M-49NT and 1 µg pEGFP. ③ CHO-AIM. The bases corresponding to residues Arg^{173} (CGC; ApoA-I) and Cys^{173} (TGC; ApoA-I_M) are ringed.

3.5.2 Discussion

Curiously, co-transfection with an EGFP expressing plasmid appeared to increase the correction rate, though once more the imprecise nature of quantifying PCR-RFLP bands made it impossible to be sure. At the time, the apparent high correction rate was attributed to high transfection efficiency and not followed up any further, but Igoucheva *et al.*[290] recently reported a similar finding. They demonstrated that the elevated correction rate was due to the higher level of activation of the ATM/ATR pathway induced by plasmid co-transfection in comparison to ssODN only transfection.

Cell sorting for EGFP positive and negative populations did not appear to enrich for corrected cells in the positive population, but on top of the already discussed problems with PCR-RFLP band quantification, this analysis was complicated by the re-amplification step needed for the EGFP positive population. Further experiments to quantify this potential increase are shown in section 5.2.

It became clear from these results, that PCR-RFLP analysis was unsuitable for determining absolute correction numbers and that a more accurate method was required to gain meaningful and usable data. A flow cytometry-based read-out has been widely used in conjunction with stably integrated mutated reporter genes[228,250] and I considered a similar system for detection of ApoA-I_M in live cells. Unfortunately, I was unable to source an antibody capable of distinguishing the wild-type ApoA-I protein from the mutated Milano form at that time.

Moreover, ApoA-I is rapidly secreted from liver cells and although Brefeldin A could be used to block protein excretion, intracellular antigen staining requires fixation and permeabilization of cells. Consequently, I would be unable to recover live cells, although such an assay could theoretically be used to analyse gene editing efficiencies of ApoA-I to ApoA-I_M targeting.

Thus, I decided to change the experimental system from targeting ApoA-I to targeting a mutated reporter gene, in the hope that quantitative data gained in such a system could in time be applied to ApoA-I gene editing. Prof. Krauss from the Rikshospital in Oslo, Norway kindly provided CHO and HEK293 cell lines with the mutated EGFP stably integrated. Both these cell lines have been used extensively in their gene editing studies.[242,251]

CHAPTER 4: OPTIMIZATION OF OLIGONUCLEOTIDE-MEDIATED GENE EDITING

INTRODUCTION

One possible explanation for the problems with isolating viable corrected cell clones could be cytotoxicity associated with the transfection reagent[291,292] and/or ssODN modification[293,294] used, which could result in reduced viability or cell death. Therefore, I investigated less cytotoxic transfection reagents and oligonucleotide modifications, as well as the influence of the number of mismatches between ssODN and target sequence. However, to establish the influence of these parameters on gene editing efficiency, quantitative experimental data was required. PCR-RFLP and DNA sequencing are unsuited for this purpose, as the read-out cannot be accurately quantified (previously discussed in section 3.2.2). Fortunately, ingenious model systems using mutated forms of enhanced green fluorescent protein (EGFP) have been developed.[182,228,242,245,249]

These systems allow the direct cell-based read-out of correction events, as successful gene editing is demonstrated by restoration of fluorescence, which can be visualized by microscopy or quantified with flow cytometry. Flow cytometry establishes exact numbers or percentages of fluorescent cells in mixed populations with a minimum sensitivity of around 0.05%. It can examine large numbers of cells and therefore distinction of small differences between two or more populations is possible.

Most mutated EGFP systems are based on a premature stop codon in the EGFP gene and are thus vulnerable to a low percentage of read-throughs. To solve this problem Tran *et al.*[228] introduced a Tyrosine to Serine mutation at amino acid 66 in the central chromophore region of EGFP that rendered the protein non-fluorescent, but preserved full-length expression. If this mutation is corrected by gene editing the EGFP fluorescence is restored. The CHO cell line kindly provided by Prof. Krauss is based on this system. These cells have multiple copies of the mutated pEGFP-C1 plasmid stably integrated and are therefore called CHO-mEGFP cells. An additional silent mutation at Lysine 45 in the EGFP permits exclusion of possible contamination with wild-type EGFP plasmid or cells by sequencing.

Figure 4-1 illustrates the differences between wild-type (wtEGFP), mutated (mEGFP) and gene-edited (edEGFP) sequences as well as the targeting oligonucleotides used here. Gene editing in these CHO-mEGFP cells has been widely studied by Prof. Krauss' group[242,251] and strand-bias for the non-transcribed strand was shown. Therefore, all targeting ssODNs were complementary to the non-transcribed strand (NT).

Using these ssODNs, I determined that Lipofectamine 2000 was the optimal transfection reagent for oligonucleotide-mediated gene editing in CHO-mEGFP cells. I also compared the efficiency of ssODNs with phosphorothioate (PTO) or locked nucleid acid (LNA) backbone modifications with unmodified (DNA) ssODN (Figure 4-2). PTO ssODNs have been the standard used for oligonucleotide-mediated gene editing in recent years, but their cytotoxicity is well-documented.[294] According to Andrieu-Soler *et al.*[184] LNA-modified ssODNs produce similar correction efficiencies to PTO-modified ssODNs, but exhibit reduced cytotoxicity. In my experiments, LNA- were less effective than PTO- modified ssODNs. Only DNA ssODNs had very low cytotoxicity, but unfortunately also induced very low levels of gene editing.

Andrieu-Soler *et al.*[184] also reported that gene-edited cells treated with LNA-modified ssODN expanded into colonies, whereas cells treated with PTO ssODN did not. They concluded that LNA modifications were less cytotoxic than PTO ones. On the other hand, Radecke *et al.*[241] stated that unmodified ssODNs, especially the shorter ones with only 13 or 19 nucleotides, were significantly less toxic than PTO-modified ssODNs. Also, Olsen *et al.*[251] have shown that the use of DNA ssODN leads to ~3 fold higher number of colonies than PTO ssODN. Therefore I compared viability of gene-edited cells treated with LNA- or PTO-modified ssODNs to unmodified ssODN. I found that both backbone modifications reduced viability of gene-edited cells in comparison to unmodified ssODN.

Initially, I used oligonucleotides with two mismatches (2A), as these were more effective in mediating gene editing than one mismatch ssODNs according to Olsen *et al.*[242] However, I needed to create a new restriction endonuclease recognition site for validation of oligonucleotide-mediated gene editing (chapter 6) and this required three mismatch (3A) ssODNs (Figure 4-1c). In contrast to the published data,[242] the mEGFP-3A-NT-PTO ssODN performed better than the mEGFP-2A-NT-PTO ssODN in my hands.

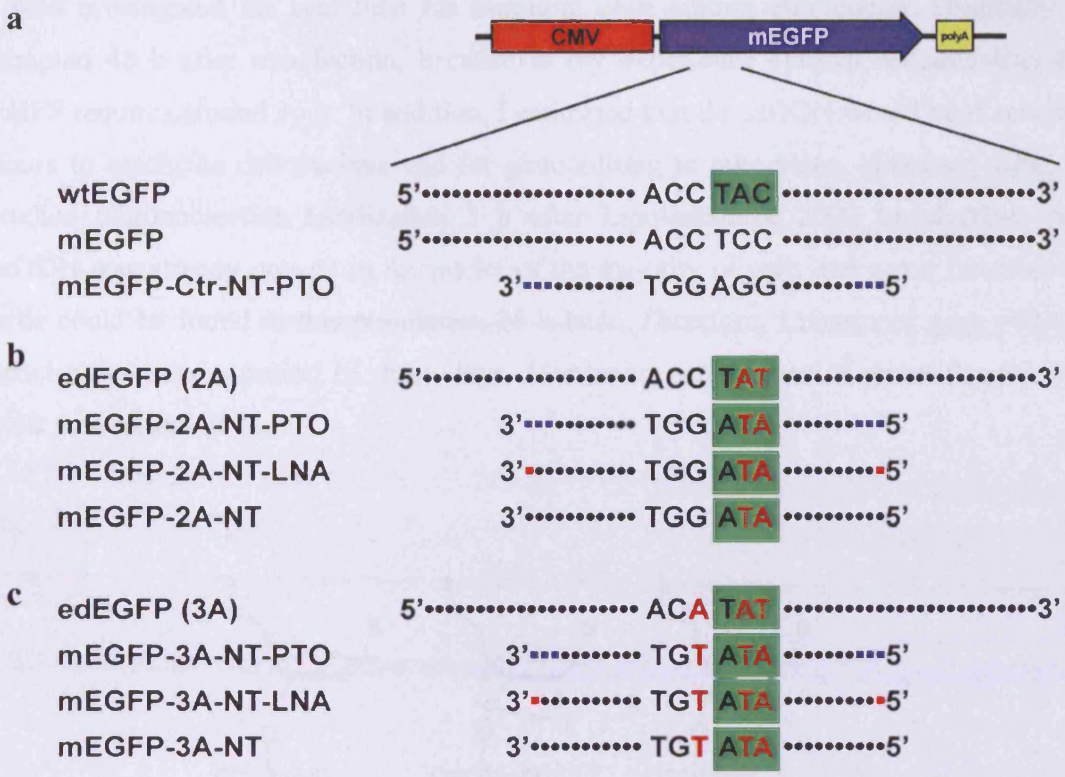


Figure 4-1 Schematic of the mEGFP system

a) wtEGFP has a Tyrosine (TAC) at amino acid 66, this has been mutated to a Serine (TCC) in mEGFP. All control ssODNs, such as mEGFP-Ctr-NT-PTO, are completely homologous to the mEGFP sequence.

b) Targeting ssODNs containing two mismatches (2A), such as mEGFP-2A-NT-PTO, change the Serine (TCC) at position 66 in mEGFP to an alternative Tyrosine codon (TAT). Thus, it is possible to distinguish wtEGFP and edEGFP (2A).

c) Targeting ssODNs containing three mismatches (3A), such as mEGFP-3A-NT-PTO, change the Serine (TCC) at position 66 in mEGFP to an alternative Tyrosine codon (TAT) and additionally introduce a silent mutation in the upstream Threonine codon (ACC to ACA). Thus, it is possible to distinguish wtEGFP and edEGFP (3A). All 3A ssODNs create a NdeI restriction endonuclease recognition site (5'- CA/TATG -3').

Green boxes denote sequences translating to functional green fluorescence; red letters draw attention to differences from the wtEGFP sequence; ● DNA backbone, ■ phosphorothioate (PTO) and ■ locked nucleic acid (LNA) backbone modifications of ssODNs. All ssODNs in this figure are complementary to the non-transcribed (NT) strand.

I also investigated the best time for sampling gene editing efficiencies. Originally I sampled 48 h after transfection, because in my experience optimal accumulation of EGFP requires around 36 h. In addition, I estimated that the ssODN would need several hours to reach the cell nucleus and for gene editing to take place. However, when I studied oligonucleotide localization 3 h after Lipofectamine 2000 transfection, the ssODN was already present in the nuclei of the majority of cells and green fluorescent cells could be found in this population 24 h later. Therefore, I measured gene editing efficiencies over a period of three days. Maximum percentages of green fluorescent cells were found at 24 h.

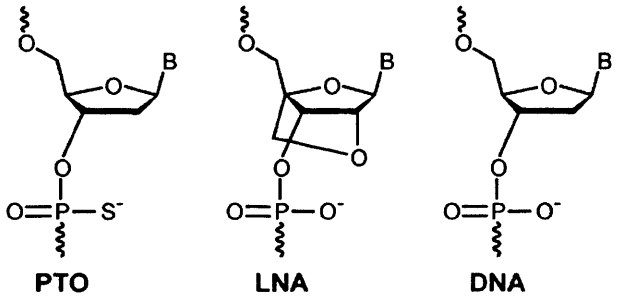


Figure 4-2 Chemical structure of oligonucleotide backbone modifications
In phosphorothioate (PTO) backbone modifications, sulphur replaces an oxygen anion in the linking phosphate group. Locked nucleic acid modifications are bridges connecting the 2' and 4' carbons of the 2-deoxyribose. This bridge "locks" the ribose in the 3'-endo structural conformation.

4.1 Effect of transfection reagents on gene editing efficiencies

4.1.1 Results

In this experiment I used the CHO-mEGFP cell line to compare the transfection reagents Lipofectamine 2000 and jetPEI (a commercially available linear PEI). Three wells each were transfected with 4 μg mEGFP-Ctr-NT-PTO or 4 μg mEGFP-2A-NT-PTO ssODN as described in section 2.2.1.4. Tissue culture plates containing jetPEI transfections were centrifuged for 5 min at 280 *g* to enhance cell-reagent contact and hence maximize transfection efficiency.[236]

Cells were trypsinised 48 h post-transfection, pelleted by centrifugation and resuspended in PBS. Propidium iodide (PI) was added at 50 μg/ml to permit dead cell discrimination. Samples were then examined by flow cytometry and $1x10^5$ events were acquired for each sample. Cells were gated based on forward scatter (FSC) versus side scatter (SSC) and the EGFP fluorescence of each cell was plotted against its PI fluorescence on a density plot.

As shown in Figure 4-3a left panel, cells treated with the fully complementary control ssODN (mEGFP-Ctr-NT-PTO) did not express EGFP (x-axis), although some of the cells had taken up PI (y-axis). When CHO-mEGFP cells were transfected with the targeting ssODN (mEGFP-2A-NT-PTO), a number of cells clearly expressed EGFP (Figure 4-3a right panel).

Based on the localization of the cells treated with control ssODN on the density plot, a region was drawn around EGFP positive cells (e.g. R1 in Figure 4-3a). The mean and standard deviation of the percentages of green fluorescent cells in this region were displayed in a column chart (Figure 4-3b). Samples in the same experiment were analysed using identical machine and region settings to allow comparison.

Results clearly show that cells transfected with Lipofectamine 2000 exhibit a far higher percentage of green fluorescent cells than those transfected with jetPEI. Two-way ANOVA with Bonferroni post-tests yielded a *P* value of less than 0.001 for Lipofectamine 2000 versus jetPEI for cells transfected with mEGFP-2A-NT-PTO.

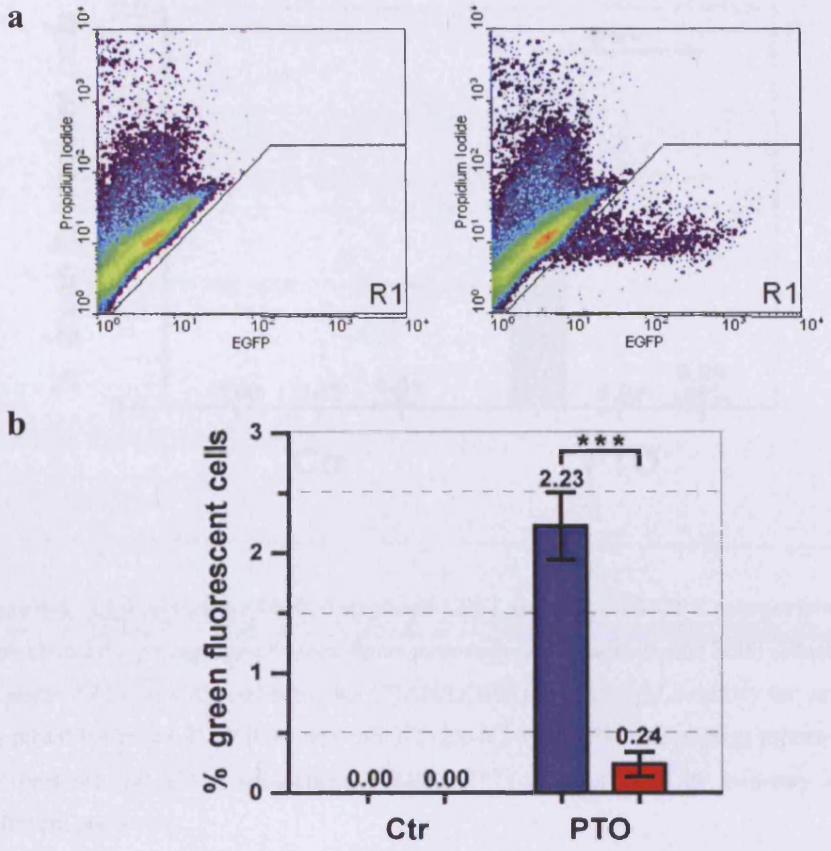


Figure 4-3 Lipofectamine 2000 and jetPEI comparison

a) EGFP versus PI density plots of cells transfected with mEGFP-Ctr-NT-PTO (left) and mEGFP-2A-NT-PTO (right) using Lipofectamine 2000. Region 1 (R1) was set using the cells transfected with mEGFP-Ctr-NT-PTO as negative control.

b) Graph shows the percentage of green fluorescent cells in R1 for Lipofectamine 2000 (Blue) and jetPEI (red) for cells transfected with mEGFP-Ctr-NT-PTO (Ctr) and mEGFP-2A-NT-PTO (PTO). Numbers represent means and error bars are the standard deviation (SD) of triplicates. $P<0.001$ (***) as determined by two-way ANOVA with Bonferroni post-tests.

In a similar experiment I compared deacylated LPEI and CDAN/DOPE with Lipofectamine 2000. Fully deacylated LPEI has been shown to have increased efficiency and less toxicity than the commercial LPEI from which it was derived in plasmid transfections.[140] CDAN/DOPE also demonstrated increased down-regulation and reduced toxicity compared to Lipofectamine 2000 in siRNA transfections.[148]

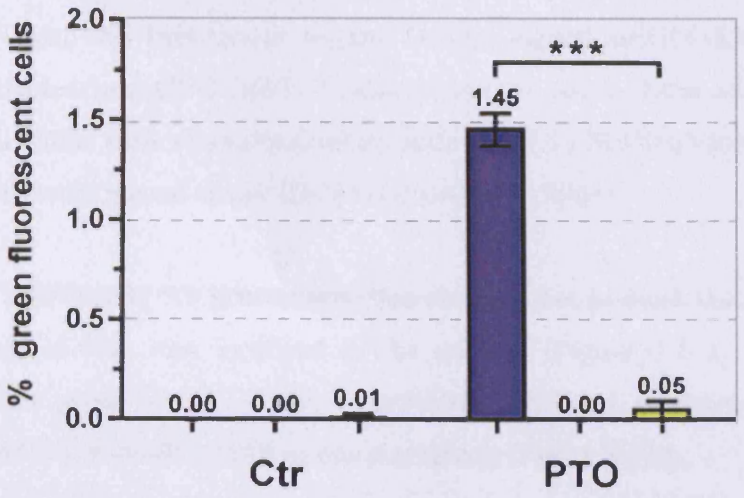


Figure 4-4 Lipofectamine 2000, deacylated LPEI and CDAN/DOPE comparison
Graph shows the percentage of green fluorescent cells for Lipofectamine 2000 (Blue; left column), deacylated LPEI (middle column) and CDAN/DOPE (yellow; right column) for cells transfected with mEGFP-Ctr-NT-PTO (Ctr) and mEGFP-2A-NT-PTO (PTO). Numbers represent means and error bars are the SD of triplicates. $P<0.001$ (***) as determined by two-way ANOVA with Bonferroni post-tests.

Deacylated LPEI was transfected in the same way as normal LPEI with N/P ratio of 3 (section 2.2.1.4). CDAN/DOPE transfection was performed as described in section 2.2.1.4. Cells were prepared for flow cytometry and results evaluated as described above.

As seen in Figure 4-4 Lipofectamine 2000 is significantly more effective than either deacylated LPEI or CDAN/DOPE. Extensive studies using the biopolymers tri-methylated chitosan and atelocollagen for transfection did not lead to any green fluorescent cells (data not shown). In both cases two different batches of the reagent were tried and for atelocollagen the protocol described by Nakamura *et al.*[145] was followed closely. Optimisation of ssODN electroporation conditions by Wenyan Li, a visiting scientist, yielded green fluorescent cells but less than Lipofectamine 2000 and electroporation-associated cytotoxicity was higher (data not shown). Thus Lipofectamine 2000 was the most effective transfection reagent for gene editing.

To further evaluate the efficiency of Lipofectamine 2000, I studied the nuclear delivery of ssODN with this transfection reagent. A Cy3-tagged mEGFP-2A-NT-PTO ssODN was transfected into CHO-mEGFP cells grown on cover slides using Lipofectamine 2000. Cell nuclei were counterstained by addition of 5 µM Draq5 for 10 min before the cover slides were placed upside down on microscopy slides.

Confocal microscopy 3 h post-transfection showed that in more than 90% of cells, the majority of ssODN was localized in the nucleus (Figure 4-5 a). When cells were permitted to grow for 24 h after transfection, confocal microscopy confirmed the presence of EGFP positive cells in this population (Figure 4-5b).

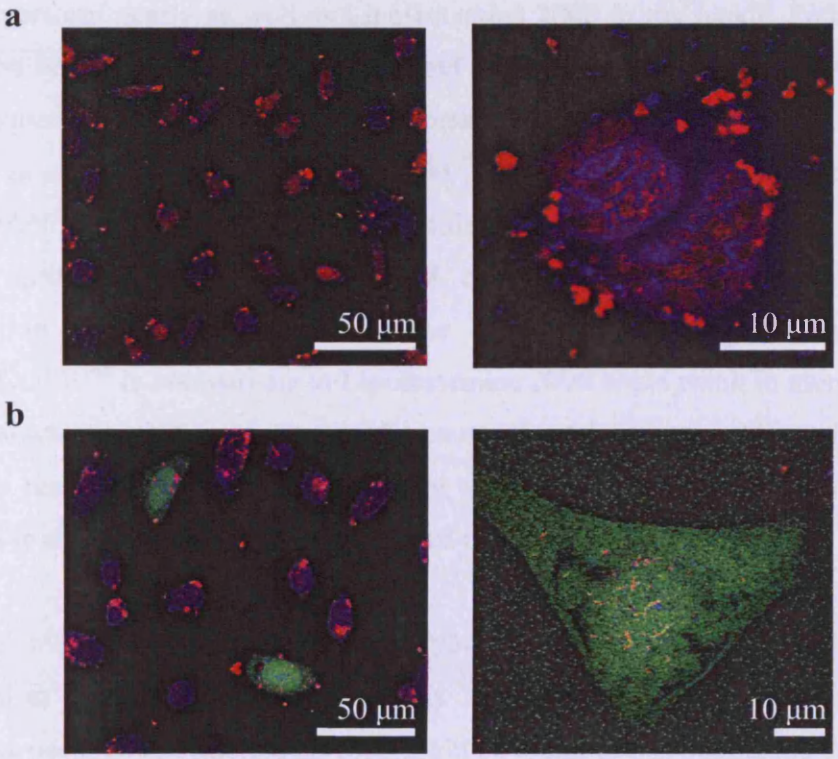


Figure 4-5 Confocal microscopy of gene-edited cells

CHO-mEGFP cells transfected with Cy3-mEGFP-2A-NT-PTO (pink) were examined (**a**) 3 h or (**b**) 24 h post-transfection. Cell nuclei were counterstained with Draq5 (blue), EGFP in green.

4.1.2 Discussion

The experiments described in this section clearly demonstrate that flow cytometry is an easy and quick way to determine exact gene editing efficiencies in mutated reporter gene systems. In contrast to PCR-RFLP it is a direct cell-based read-out and thus the possibility of artefacts during analysis can be completely excluded. This system was used to evaluate gene editing efficiencies associated with several transfection reagents. Lipofectamine 2000 has quickly become the standard transfection reagent used in oligonucleotide-mediated gene editing in recent years, as it guarantees nuclear delivery of a high proportion of the ssODN and this was completely justified by my data.

LPEI has also been used extensively for oligonucleotide-mediated gene targeting, but did not perform nearly as well as Lipofectamine 2000 in my hands. Perhaps polyplex formation between the 27 base ssODN and LPEI is not as strong as it is with longer DNA sequences. LPEI has been used successfully to transfect 20 - 23 base siRNAs, but was not as effective as Lipofectamine 2000.[295] Deacylated LPEI was not effective at all and CDAN/DOPE did not mediate significant correction above background levels. Further optimisation of the lipid to DNA ratio for CDAN/DOPE might yield better transfection efficiencies. However, the slower cellular uptake described for CDAN/DOPE[148] in comparison to Lipofectamine 2000 could result in more degradation of the targeting ssODN and thus be the cause of reduced gene editing efficiency. The negative results from atelocollagen were particularly disappointing as it has been reported to mediate gene editing in ~10% of treated HepG2 cells.[145]

Confocal microscopy showed that the Cy3-tagged mEGFP-2A-NT-PTO ssODN was localized in the nucleus of >90% of cells 3 h after transfection when Lipofectamine 2000 was used. Green fluorescent cells could be found in this population 24 h later. It is interesting that transfection of ssODN into the nucleus using Lipofectamine 2000 does not seem to require cellular division. It has been described that successful plasmid transfection is dependent on mitosis and fragmentation of the nuclear membrane, though a small percentage of non-dividing cells is also transfected.[296] However, the presence of oligonucleotide in the nucleus of the majority of unsynchronized CHO-mEGFP cells only 3 h after transfection suggests that this is not the case for ssODN transfection. Lipofectamine 2000 proved to be the most effective transfection reagent for oligonucleotide-mediated gene editing and was therefore used in all further studies.

4.2 Effect of ssODN backbone modifications on gene editing efficiency

4.2.1 Results

In similar experiments, I compared different oligonucleotide backbone modifications. The percentages of green fluorescent cells resulting from treatment with PTO- and LNA-modified as well as unmodified ssODN were compared. CHO-mEGFP cells were transfected with 4 μg of the various ssODNs and prepared for flow cytometry 48 h post-transfection as described previously.

Cell death associated with mEGFP-2A-NT ssODN transfection was minimal as observed under the microscope; unfortunately, correction efficiency was very low too (Figure 4-6a). The LNA-modified ssODN resulted in a lower correction efficiency than the PTO ssODN (Figure 4-6a). One-way ANOVA with Bonferroni post-tests yielded $P<0.05$ for this difference. All other differences except between mEGFP-Ctr-NT-PTO and mEGFP-2A-NT had $P<0.001$.

Percentages obtained in a second independent experiment varied significantly from the first as determined by two-way ANOVA. However, when the percentages from both experiments were normalized with the mean of the highest percentage (PTO) set as 100%, it became clear that the differences between the two experiments did not invalidate the differences between PTO- and LNA-modified ssODNs (Figure 4-6b). In fact, all differences except between mEGFP-Ctr-NT-PTO and mEGFP-2A-NT had $P<0.001$ as determined by one-way ANOVA with Bonferroni post-tests (Table 4-1).

Furthermore, the difference between PTO- and LNA-modified ssODNs was also seen in comparisons of three mismatch (3A) ssODNs (section 4.5.1).

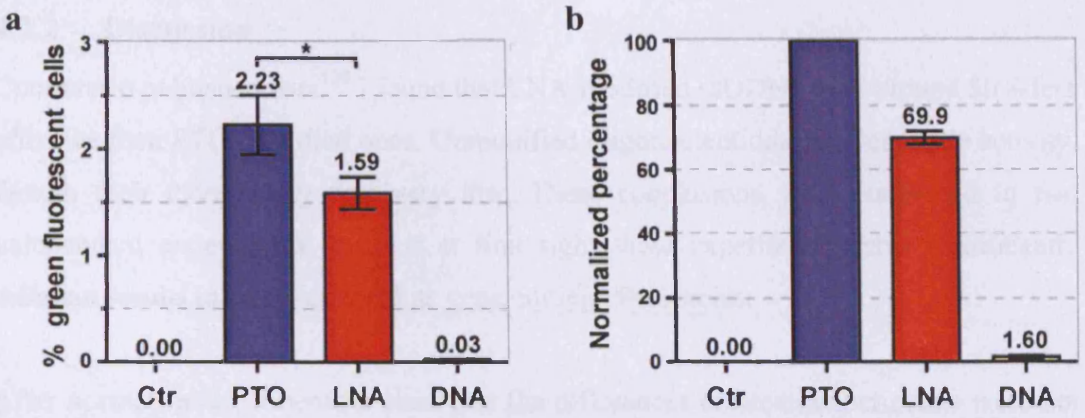


Figure 4-6 Effect of ssODN backbone modifications on gene editing efficiency

a) Percentage of green fluorescent cells in cells treated with mEGFP-Ctr-NT-PTO (Ctr; black) mEGFP-2A-NT-PTO (PTO; blue), mEGFP-2A-NT-LNA (LNA; red) and mEGFP-2A-NT (DNA; yellow). Numbers represent means and error bars are the SD of triplicates. $P<0.05$ (*) as determined by one-way ANOVA with Bonferroni post-tests.

b) Normalized percentage of green fluorescent cells in cells treated with mEGFP-Ctr-NT-PTO (Ctr; black), mEGFP-2A-NT-PTO (PTO; blue), mEGFP-2A-NT-LNA (LNA; red) and mEGFP-2A-NT (DNA; yellow). Numbers represent means and error bars are the SD of two independent experiments each with triplicates.

ssODN	*P* value
Ctr vs PTO	***
Ctr vs LNA	***

ssODN	*P* value
Ctr vs DNA	ns
PTO vs LNA	***

ssODN	*P* value
PTO vs DNA	***
LNA vs DNA	***

Table 4-1 Effect of ssODN backbone modifications on gene editing efficiency

Statistical significance of differences between normalized gene editing efficiencies mediated by mEGFP-Ctr-NT-PTO (Ctr), mEGFP-2A-NT-PTO (PTO), mEGFP-2A-NT-LNA (LNA) and mEGFP-2A-NT (DNA) in Figure 4-6b. *P* values for differences between ssODNs were determined by one-way ANOVA with Bonferroni post-tests. *** $P<0.001$; ns $P>0.05$ grey text.

4.2.2 Discussion

Contrary to published data[184] I found that LNA-modified ssODNs were around 30% less effective than PTO-modified ones. Unmodified oligonucleotides had very little activity, though their cytotoxicity was very low. These conclusions were confirmed in two independent experiments, even if at first sight these experiments gave significantly different results in terms of absolute gene editing efficiencies.

After normalization it became clear that the differences between experiments were not caused by different results, but rather by unrelated factor(s) that influenced all samples in the same experiment. Two possibilities spring to mind: the general condition of the transfected cells or differing efficiencies of the Lipofectamine 2000 complex formation.

Whatever the cause, the small standard deviations gained after normalization of only two experiments together with data from 3A ssODNs (section 4.5.1) are convincing evidence that the observed difference between PTO- and LNA-modified ssODNs is not due to experimental artefacts.

4.3 Effect of ssODN backbone modifications on viability of edited cells

4.3.1 Results

CHO-mEGFP cells were transfected as described before and cells were sorted according to EGFP status 24 h after transfection as explained in section 2.2.1.9. Single cells were seeded in each well of 96-well plates and incubated for 10 d. Wells showing colony formation under the microscope were scored as positive and percent viability was determined as percent of colony forming wells out of the total number of wells seeded.

	Colony formation		Total	% viable	Viability ratio
	Yes	No			
PTO					
EGFP -	75	404	479	15.66	6:1
EGFP +	47	1852	1899	2.47	
LNA					
EGFP -	93	387	480	19.38	13:1
EGFP +	26	1785	1811	1.44	
DNA					
EGFP -	124	356	480	25.83	4:1
EGFP +	31	448	479	6.47	

Table 4-2 Effect of ssODN backbone modifications on viability of edited cells

Total numbers (from three independent experiments) of cells treated with mEGFP-3A-NT-PTO (PTO), mEGFP-3A-NT-LNA (LNA) or mEGFP-3A-NT (DNA) that either formed colonies or did not after being sorted as EGFP negative or positive. Percent of viable cells is the proportion of colony forming cells out of the total number of cells for each condition. Viability ratio between the percentage of EGFP- and EGFP+ viable cells for each backbone modification.

Table 4-2 summarizes the results from three independent experiments. There is a clear reduction in viability between EGFP negative and positive cells treated with either LNA- or PTO-modified, as well as DNA ssODNs as demonstrated by the variances in viability ratio (13:1 for LNA, 6:1 for PTO and 4:1 for DNA). This reduction is less pronounced in cells treated with unmodified oligonucleotide, but is highly statistically significant (analysis using contingency tables with Fisher's exact test; $P<0.001$) in all cases. Interestingly, there is also a decrease in viability of EGFP negative sorted cells of LNA or PTO treated cells compared to DNA treated cells.

Notably, only three colonies grown from cells treated with unmodified oligonucleotide and originally sorted as EGFP positive showed partial green fluorescence two weeks after sorting. Data shown in section 6.2 confirmed successful gene editing in two of them.

4.3.2 Discussion

Treatment with any ssODN, whether modified or not, influenced viability of non-edited EGFP- cells, presumably because transfection of oligonucleotide per se was cytotoxic to some degree. This agrees well with previously reported evidence that single-stranded DNA induces phosphorylation of histone 2AX, which is an indicator for activation of the ATM damage response pathway.[297]

However, it is evident that there are additional reductions in viability associated with successful gene editing and that these decreases are dependent on the specific nature of the oligonucleotide backbone. Thus, both LNA and PTO modifications led to a more pronounced reduction in viability of edited cells than the unmodified DNA backbone. This should perhaps not come as a surprise. Olsen *et al.*[251] reported that the presence of 5' and/or 3' octyl or hexaethylene glycol blocking groups on the targeting ssODN reduced gene editing by up to 80%. Radecke *et al.*[259] confirmed these results using ssODNs with non-canonical termini and also showed that targeting oligonucleotides were physically integrated into the genome.

If the oligonucleotide is integrated into the genome while still containing the backbone modifications, these abnormalities in the DNA could make it more difficult for affected cells to progress through the cell cycle. Perhaps the difference in the viability ratio of gene-edited cells treated with LNA (13:1) in comparison to PTO (6:1) reflects the fact that the occasional sulphur atom instead of oxygen can be tolerated better than the fairly constrained formation of deoxyribose imposed by the LNA modification.

It is also conceivable that the remaining viability in PTO and LNA treated cells is based on partial degradation of the ssODNs by endonucleases that would result in a shorter remnant consisting of pure DNA. In this case the differences in viability between PTO and LNA could be explained by a differential susceptibility to endonucleases. However gene-edited cells treated with DNA ssODN also show a significant reduction in viability. Thus, the gene editing event itself must play a part in the observed cytotoxicity.

4.4 Effect of the number of mismatches

4.4.1 Results

For the validation of oligonucleotide-mediated gene editing a new restriction endonuclease recognition site had to be created with the targeting ssODNs. However, the only feasible restriction site required three changes of the mEGFP sequence. Olsen *et al.*[242] have previously evaluated the influence of the number of mismatches on gene editing efficiency (Figure 4-7a and Table 4-3) and found that a two mismatch ssODN generated the highest percentage of green cells.

When I compared the gene editing efficiency of mEGFP-2A-NT-PTO (2A) with mEGFP-3A-NT-PTO (3A), I found that the 3A ssODN was significantly *more* effective (Figure 4-7b and Table 4-3).

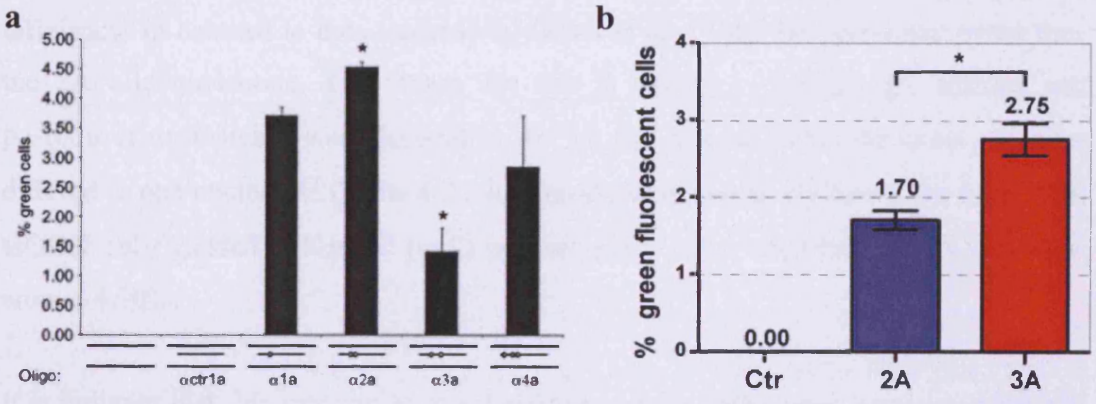


Figure 4-7 Effect of the number of mismatches on gene editing efficiency

a) Figure taken from Olsen *et al.*[242] for comparison.

b) Percent of green fluorescent cells in cells treated with mEGFP-Ctr-NT-PTO (Ctr; black), mEGFP-2A-NT-PTO (2A; blue) and mEGFP-3A-NT-PTO (3A; red). Numbers represent means and error bars are the SD of two independent experiments each with triplicates. $P<0.05$ (*) as determined by one-way ANOVA with Bonferroni post-tests.

ssODN	sequence 3' – 5'	% green cells
αctr1a	TGG GAC TGG AGG CCG CAC GTC ACG AAG	0.00
α1a	TGG GAC TGG A**T**G CCG CAC GTC ACG AAG	~3.65
α2a	TGG GAC TGG A**TA** CCG CAC GTC ACG AAG	~4.50
α3a	TGG GAC TG**C** A**T**G CCG CAC GTC ACG AAG	~1.45
α4a	TGG GAC TG**C** A**TA** CCG CAC GTC ACG AAG	~2.90
2A	TGG GAC TGG A**TA** CCG CAC GTC ACG AAG	1.70±0.12
3A	TGG GAC TG**T** A**TA** CCG CAC GTC ACG AAG	2.75±0.21

Table 4-3 Effect of the number of mismatches on gene editing efficiency

Sequence of ssODNs and resulting percentage of green fluorescent cells; αctr1a, α1a, α2a, α3a and α4a used by Olsen *et al.*;[242] 2A and 3A used here. Note that α2a and 2A have the same sequence.

4.4.2 Discussion

My experiments demonstrated an effect of the number of mismatches on gene editing efficiency. In contrast to data reported by Olsen *et al.*,[242] the 3A performed better than the 2A oligonucleotide. The reason for this is unclear. Although the number and position of mismatches was identical in the 3A and α4a ssODNs, the exact sequence differed in one nucleotide (Table 4-3). It is noteworthy that in my hands the 2A (=α2a) ssODN only yielded 1.70±0.12 (n=2) percent green cells, whereas Olsen's data were around 4.50%.

It is unlikely that this incongruence is based on quality differences between individual ssODNs as I used different batches of ssODN for the two independent experiments. Perhaps there is a quality difference in ssODNs from different suppliers (Integrated DNA Technologies in my case and MWG Biotech or Eurogentec in Olsen's). It is more likely though that there are slight differences between investigators in transfection methodology or flow cytometer setup resulting in inconsistent rates of green positive cells. These factors, however, would influence all experiments done by the same investigator. Thus, relative comparisons between experiments by different investigators would be valid.

It is possible that the single nucleotide sequence difference between the 3A and α4a ssODNs could result in differential recognition of the mismatch by the mismatch repair system. Dekker *et al.*[260] showed that in MSH2 positive embryonic stem cells replacement of a four-nucleotide sequence was more effective than one, two or three nucleotide exchanges. However, a slight alteration in positioning and therefore the sequence of the four mismatch substitution did not lead to a similar increase.

They also demonstrated that two three-nucleotide substitutions with the same sequence resulted in significantly different gene editing efficiencies, even though they were only three bases apart, possibly because of the differences in the target sequence. Dekker *et al.*[260] were unable to derive rules to determine which sequence substitutions were more effective, but their data hint that specific nucleotide differences between target and ssODN could be responsible for differential recognition by the mismatch repair system.

However, this differential recognition would not explain the differences in gene editing efficiencies between the 2A and α2a ssODNs.

4.5 Effect of sampling time

4.5.1 Results

To investigate the effect of sampling time on gene editing efficiencies, CHO-mEGFP cells were transfected as previously described with 4 µg of mEGFP-Ctr-NT-PTO (Ctr), mEGFP-3A-NT-PTO (PTO), mEGFP-3A-NT-LNA (LNA) and mEGFP-3A-NT (DNA) ssODN. At 24, 48 and 72 h triplicate samples were prepared for flow cytometry. Two independent experiments demonstrated significantly different gene editing efficiencies for samples from the same time point as determined by two-way ANOVA. Therefore, the 24 h PTO sample mean of each experiment (1.06±0.04 and 3.36±0.20 percent respectively; ±SD of triplicates) was set as 100%; all other sample means were normalized to this value (Figure 4-8). The normalized percentages from both experiments were compared in a two-way ANOVA with repeated measures and Bonferroni post-tests (Table 4-4).

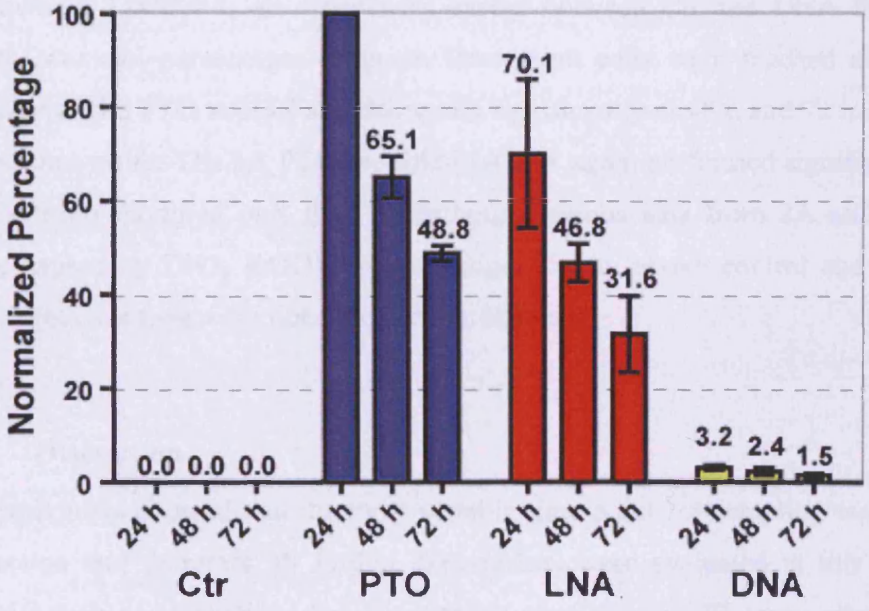


Figure 4-8 Effect of sampling time

Normalized percentage of green fluorescent cells in cells treated with mEGFP-Ctr-NT-PTO (Ctr; black), mEGFP-3A-NT-PTO (PTO; blue), mEGFP-3A-NT-LNA (LNA; red) and mEGFP-3A-NT (DNA; yellow) 24, 48 and 72 h after transfection. Numbers represent means and error bars are the SD of two independent experiments each with triplicates.

ssODN	*P* value
Ctr vs PTO	*** for all time points
Ctr vs LNA	*** for all time points
Ctr vs DNA	ns for all time points
PTO vs LNA	*** for 24 h * for 48 and 72 h
PTO vs DNA	*** for all time points
LNA vs DNA	*** for all time points

time	*P* value
Ctr	ns for all time points
DNA	ns for all time points
PTO	*** for 24 h vs. 48 h ns for 48 h vs 72 h *** for 24 h vs. 72 h
LNA	** for 24 h vs. 48 h ns for 48 h vs 72 h *** for 24 h vs. 72 h

Table 4-4 Effect of sampling time

Statistical significance of differences between mEGFP-Ctr-NT-PTO (Ctr), mEGFP-3A-NT-PTO (PTO), mEGFP-3A-NT-LNA (LNA) and mEGFP-3A-NT (DNA) 24, 48 and 72 h after transfection. *P* values for differences between ssODNs and sampling time points were determined by two-way ANOVA with repeated measures and Bonferroni post-tests. *** $P<0.001$; ** $P<0.01$; * $P<0.05$; ns $P>0.05$ grey text.

As shown in Table 4-4, all differences except between Ctr and DNA have at least $P<0.05$. Maximal percentages of green fluorescent cells were reached at 24 h after transfection with PTO ssODN and decreased significantly at 48 h and 72 h compared to the first time point. The 3A PTO-modified ssODN again performed significantly better than the LNA-modified one, thus confirming previous data from 2A ssODNs. Gene editing caused by DNA ssODN was not significantly above control and normalized percentages over time were not significantly different.

4.5.2 Discussion

This experiment showed that the most suitable time point for sampling was 24 h after transfection and therefore all further experiments were evaluated at this time point. Several possible explanations for decreases in gene editing efficiency after 24 h have already been discussed (section 3.2.2) and therefore will not be repeated here. Results with DNA ssODN again were not significant due to the low percentages of green fluorescent cells derived from treatment with this oligonucleotide. Thus, the results highlight the fact that even if viability of PTO- and LNA-treated gene-edited cells is reduced, the higher gene editing efficiencies associated with modified ssODNs are required to examine the influence of small differences in treatment conditions.

CHAPTER 5: TOWARDS INCREASING GENE EDITING EFFICIENCY AND SELECTING EDITED CELLS

5.1 Repeated targeting

5.1.1 Introduction

Repeated targeting of the same population has been reported to enhance chromosomal gene editing efficiencies with LNA ssODN in HEK293T cells.[184] According to Andrieu-Soler *et al.*[184] three consecutive transfections at 24 h intervals (t=0, 1 and 2 d) increased the percent of green fluorescent cells compared to a single (t=0) and two successive (t=0 and 1 d) control transfections when measured 3 days after the last transfection (t=5 d). However, as I have shown in the previous chapter (section 4.5.1), the fraction of green fluorescent cells decreases significantly over as little as 24 h.

Consequently, the increase seen by Andrieu-Soler *et al.*[184] with triple transfection could simply be due to the different time delays between the single (5 d delay), double (4 d delay) and triple (3 d delay) transfections and measurement. Thus, I evaluated whether triple transfection would increase the percentage of green fluorescent cells when single and double control transfections with identical time delays between targeting and measurement were included in the experiment. Consistent with my previous tentative data with HepG2 cells (section 3.4.1), I found that repeated targeting did not increase gene editing efficiencies in CHO-mEGFP cells under these circumstances.

5.1.2 Results

CHO-mEGFP cells were transfected three times with mEGFP-2A-NT-PTO (Figure 5-1a sample C) or mEGFP-2A-NT ssODN (Figure 5-1b sample C) as previously described at t=0, 2 and 5 d. Samples with single transfections at t=0 (E) or t=5 d (A) as well as double transfections at t=0 and 2 d (D) or t=2 and 5 d (B) were included as gene editing efficiency controls. The corresponding mEGFP-Ctr-NT-PTO and mEGFP-Ctr-NT ssODNs were also transfected and used to set gates for flow cytometric analysis. Cells were harvested and analysed 1 d after the last transfection (i.e. t=6 d).

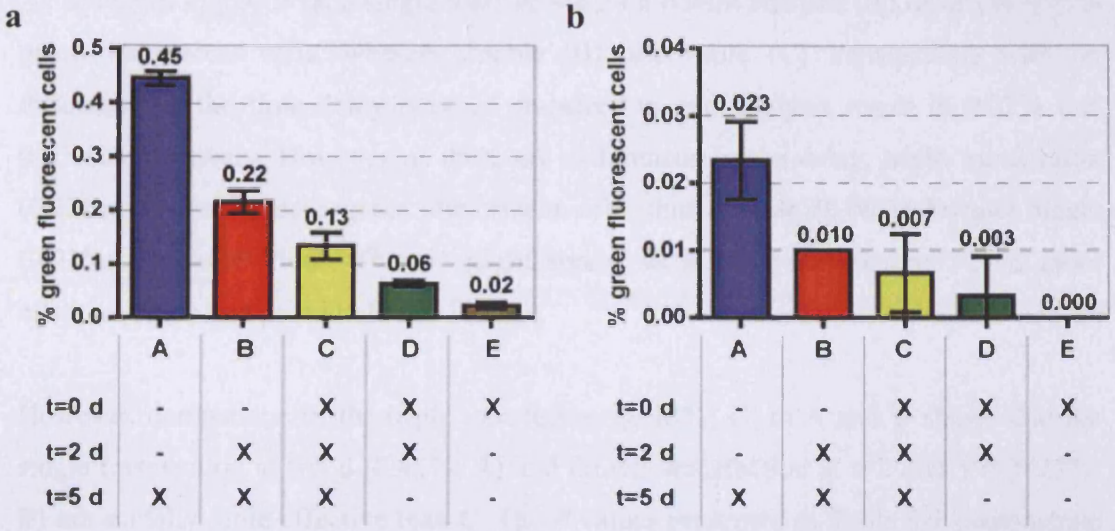


Figure 5-1 Repeated targeting

Flow cytometry results of CHO-mEGFP cells transfected with (**a**) mEGFP-2A-NT-PTO or (**b**) mEGFP-2A-NT at the time-points indicated with an X. Numbers represent means and error bars are SD of triplicates.

Bonferroni post-tests a)	*P* value
A vs B	***
A vs C	***
A vs E	***
B vs C	**
B vs D	***
C vs D	**
C vs E	***

Bonferroni post-tests b)	*P* value
A vs B	*
A vs C	*
A vs E	***
B vs C	ns
B vs D	ns
C vs D	ns
C vs E	ns

Table 5-1 Repeated targeting

Statistical significance of differences between single, double and triple transfected samples. *P* values for differences between samples were determined by one-way ANOVA and Bonferroni post-tests. *** $P<0.001$; ** $P<0.01$; * $P<0.05$; ns $P>0.05$ grey text.

As shown in Figure 5-1a, a single transfection 24 h before analysis (A) results in 0.45% green fluorescent cells, whereas double (B) and triple (C) transfections with no difference in the time delay between transfection and analysis result in 0.22% and 0.13% respectively. However, if there are differences in the delay, triple transfection (0.13%; C) yields more green fluorescent cells than double (0.06%; D) and single (0.02%; E) transfections. Thus it might appear as if triple transfection (C) is more effective when compared to D and E only.

However, comparison of the triple transfection (0.13%; C) to A and B shows that the single transfection at t=5 d (0.45%; A) and double transfection at t=2 and 5 d (0.22%; B) are actually more effective than C. The *P* values presented in Table 5-1 demonstrate that at least for the PTO ssODN, these differences in gene editing efficiencies between samples are highly significant. Similar results were obtained with the mEGFP-2A-NT ssODN, although gene editing efficiencies were very low. Consequently, differences did not reach statistical significance in all comparisons, though A (0.023%) vs C (0.007%) was still significant ($P<0.05$).

5.1.3 Discussion

Similar to data reported by Andrieu-Soler *et al.*;[184] I found that triple transfection results in higher gene editing efficiencies than double or single transfection if the time delay between transfection and analysis is different. This was the case even though the cells I used here (CHO-mEGFP) are different from the cells used by Andrieu-Soler (HEK293T with blue fluorescent protein). However, when the delay is the same, single transfection is more effective than either double or triple transfection. Thus, it is clear that the apparent increase in gene editing efficiency through triple transfection seen by Andrieu-Soler *et al.*[184] was due to an experimental artefact.

Furthermore, the published data were from one experiment only, where the single transfection gave comparable gene editing efficiencies to previous experiments and the triple transfection gave a 2.5 times higher outcome. This discrepancy is easily explained by factors amply demonstrated in previous chapters of this thesis; even when every effort is made to standardize cell treatment and transfection method, significant differences in gene editing efficiencies between repeat experiments happen.

In my opinion this does not reflect the fact that oligonucleotide-mediated gene editing efficiencies per se are widely variable, but rather that many factors cannot conveniently be controlled for when working with cell cultures. Factors such as the general state of the cells in culture or the efficiency of complex formation between transfection reagent and DNA can have a major influence on results. Even though these factors can be investigated in detail if required, the actual time and expenditure involved precludes this option in practice. Thus, it is imperative to include accurate experimental controls, interpret resulting data carefully and repeat experiments at least once.

Although the data represented here are also based on only one experiment each for mEGFP-2A-NT-PTO and mEGFP-2A-NT, the fact that two different ssODNs gave similar results is more convincing than a simple repeat would be. Thus, triple transfection does not increase gene editing efficiencies compared to single transfection in CHO-mEGFP cells. This observation seems counterintuitive. Repeated transfection should in theory enrich for gene-edited cells, as the large majority of non-edited cells left over after the first transfection should then be targetable in the second and third transfections. Two main possibilities could explain why this is not the case.

Firstly, only a certain percentage of the whole population might be susceptible to gene editing. Considering that nearly all CHO-mEGFP cells show large amounts of fluorescently-labelled ssODN in the nucleus with Lipofectamine 2000 transfection (see section 4.1.1), but less than 2.5% of cells are gene-edited, this could well be the case. However, results presented in a later chapter (section 7.2.1) demonstrate that a much higher percentage of CHO-mEGFP cells can be gene-edited under certain circumstances. Thus, this explanation is unlikely.

Secondly, and the more likely explanation, is cytotoxicity associated with repeated administration of ssODN/Lipofectamine 2000 complexes. This explanation is certainly supported by routine visual observations of cells after transfection, as well as data from propidium iodide staining not shown here. Furthermore, when CHO-mEGFP cells were transfected with siRNA, followed by ssODN 24 h later using Lipofectamine 2000, gene editing efficiency decreased significantly even with control siRNA (James Booth and Ioannis Papaioannou, personal communication).

5.2 Plasmid co-transfection

5.2.1 Introduction

Another attempt to improve gene editing efficiencies, and perhaps enrich for edited cells, was based on a report by Igoucheva *et al.*[290] that co-transfection of heterologous plasmid with targeting ssODN led to a 5-fold increase over targeting ssODN alone. Instead of a random plasmid, I chose to use a red fluorescent protein (RFP) expressing plasmid. I hoped that I would be able to use RFP expression to enrich for cells that had received a high amount of not only plasmid but ssODN as well.

Such enrichment is possible for co-transfections of two plasmids if one of the plasmids expresses a fluorescent protein.[288,289] Also, dose-response has been clearly established for oligonucleotide-mediated gene editing.[184,245] Thus, I theorized that cells which expressed detectable RFP would be enriched in gene-edited cells. However, my results showed that this was not the case. Interestingly there was a slight but significant increase in gene editing efficiency with co-transfection compared to ssODN transfection alone, even though I was unable to identify cells that were both RFP and EGFP positive.

5.2.2 Results

CHO-mEGFP cells were transfected with 3 μg of mEGFP-3A-NT-PTO together with decreasing amounts (2, 1 and 0.5 μg) of pRFP (Figure 5-2; samples B, C and D). Control samples included cells treated with 5 μg mEGFP-Ctr-NT-PTO (A), 3 μg mEGFP-3A-NT-PTO (E) or 2 μg pRFP (F). A ssODN with no sequence homology to the CHO-mEGFP genome (huApoA-I_M-49NT-PTO; Scr) was added where necessary to keep the total amount of transfected DNA identical between samples.

Cells were transfected as described before with a DNA to Lipofectamine 2000 ratio of 1:2, though the total amount of transfected DNA was 5 μg, not 4 μg as usual with ssODN transfection alone. The total amount of transfected DNA was limited by the fact that Lipofectamine 2000 was highly cytotoxic at more than 10 μl per well.

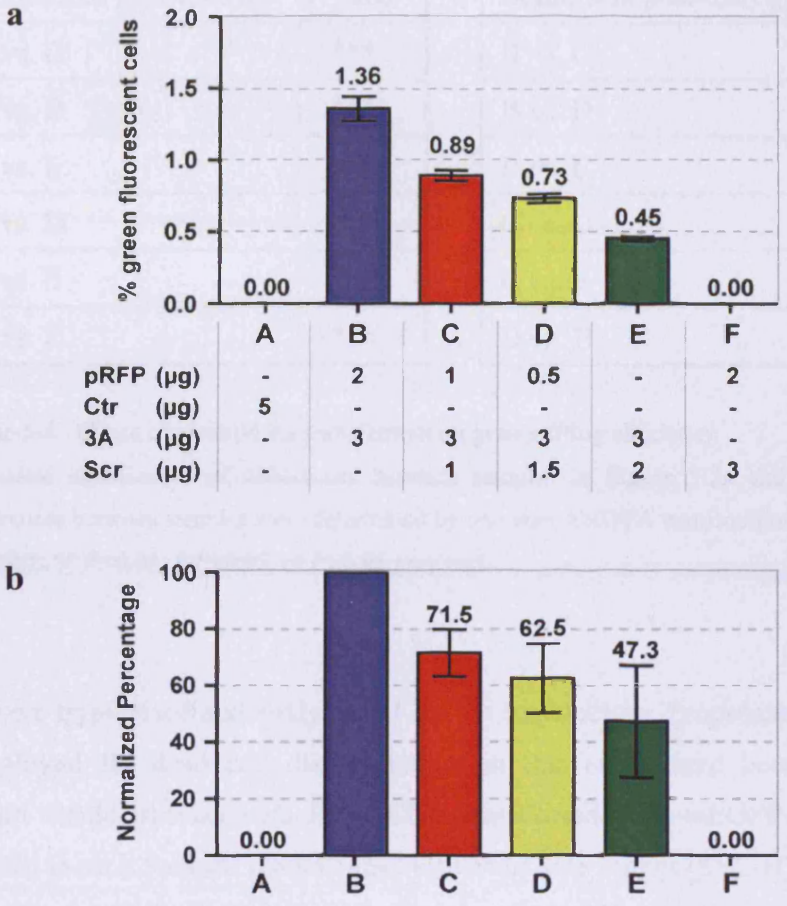


Figure 5-2 Effect of plasmid co-transfection on gene editing efficiency

a) Samples transfected with the indicated amounts of pRFP, mEGFP-Ctr-NT-PTO (Ctr), mEGFP-3A-NT-PTO (3A) and huApoA-I_M-49NT-PTO (Scr). Numbers represent means and error bars are the SD of triplicates.

b) Normalized percentage of green fluorescent cells achieved by setting the means of sample B in two independent experiments as 100% and normalizing all other samples in that experiment to this value. Numbers represent means and error bars are the SD of two independent experiments with triplicates each.

Bonferroni post-tests a)	*P* value
B vs. C	***
B vs. D	***
B vs. E	***
C vs. D	**
C vs. E	***
D vs. E	***

Bonferroni post-tests b)	*P* value
B vs. C	ns
B vs. D	ns
B vs. E	*
C vs. D	ns
C vs. E	ns
D vs. E	ns

Table 5-2 Effect of plasmid co-transfection on gene editing efficiency

Statistical significance of differences between samples in Figure 5-2a and b. P values for differences between samples were determined by one-way ANOVA and Bonferroni post-tests. *** $P<0.001$; ** $P<0.01$; * $P<0.05$; ns $P>0.05$ grey text.

Cells were trypsinised and analysed 24 h after transfection. Propidium iodide could not be employed for dead cell discrimination in this experiment because its emission spectrum would overlap with RFP. Cells transfected with mEGFP-Ctr-NT-PTO (A) were used to set a forward scatter (FSC-H) versus side scatter (SSC-H) gate (R1), which excluded most of the dead cells independent of propidium iodide staining (Figure 5-3 left panel). One sample treated with mEGFP-3A-NT-PTO but no pRFP (E), and another sample treated with pRFP but no mEGFP-3A-NT-PTO (F), were used as the EGFP and RFP positive controls for autocompensation with WinList 6.0 (Figure 5-3 right panel). Figure 5-3 shows example data from one of the triplicates for each sample.

As shown in Figure 5-2a and Table 5-2, co-transfection with plasmid and targeting ssODN did significantly increase the gene editing efficiency over transfection with ssODN alone (1.36, 0.89 and 0.73% compared to 0.45%; $P<0.001$ for all pairs). This increase in gene editing efficiencies was dose-dependent as the percentage of green fluorescent cells was also significantly different between the samples with decreasing amounts of plasmid (samples B, C and D; Table 5-2). However, percentages of green fluorescent cells in a second independent experiment were significantly different as determined by two-way ANOVA.

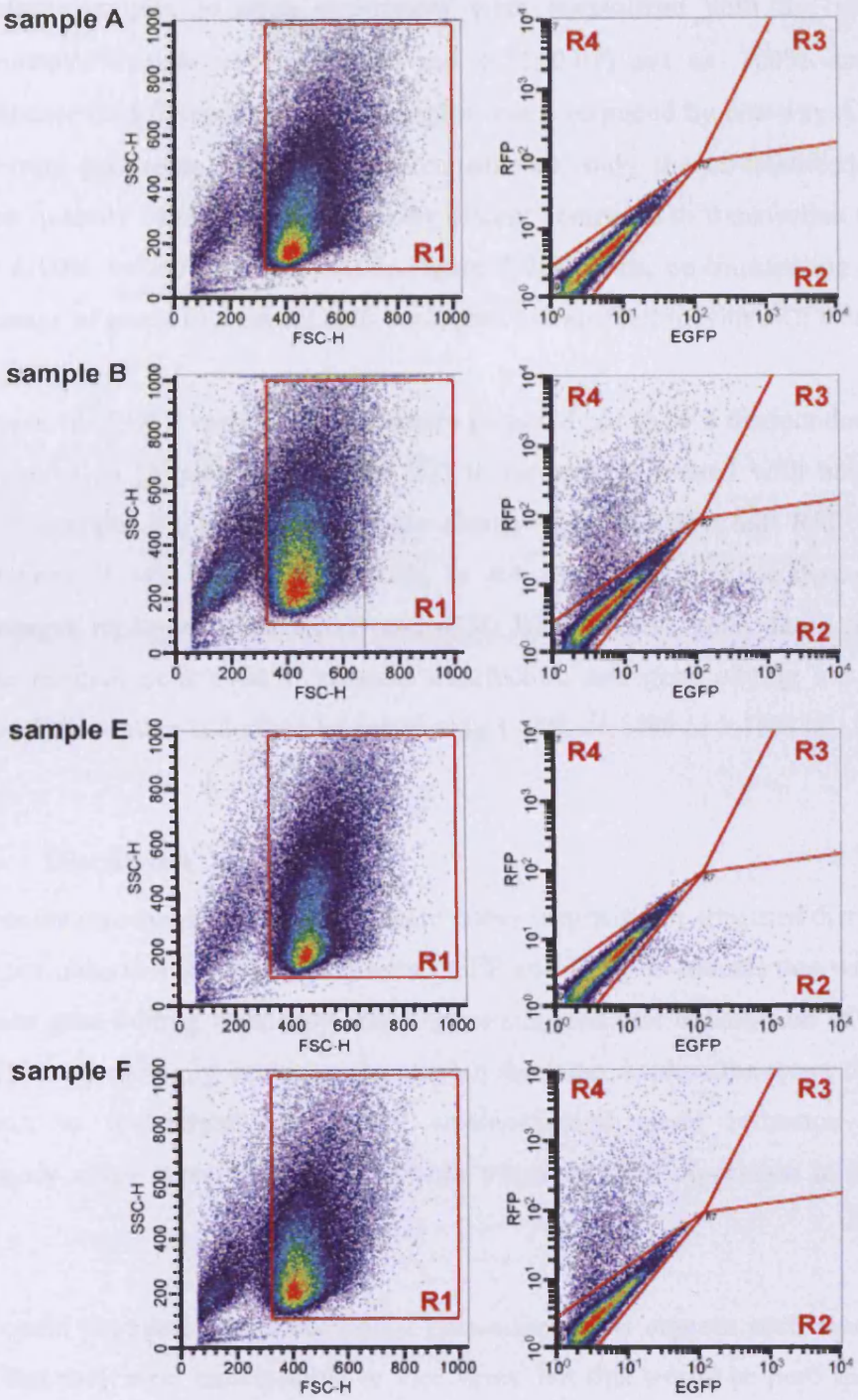


Figure 5-3 Flow cytometry plots of plasmid co-transfection

The FSC versus SSC plot of a cells transfected with mEGFP-Ctr-NT-PTO (sample A) was used to set R1 (left panel). Cells were gated based on R1 and EGFP fluorescence was plotted against RFP (right panel). EGFP+ cells appear in R2, RFP+ in R4 and double positive cells in R3. Cells treated with 3 µg mEGFP-3A-NT-PTO and 2 µg pRFP (B), 3 µg mEGFP-3A-NT-PTO (E) or 2 µg pRFP (F).

Therefore, samples in each experiment were normalized with the mean of that experiment's sample B (1.36±0.09 and 0.71±0.05) set as 100% and statistical significance of differences between samples was determined by one-way ANOVA with Bonferroni post-tests. Under these circumstances, only the co-transfection with the highest quantity of plasmid remained significant compared to transfection with ssODN alone (100% versus 47.5%; $P<0.05$; Figure 5-2b). Thus, co-transfection doubled the percentage of green fluorescent cells compared to transfection with ssODN alone.

However, the EGFP versus RFP cytometry plots did not show a distinct double positive cell population (Figure 5-3; cells in R3) in the sample treated with both pRFP and ssODN (sample B), although there are clearly defined EGFP and RFP positive cell populations (1.44% in R2 and 5.16% in R4 respectively). Considering that these percentages represent 1168 EGFP and 4190 RFP positive cells, there should be 60 double positive cells even if plasmid transfection and gene editing are independent events. This number is derived by calculating 1.44% of 4190 or 5.16% of 1168 cells.

5.2.3 Discussion

The results presented in this chapter were rather surprising. It appeared that even though I did not observe cells expressing both EGFP and RFP, co-transfection with pRFP did increase gene editing efficiency. The former suggests that transfection of plasmid and ssODN were mutually exclusive events, but the latter implies the exact opposite. It is difficult to understand how pRFP co-transfection could influence gene-editing efficiency when there were no detectable traces of RFP expression in GFP positive cells.

One could postulate that successfully gene-edited cells express such low amounts of RFP that they were undetectable or vice versa, but this would be pure speculation. In theory, dimly double positive cells could be hidden by the background fluorescence of the non-edited and non-transfected cells. It is difficult to distinguish such weak positive cells from background fluorescence reliably with flow cytometry, though it could be done with RT-PCR. That, however, would require cell lysis and the goal of this experiment was to select or at least enrich for gene-edited live cells.

A possible explanation for the conundrum of increased gene-editing efficiency but no double positive cells could be that the increasing amounts of PTO-modified residues (the amount of huApoA-I_M-49NT-PTO ssODN increased with the reduction in pRFP) led to an increase in cytotoxicity. If this was the case, the reduction in gene editing efficiency in samples with less or no pRFP would actually be due to an increase in cytotoxicity in those samples, not a real difference in gene editing. More experiments with an unmodified ssODN as the scrambled control would be needed to investigate this possibility.

Noteworthy is also that the limitation on the maximum amount of transfected DNA imposed by the cytotoxicity of Lipofectamine 2000 meant that $1/4^{th}$ less ssODN than usual was transfected. Without further experiments incorporating transfection of 4 or even 5 μg ssODN as gene-editing efficiency controls, no conclusions can be drawn about which transfection condition would give the highest efficiency.

Thus, the main point to take from these results is that plasmid co-transfection cannot be used to enrich for gene-edited cells.

5.3 Selecting G2/M arrested cells by sorting

5.3.1 Introduction

A further attempt to enrich for gene-edited cells was based on data by Olsen *et al.*,[242] which showed that these cells accumulate in the G2/M phase of the cell cycle. Most of the gene-edited cells were terminally arrested, but a small percentage proceeded to form viable colonies. Considering the evidence that the targeting oligonucleotide is physically integrated into the genome of gene-edited cells[259] it is likely that colony-forming edited cells escape from this G2/M arrest rather than not entering arrest at all.

Therefore, I thought that if I transfected synchronized CHO-mEGFP cells with ssODN and released the cells to proceed through the cell cycle, there would be a time point where all normally cycling cells would have gone through the G2/M stage. Any cells remaining in the G2/M phase at that time would then be enriched for gene-edited arrested cells.

Although technically challenging this experiment was feasible in my opinion. Cell-permeable DNA stains that allow flow cytometric cell cycle analysis without the need to fix cells are available. Such DNA stains bind stoichiometrically to DNA, so flow cytometric cell cycle analysis generates distinctive histograms: a high peak consisting of G0/G1 phase cells (one set of paired chromosomes per cell), a smaller peak for the G2/M phase (two sets of paired chromosomes per cell, prior to cell division) and a spread between these peaks with S phase cells (DNA synthesis with variable amount of DNA). Cell cycle synchronization and cell sorting are also widely employed in cell biology. Thus, a combination of these three techniques should be possible.

5.3.2 Results

Unfortunately, all DNA stains are cytotoxic to various degrees, but after researching the available options I settled upon Hoechst 33342. In preliminary experiments, I titrated the Hoechst 33342 to minimize cytotoxicity (data not shown). I found that good histogram data could be achieved at a minimum concentration of 5 µg per ml and $1x10^6$ cells. Even though the stain induced growth inhibition lasting 24-48 h at this concentration, the majority of cells recovered without difficulty (data not shown).

CHO-mEGFP cells were synchronized using 48 h serum starvation and subsequent incubation with 100 µM mimosine as described in section 2.2.1.5. This synchronization method was chosen because Dr. Petter Olsen found that it was optimal for inducing cell cycle arrest in CHO-mEGFP cells (personal communication). Immediately after release from the mimosine induced G1 block, cells were transfected with mEGFP-2A-NT-PTO as described before.

At 13 h after release from the cell cycle block, when non-arrested cells had proceeded through the G2/M phase according to Olsen *et al.*,[242] 5 µg per ml and $1x10^6$ cells Hoechst 33342 was added and $1x10^5$ cells each were sorted for the G0/G1 and G2/M phase (Figure 5-4a). Unsynchronized control cells were also treated with mEGFP-2A-NT-PTO and sorted (Figure 5-4b). Cells were gated based on FSC versus SSC for viability and FSC-area versus FSC-width for doublet discrimination. Cells were sorted directly into full-growth medium containing 10% FBS and plated at $1x10^5$ cells per well of a 6-well plate.

Unfortunately, 24 h after sorting total cell death was observed in the synchronized samples, whereas the majority of non-synchronized cells recovered after 48 h (data not shown). Cells that were synchronized and transfected but not treated with the DNA stain did show moderate cell death.

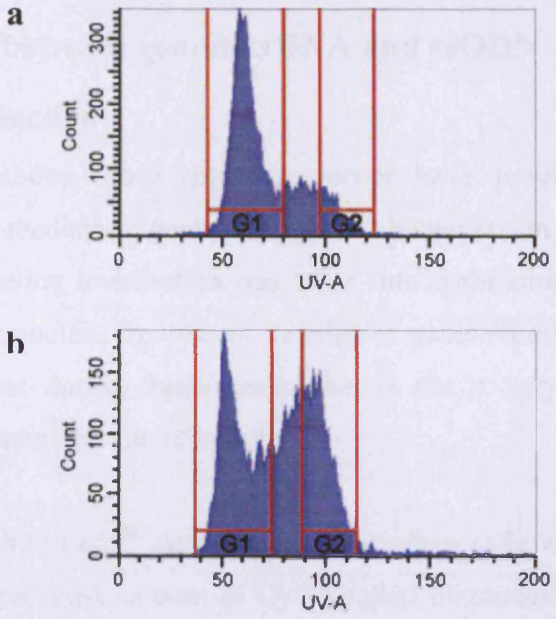


Figure 5-4 Sorting for G2/M arrested cells

Cells synchronized with serum starvation and 100 μM mimosine (**a**) or unsynchronized cells (**b**) transfected with mEGFP-2A-NT-PTO. G1 and G2 denote gates set for cells in the G0/G1 and G2/M cell cycle phases.

5.3.3 Discussion

I concluded that the cytotoxicity of the synchronization treatment, combined with the toxicity of the DNA stain required for cell cycle analysis, was responsible for the total cell death observed in the synchronized and sorted cells. Perhaps a less toxic combination of synchronization method and DNA stain would lead to the recovery of live cells. Encouraging results show that the newly-developed Vybrant DyeCycle stains can be used to sort according to cell cycle though they do cause some growth retardation.[298] However, the cells were not synchronized, so it is difficult to predict if these stains would be less toxic in combination with synchronization.

Depending on the time the stains remain attached to the DNA it might be possible to stain unsynchronized cells with a DNA-specific dye and then sort for cells in the G1 phase. These "pseudo-synchronized" cells could then be transfected with targeting oligonucleotide and sorted again for cells arrested in the G2/M phase at a time when normally growing cells have passed the G2/M checkpoint. Thus, this approach might still be practicable to enrich for edited cells, if the technical problems can be overcome.

5.4 FRET between genomic DNA and ssODN

5.4.1 Introduction

As discussed above, dose response curves have previously been described for oligonucleotide-mediated gene editing,[184] however in these studies only the concentration during transfection was taken into consideration. As the oligonucleotide has to reach the nucleus in order to function in gene editing and undefined amounts of complex are lost during transfection, this is not a very accurate measurement of functional reagent at the site of activity.

Recently, Murphy *et al.*[244] demonstrated that when cells were sorted according to the successfully transfected amount of Cy5-labelled oligonucleotide, the cells with higher ssODN amounts had a higher probability of correction. However, flow cytometry can only determine if and how much Cy5-ssODN is present in each cell, but not where in the cell it is located. Again this is not an accurate measurement of functional reagent at the site of activity.

In this experiment I attempted to employ fluorescence resonance energy transfer (FRET) between Cy5-labelled ssODN and Cy3-labelled genomic DNA to gain a more accurate measurement of the amount of close spatial interaction between the targeting oligonucleotide and the genomic DNA. From the previously published data I predicted that I would be able to enrich for cells most likely to be gene-edited by this method, hopefully to a higher degree than Murphy *et al.*[244] achieved.

The first step was to label genomic DNA with Cy3-dCTP. Labelling genomic DNA with fluorescent nucleotides is a fairly standard method in chromosome studies,[299] but usually involves microinjection (limiting the number of labelled cells), cell permeabilization with glass-beads (not overly successful according to literature) or a "micro-scratch" technique (which also limits cell numbers). However, another publication reported successful use of electroporation to transfect a biotin-labelled dUTP.[300]

5.4.2 Results

Therefore, I electroporated $3x10^5$ CHO-mEGFP cells with 20 µM of Cy3-dCTP and seeded the cells into two wells of a 6-well plate. Another $1.5x10^5$ cells were not electroporated, but 20 µM Cy3-dCTP was added before seeding (section 2.2.1.6). The cells were examined under the fluorescence microscope 2 h after electroporation, or at the start of incubation with Cy3-dCTP. As seen in Figure 5-5a more than 75% of electroporated cells had taken up the labelled nucleotide. There was no accumulation of fluorescence in the cells incubated, but not electroporated, with Cy3-dCTP up to 24 h after the start of the experiment (data not shown).

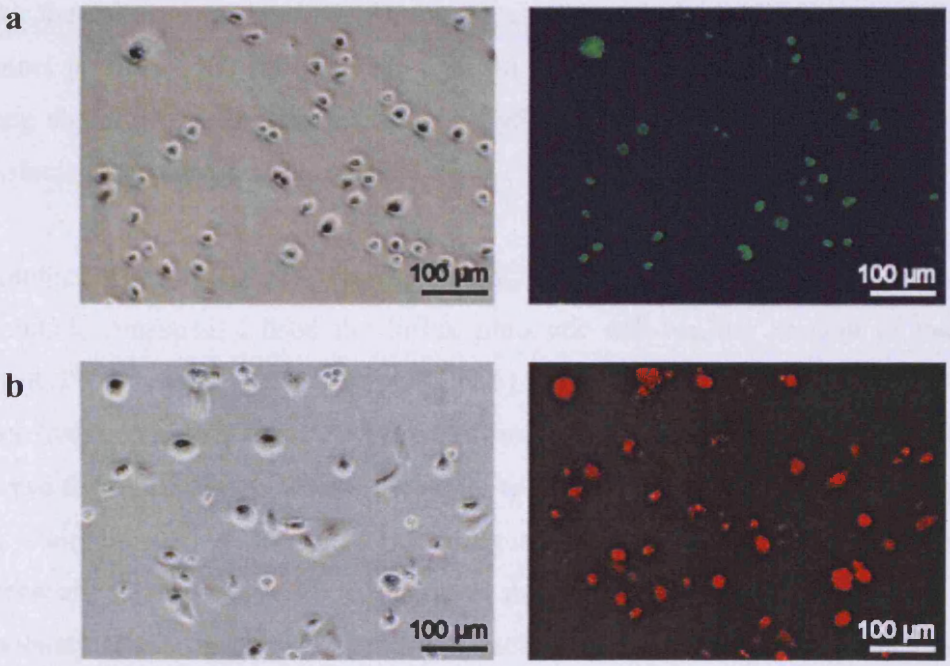


Figure 5-5 Cy3-dCTP electroporated and Cy5-ssODN transfected cells

a) CHO-mEGFP cells electroporated with Cy3-dCTP; left panel phase contrast, right panel Cy3.

b) CHO-mEGFP cells from a) transfected with Cy5-ssODN; left panel phase contrast, right panel Cy5.

The electroporated cells were transfected using Lipofectamine 2000 and Cy5-huAIM-49NT-PTO ssODN 24 h later to allow time for the Cy3-dCTP to be integrated into the genomic DNA. Fluorescence microscope observation 24 h after the lipofection showed that the Cy5-ssODN was easily transfected (Figure 5-5b), however the Cy3 signal was no longer detectable (data not shown).

Therefore, I attempted to transfect Cy3-dCTP (0.2 µg) at the same time as Cy5-ssODN (2 µg) with 4 µl Lipofectamine 2000. Cells transfected with Cy3-dCTP only and Cy5-ssODN only were used as fluorescence controls. The Cy5-ssODN/Cy3-dCTP solution and the Lipofectamine 2000 were diluted in 125 µl OptiMEM each, mixed and added to one well of a 12-well plate containing $1x10^5$ cells in 500 µl antibiotic-free medium. Cells were washed into fresh medium 1.5 h after the start of transfection and observed under the microscope 1 h later. Unfortunately, no signal could be detected in the Cy3 channel in any of the samples, not even in the Cy3-dCTP only control. There was a strong signal in the Cy5 channel in the Cy5-ssODN only and Cy5-ssODN/Cy3-dCTP transfected samples (data not shown).

In another effort to transfer both the Cy3-dCTP and the Cy5-ssODN into cells within a reasonable timespan, I used the Influx pinocytic cell-loading reagent to load 0.2 µg Cy3-dCTP into $1x10^6$ cells (section 2.2.1.6). This reagent has been used to load Alexa Fluor hydrazides and Texas Red dextran into cells. Observation under the microscope showed that immediately after loading, up to 75% of cells had taken up the Cy3-dCTP. The staining pattern 14 h after loading was congruent with that observed after fluorescent ssODN transfection. Confocal microscopy of ssODN transfected cells has previously shown that this pattern is characteristic for staining of the nucleus (section 4.1.1).

Cells were then transfected with 4 µg mEGFP-Cy3-3A-NT-PTO oligonucleotide using Lipofectamine 2000. The incubation time with transfection complex was reduced to 45 min, as cells appeared stressed (Figure 5-6a). One hour after the end of transfection (or t = 16 h after the start of the experiment) cells were examined for both Cy3 and Cy5 signals Figure 5-6b). Another 24 h later (t = 40 h) gene-edited EGFP positive cells could also be seen (Figure 5-6c).

Several attempts to study cells treated in this manner on a flow cytometer at t = 16 h did not result in usable data as the success of the Cy3-dCTP loading seemed to be rather variable and excess Cy5 signal interfered with detection of any possible FRET. In some experiments there were very few surviving cells after loading and lipofection.

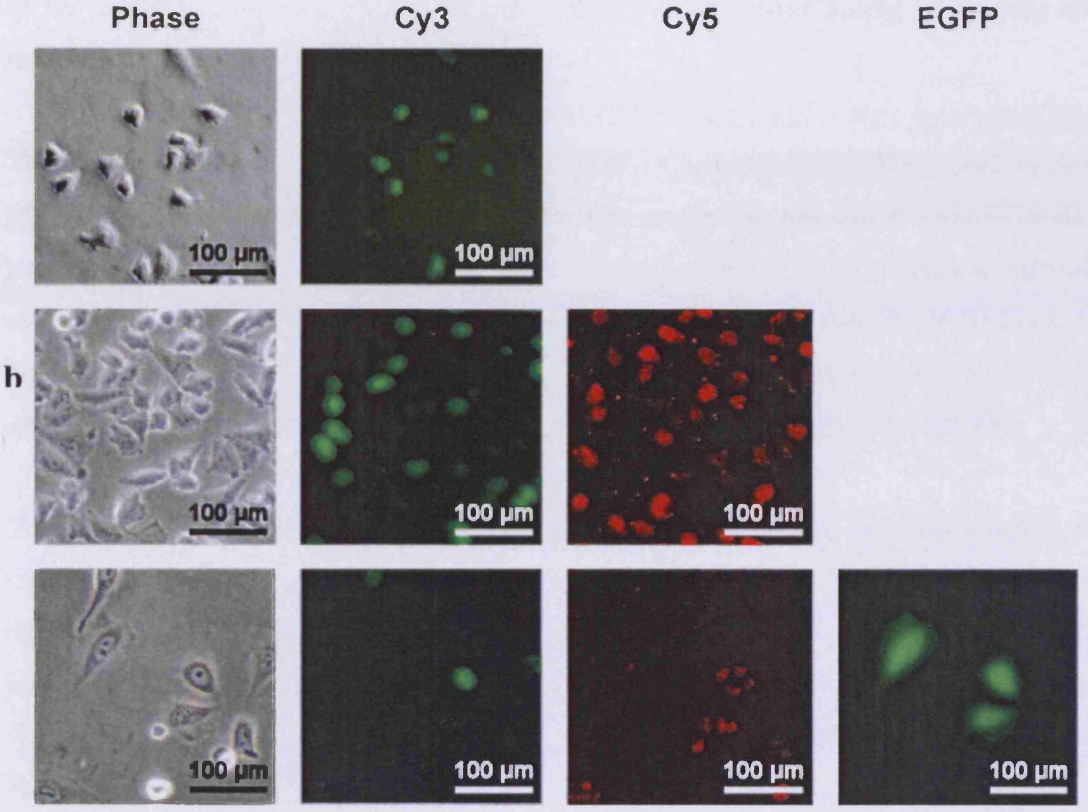


Figure 5-6 Cy3-dCTP loaded and Cy5-ssODN transfected cells

a) CHO-mEGFP cells loaded with Cy3-dCTP 14 h after start of experiment. **b)** Same cells 1 h after transfection with Cy5-ssODN.

c) Same cells 24 h after transfection with Cy5-ssODN when gene-edited EGFP positive cells were visible.

5.4.3 Discussion

Here, I attempted to employ FRET between Cy3-dCTP, which had been integrated into genomic DNA, and Cy5-ssODN to select for cells that have a high amount of spatially close interaction leading to FRET between the partners. Although I managed to transfer both the labelled nucleotide and targeting ssODN into the cells after some optimization, I could not detect any FRET signal on the flow cytometer. Perhaps further optimization of the amount of Cy3-dCTP loaded into cells or of the exact timing of loading and transfection could improve results.

There are several points to take from these results. It appears that either Lipofectamine 2000 is not able to transfect Cy3-dCTP into cells or that the amount of Cy3-dCTP (0.2 μg) was not high enough for detectable fluorescence. The latter seems unlikely though, as 2 μg Cy5-ssODN resulted in strong fluorescence. Considering that the ssODN was 49 bp long and contained one Cy5-dNTP per molecule, the transfected amount of Cy5-ssODN had about 5 times less fluorescent entities than the amount of Cy3-dCTP.

The amount of Cy3-dCTP (0.2 μg) loaded into the cells with the Influx pinocytic reagent also seemed adequate for fluorescence microscopy. However, there is a difference in sensitivity between microscopy and flow cytometry as the former permits longer exposure times, whereas the latter relies on higher laser excitation power, but very short exposure. It might be worthwhile to examine possible FRET under the confocal microscope before attempting to transfer to flow cytometry.

The main disadvantage of using confocal microscopy is that it is impossible to recover a reasonable number of live cells from a mixed population with this method, but at least it would establish if FRET detection of the interaction between targeting ssODN and genomic DNA is possible per se. Even if it is possible to detect such interaction by confocal microscopy it might be impossible to use it for sorting cells with flow cytometry because of the technical difficulties involved.

CHAPTER 6: VALIDATION OF OLIGONUCLEOTIDE-MEDIATED GENE EDITING

6.1 Introduction

In response to the chimeraplasty controversy (section 1.4.5.2 and 1.4.5.3), Albuquerque-Silva *et al.*[227] proposed four experimental criteria for rigorous validation. First, the targeting oligonucleotide must introduce a rare mutant genotype to rule out spontaneous reversion events. Second, this genotype should not exist in cells within the laboratory to prevent cross-contamination artifacts. Third, mutated clones must be studied individually to exclude PCR and sequencing artifacts caused by large molar amounts of targeting oligonucleotides. Fourth, editing efficiency must be confirmed at the protein level in pooled cells at time points that eliminate artefacts of selection and expansion of rare spontaneous mutations.

In this chapter, I set out to validate oligonucleotide-mediated gene editing as a viable technology with experiments fulfilling these requirements. I used the previously described CHO-mEGFP cells (chapter 4) as they allow easy selection of green fluorescent single cells which presumably are gene-edited. However, selection based solely on restored fluorescence cannot exclude spontaneous reversion events.

Consequently, I designed a targeting ssODN that introduces three specific base changes. Two of the nucleotide changes alter the Serine at position 66 in the mEGFP sequence (TCC) to a Tyrosine triplet (TAT) different from the codon in wtEGFP (TAC).The third creates a synonymous Threonine codon in the upstream position (ACA instead of ACC; Figure 6-1). In addition to creating a rare mutant genotype, my ssODN design generates a novel NdeI restriction endonuclease site in the target sequence, which allows successful gene-editing to be distinguished from reversion by Southern blotting. Thus, my experiments fulfilled Albuquerque-Silva's first criterium.

No cells or plasmids containing these sequences existed in the laboratory (or, to my knowledge, worldwide), thus obviating cross-contamination artefacts (Albuquerque-Silva's second criteria). I also interpreted the third requirement conservatively and avoided the potential for PCR and sequencing artefacts caused by large molar amounts of targeting oligonucleotides by using flow cytometry for detection and quantification.

The fourth condition states that editing efficiency must be confirmed at the protein level in pooled cells at time points that eliminate artefacts of selection and expansion of rare spontaneous mutations. Therefore, I transfected the CHO-mEGFP cells with targeting ssODN and isolated single green fluorescent cells 24 h later by fluorescence-activated cell sorting.

I confirmed the presence of the three specific nucleotide changes in the genomic DNA of two independently selected clones by Southern blotting. Figure 6-1 shows a schematic of the expected Southern blot band sizes depending on which enzyme combination was used for digestion of the genomic DNA and which probe was hybridized to the resulting DNA fragments. It is important to remember that the CHO-mEGFP cells used here contain multiple copies of the mEGFP gene and that it is unlikely that more than one or two of the mEGFP copies have been gene-edited. Thus, the mEGFP Southern blot bands from edited clones will not be noticeably reduced in intensity compared to the parental CHO-mEGFP cells, though there should be additional bands demonstrating the presence of at least one edited EGFP (edEGFP) copy.

The predicted band sizes when edited clones are hybridized with Probe A are 576 bp for edEGFP digested with NdeI and MfeI (①), NdeI and DraI (②) or NdeI and XbaI (③). Band sizes for the non-edited mEGFP alleles should be 1271 bp (①), 1223 bp (②) and 1168bp (③) respectively. For edited clones hybridized with Probe B, the expected band sizes are 695 bp (①), 647 bp (②) and 592 bp (③) for edEGFP, whereas the band sizes for mEGFP alleles are the same as with Probe A. Hybridization with both Probe A and Probe B will result in an amalgamated pattern of the bands, so that edEGFP will be denoted by a 576 bp band in all three enzyme combinations as well as 695 bp (①), 647 bp (②) and 592 bp (③) bands depending on which downstream restriction enzyme was used in combination with NdeI. The non-edited mEGFP alleles will show 1271 bp (①), 1223 bp (②) and 1168bp (③) bands respectively.

Sequencing of edited EGFP (edEGFP) gene alleles recovered from these clones also showed the specific nucleotide alterations. Thus, I have validated oligonucleotide-mediated gene editing as a viable technology for introducing up to three nucleotide changes in the genomic DNA of mammalian cells.

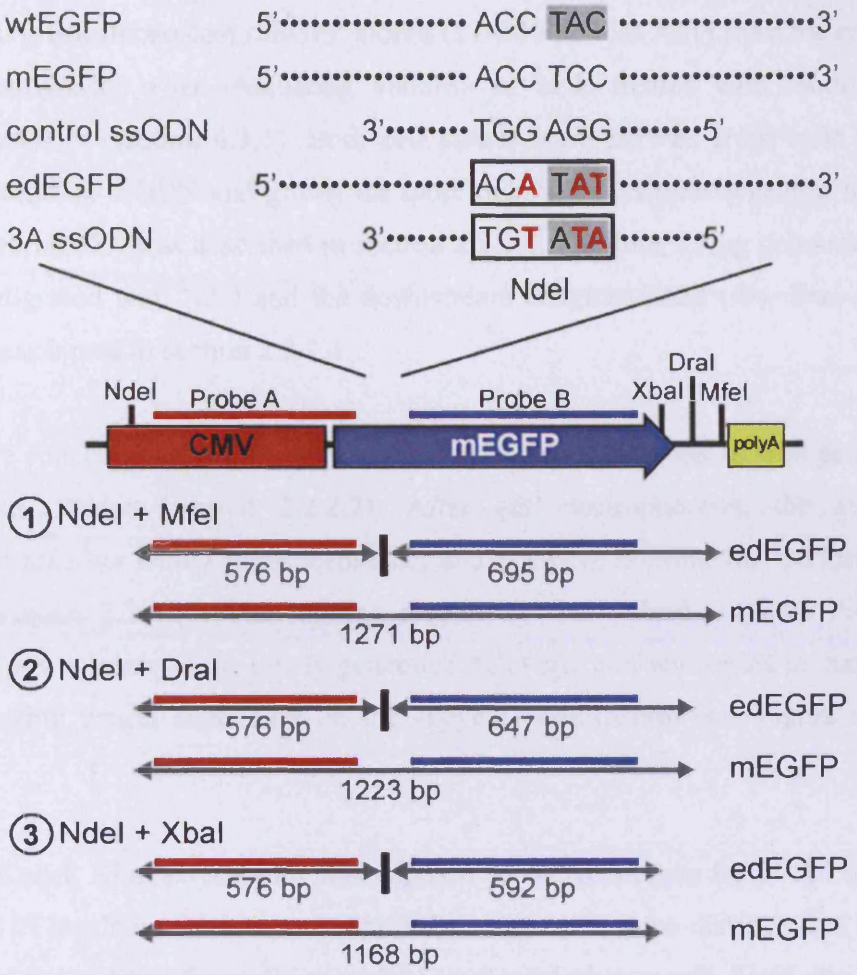


Figure 6-1 Schematic for validation of gene editing by Southern blotting

Mutated EGFP contains a Tyrosine (TAC) to Serine (TCC) substitution at amino acid 66, which destroys the central chromophore region and renders the protein non-fluorescent. The control ssODN is homologous to the mEGFP sequence, whereas the 3A ssODN contains three mismatches (red letters) that change the Serine (TCC) triplet back to an alternative Tyrosine codon (TAT) and add a silent mutation in the upstream Threonine codon (ACC to ACA).

Thus, green fluorescence is re-established (gray shading) and a novel NdeI restriction endonuclease recognition site is created (black boxes). As CHO-mEGFP cells contain multiple transgene copies, both edited EGFP (edEGFP) and mEGFP alleles will be present in gene-edited clones. Digestion with NdeI and the downstream enzymes MfeI (①), DraI (②) and XbaI (③) is predicted to yield diagnostic bands of the indicated sizes.

6.2 Results

I selected two green fluorescent mEGFP clones (11-B10 and 20-A11) from the colonies I obtained previously when evaluating viability of cells treated with modified or unmodified ssODN (section 4.3.1). Both cell clones were derived from cells treated with the unmodified ssODN and grown for more than 4 weeks after selection before I extracted genomic DNA as described in section 2.2.2.1. Per lane, 15 µg genomic DNA was double digested with NdeI and the downstream enzymes MfeI (①), DraI (②) or XbaI (③) as explained in section 2.2.2.4.

Samples were concentrated and loaded onto a 1.5% TAE-agarose gel as well as 2 µg of 100 bp DNA ladder (section 2.2.2.7). After gel electrophoresis, the gel was photographed together with a fluorescent ruler and Southern blotting was performed as detailed in section 2.2.2.10. The blotting membrane was hybridized with Probe B, which binds downstream of the newly generated NdeI site and will result in diagnostic bands of varying length depending on the enzyme combination (see Figure 6-1 for details).

Pre-flashed Kodak Biomax-MS film was exposed to the membrane for 2 - 24 h at -70 °C. The size of bands on the blot was established by marking the distance that marker bands had migrated from the wells according to the gel-photograph. Thus, the size of diagnostic bands could be ascertained.

As can be seen in Figure 6-2, both independent clones (11-B10 and 20-A11) did show the expected diagnostic bands when digested genomic DNA was hybridized with Probe B, while untreated CHO-mEGFP cells digested with the same enzyme combinations did not. The 695 bp (①), 647 bp (②) and 592 bp (③) bands characteristic for the edEGFP allele were clearly present in the genomic DNA from clone 11-B10 (Figure 6-2a). The bands were also present in digested genomic DNA from clone 20-A11 (Figure 6-2b), though this blot lacked samples digested with NdeI and XbaI (③).

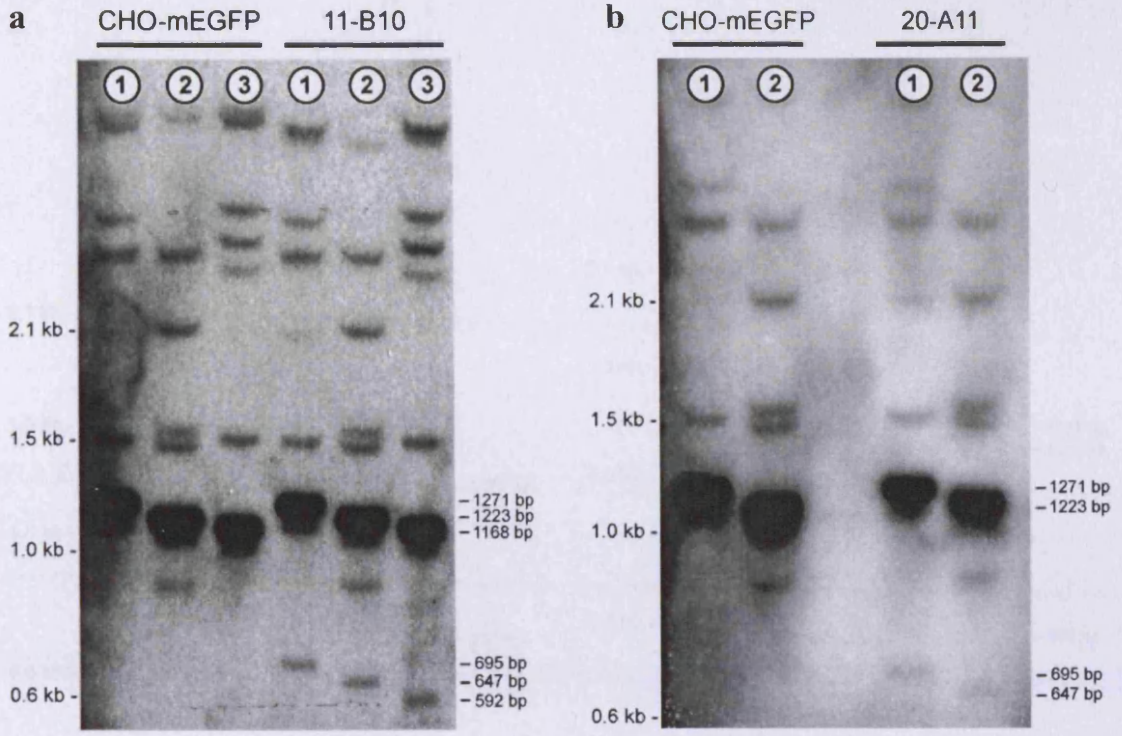


Figure 6-2 Southern blots of gene-edited clones 11-B10 and 20-A11

Southern blots of untreated CHO-mEGFP cells and independently derived clones 11-B10 (**a**) and 20-A11 (**b**) hybridized with Probe B. Labelling on the left marks the location of marker bands and on the right the position of diagnostic bands: 695 (①), 647 bp (②) and 592 bp (③) are characteristic for edEGFP; 1271 (①), 1223 (②) and 1168bp (③) for mEGFP. See also Figure 6-1.

To confirm these results, a second set of Southern blots was performed on both clones. This time, the 11-B10 Southern blot was hybridized with both Probe B and Probe A, which binds upstream of the newly generated NdeI site, but downstream of an established NdeI site. Thus, hybridization with Probe A will result in diagnostic bands of the same length in all enzyme combinations (see Figure 6-1 for details).

As shown in (Figure 6-3a) hybridization of digested 11-B10 genomic DNA with Probe A and Probe B resulted in clearly visible 576 bp diagnostic bands in all digestion samples as well as 695 bp, 647 bp and 592 bp diagnostic bands in digestions ①, ② and ③ respectively.

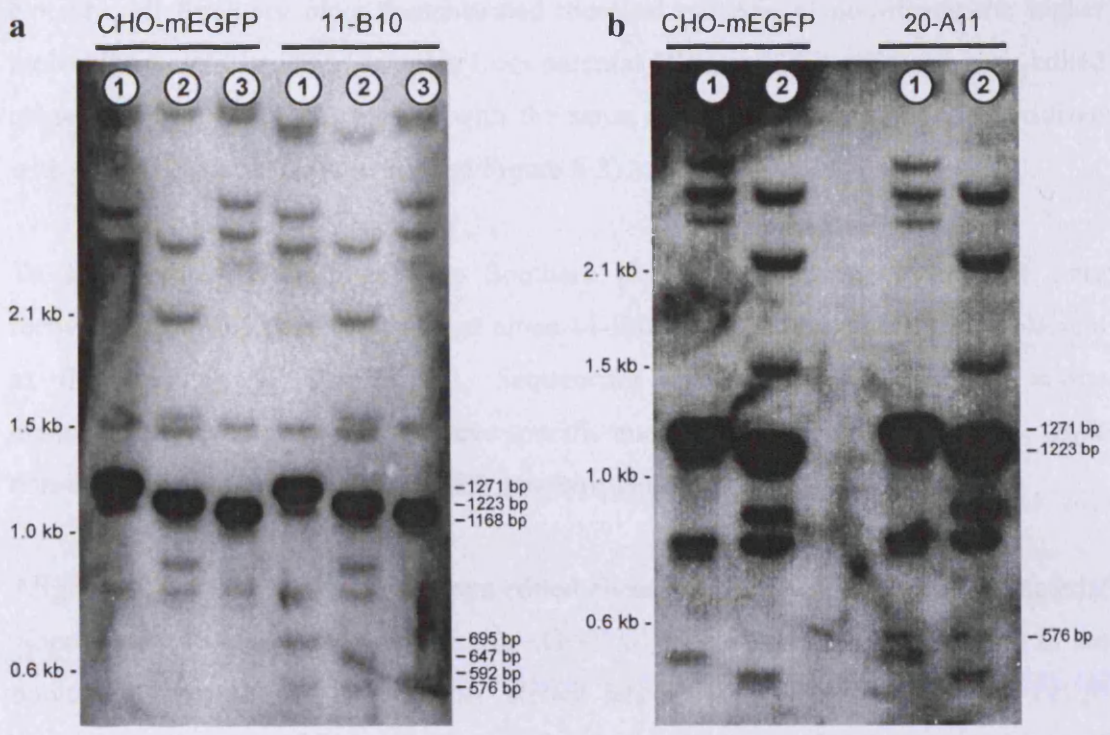


Figure 6-3 Southern blots of gene-edited clones 11-B10 and 20-A11

Southern blots of untreated CHO-mEGFP cells and independently derived clones 11-B10 (**a**) and 20-A11 (**b**) hybridized with Probe A and B (**a**) or Probe A only (**b**). Labelling on the left marks the location of marker bands and on the right the position of diagnostic bands: 576 bp (Probe A), 695 bp (①), 647 bp (②) and 592 bp (③) (Probe B) are characteristic for edEGFP; 1271 (①), 1223 (②) and 1168bp (③) for mEGFP. See also Figure 6-1.

A Southern blot of clone 20-A11 hybridized with Probe A (Figure 6-3b) also displayed the diagnostic 576 bp band, but not in untreated CHO-mEGFP cells digested with the same restriction enzymes. However, two unspecific bands with sizes of around 530 bp and 480 bp (① and ② respectively) were visible in clone 20-A11 as well as the CHO-mEGFP control cells.

As expected, 1271 bp (①), 1223 bp (②) and 1168bp (③) bands characteristic for the presence of mEGFP alleles were evident in both gene-edited and untreated CHO-mEGFP cells (Figure 6-2 and Figure 6-3). There was no discernible reduction in the intensities of these bands in the gene-edited clones, even when X-Ray film exposure times were optimized to reduce saturation of these particular bands (data not shown).

Notably, all Southern blots demonstrated identical patterns of non-diagnostic higher molecular weight bands in samples from parental CHO-mEGFP cells and gene-edited clones, when these were digested with the same enzyme combination and hybridized with the same probes (Figure 6-2 and Figure 6-3).

To corroborate the findings from Southern blotting, edited EGFP allele(s) were recovered from the genomic DNA of clone 11-B10 by PCR and cloned into a plasmid as described in section 2.2.2.11. Sequencing (section 2.2.2.12) of the inserts demonstrated the presence of the three specific nucleotide changes (Figure 6-4a), while non-edited alleles showed the mEGFP genotype (Figure 6-4b).

Alignment of the sequence traces from edited alleles revealed that in one of the bacterial clones there was an additional mutation (GGC Glycine instead of AGC Serine) in the nucleotide after the 27 bp sequence stretch targeted by the editing ssODN, i.e. 5' proximal to the start of the targeting ssODN (data not shown). However, other bacterial clones containing edited alleles did not show this mutation.

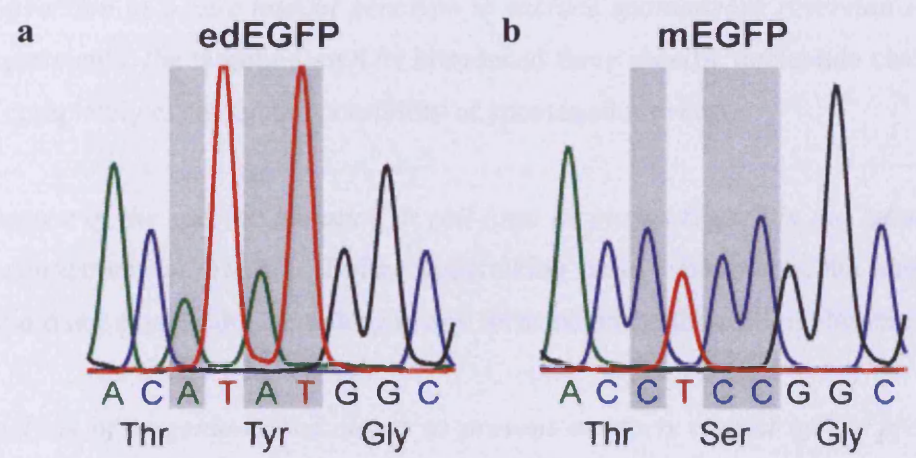


Figure 6-4 Sequencing of edEGFP and mEGFP alleles

Sequence traces of edEGFP (**a**) and mEGFP (**b**) alleles recovered from clone 11-B10. Nucleotide differences between alleles are highlighted with grey shading.

6.3 Discussion

In this chapter I have established without doubt that oligonucleotide-mediated gene editing does introduce specific nucleotide changes into the genomic DNA of mammalian cells. Southern blots of two independently derived clones (11-B10 and 20-A11) using probes from either side of the nucleotide alterations showed bands of the predicted sizes in every case. All three nucleotide changes were required for the generation of the NdeI site, so the Southern blotting results alone demonstrate that these changes were introduced into the genome.

Furthermore, comparison of the higher molecular weight non-diagnostic bands on all Southern blots verified that the gene-edited clones 11-B10 and 20-A11 are isogenic derivatives of the parental CHO-mEGFP cells. Notably, the intensities of the mEGFP diagnostic bands did not change in the gene-edited clones confirming the assumption that only one or two alleles of the multiple mEGFP copies had been gene-edited.

My experiments fully conformed to the four criteria postulated by Albuquerque-Silva *et al.*:[227]

(1) Conversion to a rare mutant genotype to exclude spontaneous reversion events. In my experiments, the targeting ssODN introduced three specific nucleotide changes and hence completely excluded the possibility of spontaneous events.

(2) Absence of the specific mutation in cell lines or plasmids used in the laboratory to avoid contamination artefacts. Before undertaking these experiments, the novel EGFP variant did not exist in the laboratory in any form except as the editing oligonucleotides.

(3) Analysis of isogenic edited clones to prevent artefacts caused by the presence of residual oligonucleotide in genomic DNA that could generate false positives with PCR-based detection. My initial analysis of clones avoided PCR altogether, relying on phenotypic separation by flow cytometry and Southern blotting.

(4) Quantitative analysis at the protein level at time points that exclude artefacts of selection and expansion of rare spontaneous mutations. I used flow cytometric analysis of editing efficiencies at very early time points, before any selection could operate.

In this way, I eliminated all possible scenarios suggesting that my results could be derived from artefacts and therefore incorporation of the three specific nucleotide sequence alterations cannot be explained other than by editing. Radecke *et al.*[241] have previously shown the presence of a novel restriction site in edited cells by Southern blotting. However, restoration of fluorescence in their system could, in principle, have resulted from a spontaneous event because only a single nucleotide change was generated.

The presence of the three specific nucleotide changes in edited clones was also substantiated by sequencing of PCR products recovered as bacterial plasmids. As the clones were derived from single cells and cultured for over 4 weeks after the initial transfection with targeting ssODN, no targeting ssODN could have remained in the cells. Thus, the sequencing results presented here cannot have been due to a PCR artefact caused by the gene editing oligonucleotide.

However, comparison of the sequence traces from different bacterial colonies derived from clone 11-B10 showed that there was an additional mutation in one of them. All other sequence traces did not show extra nucleotide changes. This could suggest that 11-B10 contains two gene-edited EGFP alleles, one with the additional mutation and one without it. More likely though, this change was introduced during PCR amplification of the EGFP gene from genomic DNA, even though the proof-reading polymerase Pfu Turbo was used.

These results highlight the fact that research has not yet answered the question of how accurate oligonucleotide-mediated gene editing is in introducing only the intended and not additional nucleotide changes. If three nucleotide changes can be made in the genomic DNA, target sequence selection will have to take into account closely matching sequences from members of the same gene family or even different families if the target sequence is in a highly conserved domain.

CHAPTER 7: GENE EDITING LEADS TO A STABLE GENOTYPE AND IS POSITIVELY ASSOCIATED WITH EXPRESSION

7.1 The genetic changes are stable but EGFP expression is not

7.1.1 Introduction

During the experiments described in the previous chapter, I observed that only a low percentage of cells in the gene-edited clones 11-B10 and 20-A11 exhibited green fluorescence and that this percentage progressively declined over time in culture. Although these cell lines were established from sorted single cells, the possibility existed that each clone was derived from more than one cell. However, when I isolated green fluorescent and non-fluorescent sub-clones of 11-B10, again by single cell sorting, every single one of the resulting sub-clones including all the clones derived from EGFP negative cells contained green fluorescent cells in varying proportions. The percentages of green fluorescent cells in these sub-clones also declined over time in culture, yet Southern blot analysis of two sub-clones derived from non-fluorescent cells confirmed the continued presence of the three specific sequence changes.

This behaviour suggested epigenetic modulation of expression. To investigate this, I treated a number of 11-B10 sub-clones with the histone deacetylase inhibitor sodium butyrate and the cytosine analogue 5-aza-2'-deoxycytidine which prevents DNA methylation. Both reagents have been shown to raise transgene expression from silenced CMV promoters[301] and, indeed, percentages of green fluorescent cells were significantly increased after such treatment.

This raised the question whether the epigenetic down-regulation was specific to the gene-edited copy or if all copies of the EGFP gene, including the uncorrected ones, were similarly affected. Silencing of integrated transgenes is a well-known phenomenon,[84,302] but if the targeting oligonucleotide is indeed integrated into the genomic sequence, the gene editing event itself could potentially result in down-regulation of the target gene. However, intracellular staining with an antibody recognizing both the mutated and edited EGFP revealed that all cells in which EGFP protein was detectable by antibody staining were also green fluorescent, and that the levels of green fluorescence and EGFP protein were directly proportional to each other. Surprisingly, the majority of cells in the corrected (sub-)clones did not detectably express any form of EGFP.

Therefore, I investigated the stability of mEGFP expression in the parental CHO-mEGFP cells. Notably, I found that even in the earliest passage cells available to me (passage 4 after the initial freeze) only 20% of cells actively expressed mEGFP and that this percentage also rapidly decreased over time in culture, despite the fact that these cells were isogenic. Thus, the epigenetic down-regulation was a result of the transgene reporter gene system and was not due to the gene editing event.

7.1.2 Results

7.1.2.1 Green fluorescent cells are lost from the population of gene-edited cells

The gene-edited cell lines 11-B10 and 20-A11 were regularly observed under the fluorescence microscope for several weeks after selection to examine viability and cell density. I found that less than 30% of cells from 11-B10 (Figure 7-1a) and less than 5% from 20-A11 (Figure 7-1b) demonstrated green fluorescence and that this percentage decreased rapidly over time in culture. Also, the intensity of fluorescence was variable between green cells of the same clone and small colonies of fluorescent cells developed after sub-culturing (Figure 7-1c).

Even though both clones were supposedly derived from single sorted cells and doublet discrimination had been employed during sorting, it was possible that the clones were derived from more than one cell. Thus, I re-cloned 11-B10, again by single cell sorting, and isolated both fluorescent and non-fluorescent cells. Remarkably, every single sub-clone derived from these cells contained green fluorescent cells on expansion, including over 60 sub-clones derived from cells identified by FACS as EGFP-negative (Figure 7-2a; and data not shown).

The proportion of fluorescent cells in sub-clones was variable, but generally higher in those derived from EGFP-positive cells (Figure 7-2b; and data not shown), although this progressively fell to the low levels of negative-sorted cells over time in culture.

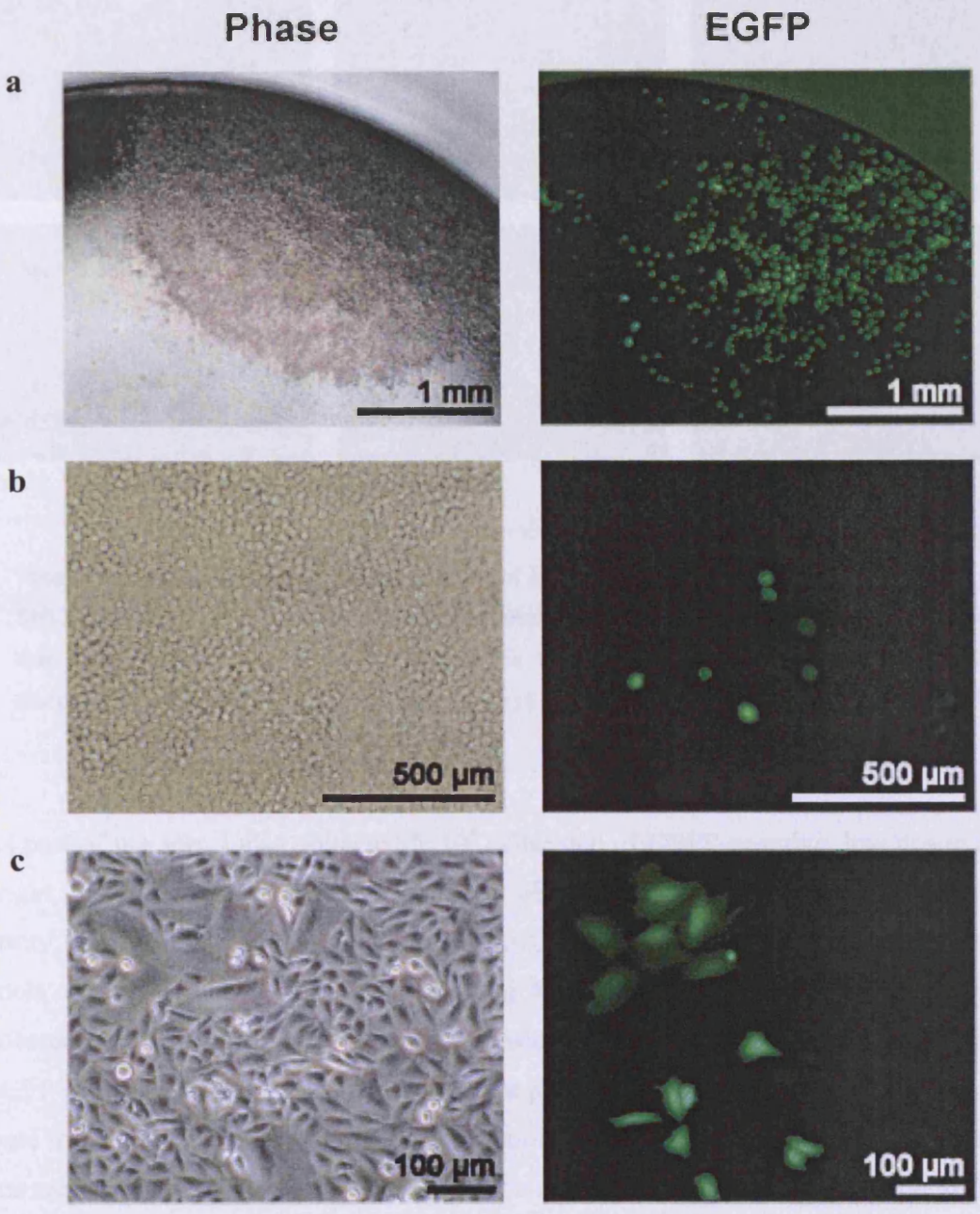


Figure 7-1 Fluorescence microscope pictures of gene-edited clones

The gene-edited clones 11-B10 (**a**) and 20-A11 (**b**) 10 d after selection. Clone 11-B10 30 d after selection (**c**).

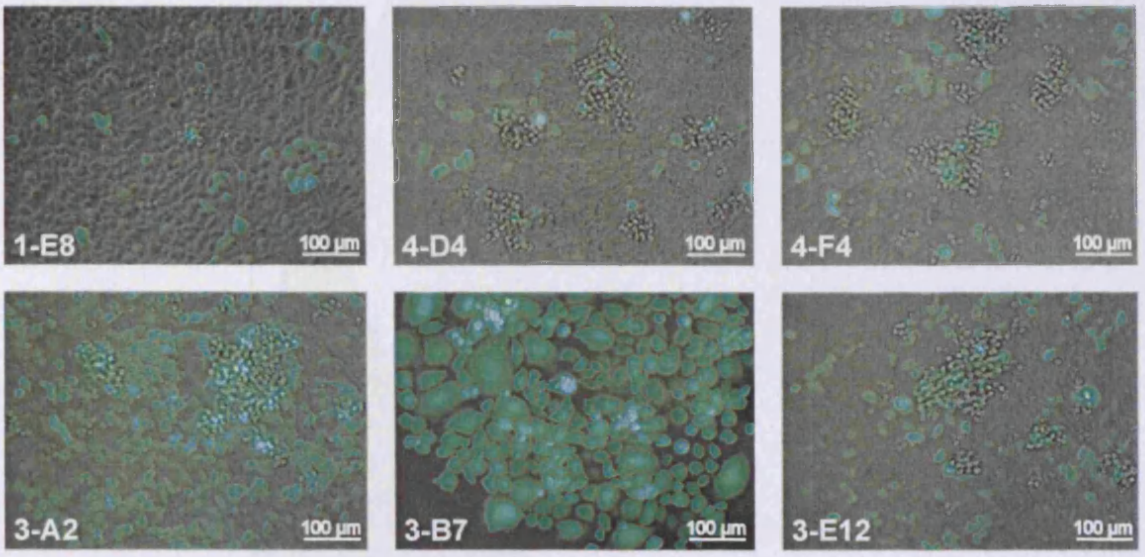


Figure 7-2 Fluorescence microscope pictures of 11-B10 sub-clones

Sub-clones of gene-edited clone 11-B10 derived from single cells that were non-fluorescent (**a**) or that were strongly green fluorescent (**b**) at the time of sorting. Phase contrast and EGFP fluorescence pictures of the same area were captured and then overlaid with LuciaG software.

As part of the sort, I also collected $5x10^4$ cells each of EGFP-negative, low positive or bright positive cells. According to post-sort observation under the microscope, sorting purity was higher than 95% (data not shown). Yet when I compared these sorted cell pools to the parental 11-B10 clone using flow cytometry 12 d later, only small differences in the percentage of green fluorescent cells remained. For this experiment, EGFP-negative cells, low positive and bright positive, as well as parental 11-B10 cells, were trypsinised as usual and treated with propidium iodide for dead cell discrimination. For each sample $1x10^5$ cells were acquired on a flow cytometer, gated based on FSC vs SSC as well as PI and plotted according to their EGFP fluorescence.

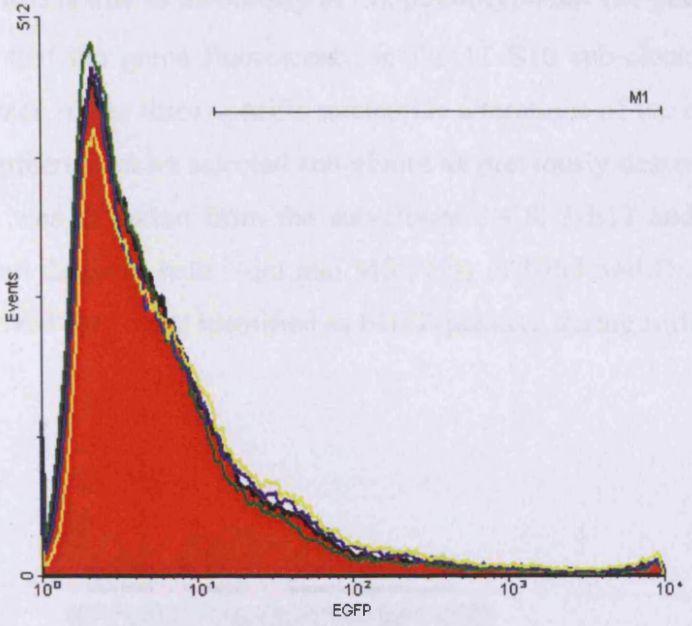


Figure 7-3 Percentage of green fluorescent cells in 11-B10 its subpopulations

Histogram overlay of unsorted 11-B10 (red fill) with negative (green), medium positive (blue) and bright positive (yellow) sorted pooled cells. Cells in region M1 are EGFP-positive.

The resulting histograms for negative, medium positive and bright positive cells were overlaid onto a histogram of unsorted 11-B10 cells using WinMDI 2.8 software. CHO-mEGFP cells were used as a negative control and region M1 was set accordingly to quantify the percentages of EGFP-positive cells. As shown in Figure 7-3, only small differences are apparent between the parental and sorted sub-populations. Thus, 20.19% of unsorted 11-B10 cells showed green fluorescence compared to 16.77%, 22.80% and 26.94% of cells in the negative, low positive and bright positive sorted cell pools, respectively.

7.1.2.2 This loss is due to instability of the phenotype not the genotype

To corroborate that the green fluorescence in the 11-B10 sub-clones was due to the continued presence of the three specific nucleotide alterations of the edEGFP variant, I performed a Southern blot on selected sub-clones as previously described (section 6.2). Genomic DNA was extracted from the sub-clones 3-C8, 3-E12 and untreated CHO-mEGFP cells and digested with NdeI and MfeI (①) or NdeI and DraI (②). Both sub-clones were derived from cells identified as EGFP-positive during sorting.

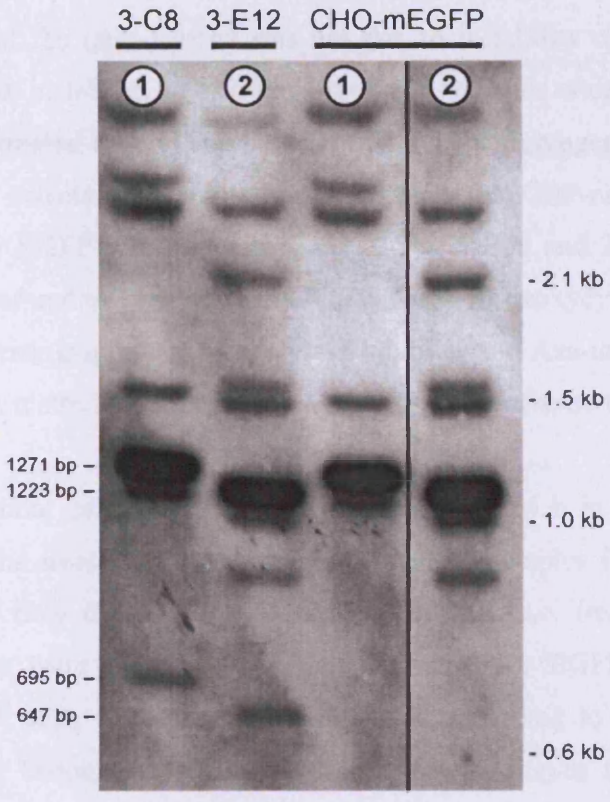


Figure 7-4 Southern blot of 11-B10 sub-clones 3-C8 and 3-E12

Southern blot of untreated CHO-mEGFP cells and sub-clones 3-C8 and 3-E12 hybridized with Probe B. Predicted edEGFP band sizes: 695 (①) and 647 bp (②); mEGFP band sizes: 1271 (①) and 1223 (②). See also Figure Figure 6-1.

After blotting and hybridization with Probe B, clone 3-C8 displayed the 695 bp band characteristic for the presence of at least one edEGFP allele when digested with NdeI and MfeI (Figure 7-4). Digestion of DNA from clone 3-E12 with NdeI and DraI resulted in the expected edEGFP diagnostic band at 647 bp. The clones also showed the 1271 or 1223 bp bands typical of mEGFP alleles after digestion with NdeI and MfeI (①) or NdeI and DraI (②). The CHO-mEGFP control samples showed the mEGFP, but not the edEGFP diagnostic bands (Figure 7-4). The patterns of non-diagnostic bands with higher molecular weight were similar in the sub-clones and CHO-mEGFP cells confirming the isogenic descent of the sub-clones.

In light of these results, it seemed that the loss of green fluorescent cells from the populations of the (sub-)clones was not due to instability of the introduced genomic alterations, but instability of EGFP expression as a result of epigenetic down-regulation. Therefore, I treated several sub-clones of 11-B10 with reagents that enhance transgene expression. I selected four sub-clones derived from EGFP-negative (1-E8, 4-D4, 4-F4 and 4-G4) or EGFP-positive cells (3-A2, 3-B7, 3-E12 and 3-H1) and incubated them with 2 mM of sodium butyrate and 4 μM 5-aza-2'-deoxycytidine (Aza-dC) for 48 h. Sodium butyrate is a histone deacetylase inhibitor and Aza-dC a cytidine analogue that inhibits DNA methylation; both have been shown to increase transgene expression.[301]

After incubation, cells were allowed to recover for 24 h in fresh growth medium. In addition to the treated samples, untreated control samples from each sub-clone were prepared for flow cytometry as described before. Also, treated and untreated CHO-mEGFP cells were included as negative control for EGFP fluorescence. For each sample $1x10^5$ cells were acquired and plotted according to their EGFP fluorescence. The resulting histograms of untreated and treated samples from each sub-clone were overlaid using WinMDI (Figure 7-5) and region M1 was set to quantify the percentage of EGFP-positive cells using the treated CHO-mEGFP negative control.

As shown in Figure 7-5, incubation with sodium butyrate and Aza-dC has a significant effect on the percentage of cells expressing detectable levels of EGFP fluorescence. For example, in clone 1-E8 only 1.41% of untreated cells, but 22.56% of treated cells are green fluorescent, a 15-fold increase.

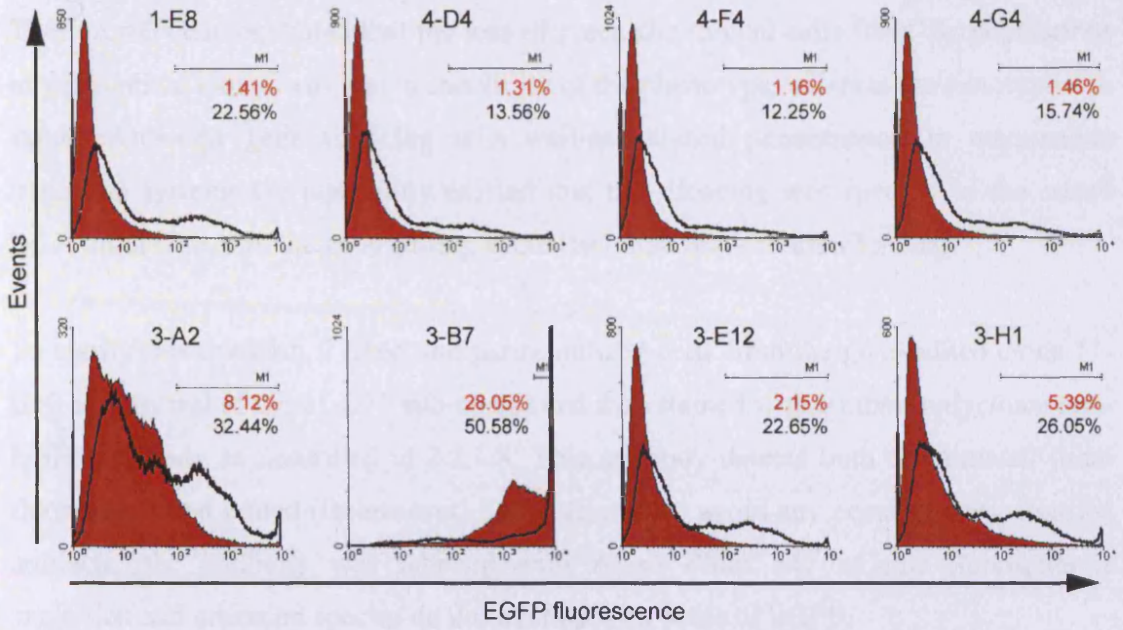


Figure 7-5 11-B10 sub-clones treated with sodium butyrate and Aza-dC

Eight sub-clones of the gene-edited clone 11-B10 derived from single cells that were non-fluorescent (1-E8, 4-D4, 4-F4 and 4-G4) or cells that were strongly green fluorescent (3-A2, 3-B7, 3-E12 and 3-H1) at the time of sorting were either untreated (red fill) or treated with 2 mM sodium butyrate and 4 μM Aza-dC (black line). Numbers represent the percentage of untreated (red) or treated (black) cells in region M1. Data are representative of two independent experiments.

In the other sub-clones that were derived from EGFP-negative cells, percentages of green cells increased about 10-fold upon treatment. Increases in the sub-clones derived from EGFP-positive cells were generally lower (4-fold for 3-A2, 2-fold for 3-B7 and 5-fold for 3-H1), though clone 3-E12 also showed a 10-fold rise.

However, region M1 had to be repositioned for clone 3-B7 as the majority of cells in this sub-clone was very strongly EGFP-positive even before treatment. Indeed, incubation with sodium butyrate and Aza-dC resulted in such high EGFP expression in this sub-clone that the majority of cells exceeded the detection range of the flow cytometer under the acquisition settings used for the other sub-clones.

7.1.2.3 EGFP down-regulation is not specific to the gene-edited copy

These results demonstrated that the loss of green fluorescent cells from the populations of gene-edited clones was due to instability of the phenotype, whereas the genotype was stable. Although gene silencing is a well-established phenomenon in mammalian transgene systems the possibility existed that the silencing was specific to the edited EGFP allele, i.e. that the gene editing event itself had induced the silencing.

To clarify this question, I fixed and permeabilized cells from the gene-edited clone 11-B10 and several of the 11-B10 sub-clones and then stained with a rabbit polyclonal anti-EGFP antibody as described in 2.2.1.8. This antibody detects both the mutated (non-fluorescent) and edited (fluorescent) EGFP forms. To avoid any possible compensation artefacts, the antibody was labelled with Alexa Fluor 647 as this fluorophore's excitation and emission spectra do not overlap with those of EGFP.

CHO-mEGFP cells were used as single stain positive control as they express EGFP, but are not green fluorescent. Fixed and permeabilized cells from the gene-edited, and therefore green fluorescent, sub-clone 3-B7 were stained with an isotype control antibody (rabbit polyclonal anti-human C reactive protein antibody) labelled with Alexa Fluor 647 for use as the second single stain positive control. CHO-K1 cells, the parental line of CHO-mEGFP cells, stained with the anti-EGFP antibody were used as the double negative control because they do not express EGFP nor are they green fluorescent.

For each sample, $2x10^4$ events were acquired on the flow cytometer, gated based on FSC versus SSC and the green fluorescence from expression of the edEGFP allele(s) was plotted against the EGFP (mEGFP and edEGFP) expression detected with the Alexa Fluor 647 labelled pan-EGFP antibody. The double-negative CHO-K1 cells together with the single stain positive CHO-mEGFP and 3-B7 isotype controls were used to determine the respective regions on the green fluorescence versus EGFP expression plots (Figure 7-6).

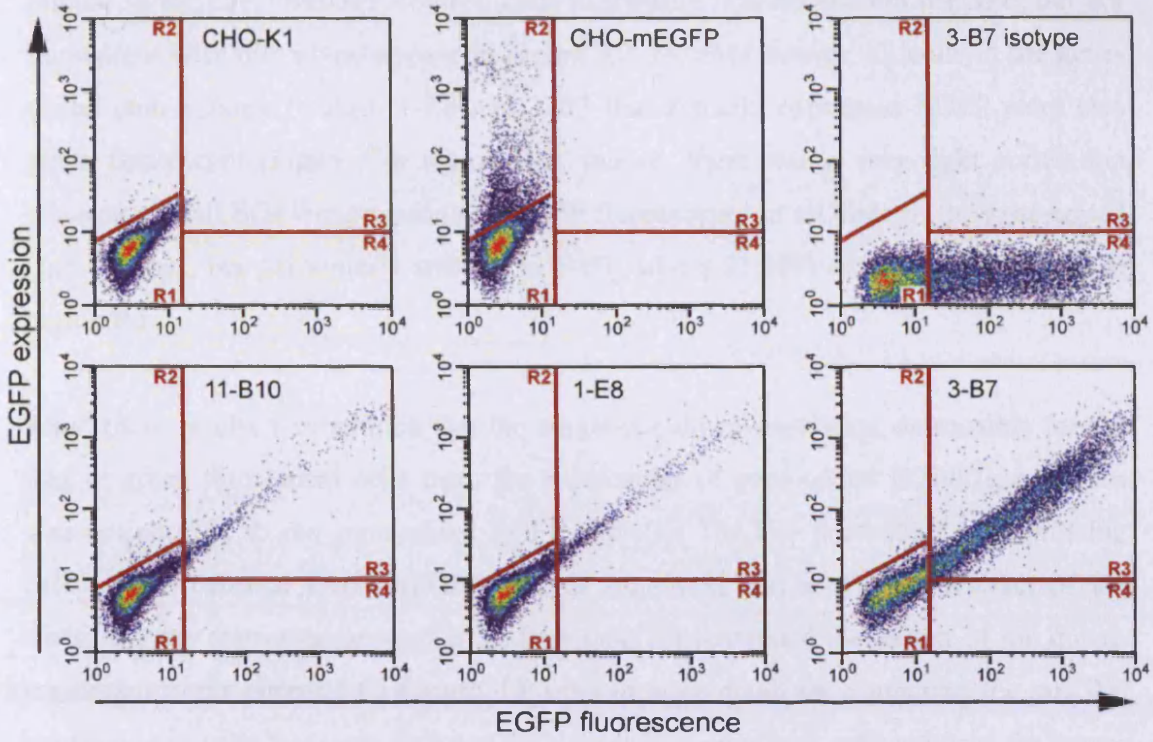


Figure 7-6 Correlation of fluorescence and EGFP expression in edited (sub-)clones

Density plots of green fluorescence from edEGFP expression (x-axis) versus intracellular staining with an Alexa Fluor 647 labelled pan-antiEGFP antibody recognizing both mEGFP and edEGFP (y-axis). CHO-K1 cells were used as a control to determine non-expressing cells; CHO-mEGFP and the 3-B7 sample stained with isotype antibody were used as single stain positive controls. Clone 11-B10 was confirmed to be gene-edited and its subclones 1-E8 and 3-B7 were derived from non-fluorescent and strongly green fluorescent single cells respectively (see Figure 7-2). Data are representative of three independent repeats.

Surprisingly, when comparing the EGFP expression in CHO-mEGFP cells with CHO-K1 cells, I found that >85% of CHO-mEGFP cells did not express mEGFP at all. Similarly, 96.05% and 95.47% of cells in the gene-edited clone 11-B10 and its sub-clone 1-E8 did not express any form of EGFP, be it fluorescent edEGFP or non-fluorescent mEGFP. This percentage remained relatively stable for 11-B10 (94.93±0.87%, n=3), but had increased from 90.48% in a previous independent repeat to 96.54% in a subsequent repeat for clone 1-E8. The percentage of non-expressing cells in the gene-edited sub-clone 3-B7 was 46.62% in this particular experiment, but had risen from 26.55% to 69.15% during the three experiments.

Similar to the CHO-mEGFP control, cells expressing non-fluorescent mEGFP, but not fluorescent edEGFP would appear in region R2. Notably though, all cells in the gene-edited (sub-)clones 11-B10, 1-E8 and 3-B7 that actually expressed EGFP were also green fluorescent (Figure 7-6 region R3). Indeed, there was a very tight correlation between overall EGFP expression and EGFP fluorescence in all three of the gene-edited (sub-) clones, but particularly striking in 3-B7, where 53.24% of cells were located in region R3.

From these results I concluded that the epigenetic down-regulation responsible for the loss of green fluorescent cells from the populations of gene-edited EGFP (sub-)clones was not specific to the gene-edited EGFP allele(s). The low percentage of expressing cells in the parental CHO-mEGFP control suggested that it was an artefact of the multiple-copy transgene array. To confirm this, I investigated the extent of the down-regulation in the parental CHO-mEGFP cells in more detail by comparing the mEGFP expression in cells that were cultured for a short time only with cells cultured for longer times.

CHO-mEGFP cells from passage number 4, 10 and 18 were stained with the Alexa Fluor 647 labelled pan-EGFP antibody, $1x10^5$ events were acquired per sample and analysed as described above. CHO-K1 cells were again used as the non-expressing control and regions were set accordingly (Figure 7-7). The flow cytometric analysis showed a rather rapid decline in the percentage of expressing cells over time: from 20.08% in passage 4 to 10.31% by passage 10 and 6.78% by passage 18. These data are derived from three different defrosted aliquots of the parental CHO-mEGFP cells and were also confirmed in independent experiments. Thus, transgene expression was already unstable in the parental CHO-mEGFP cells.

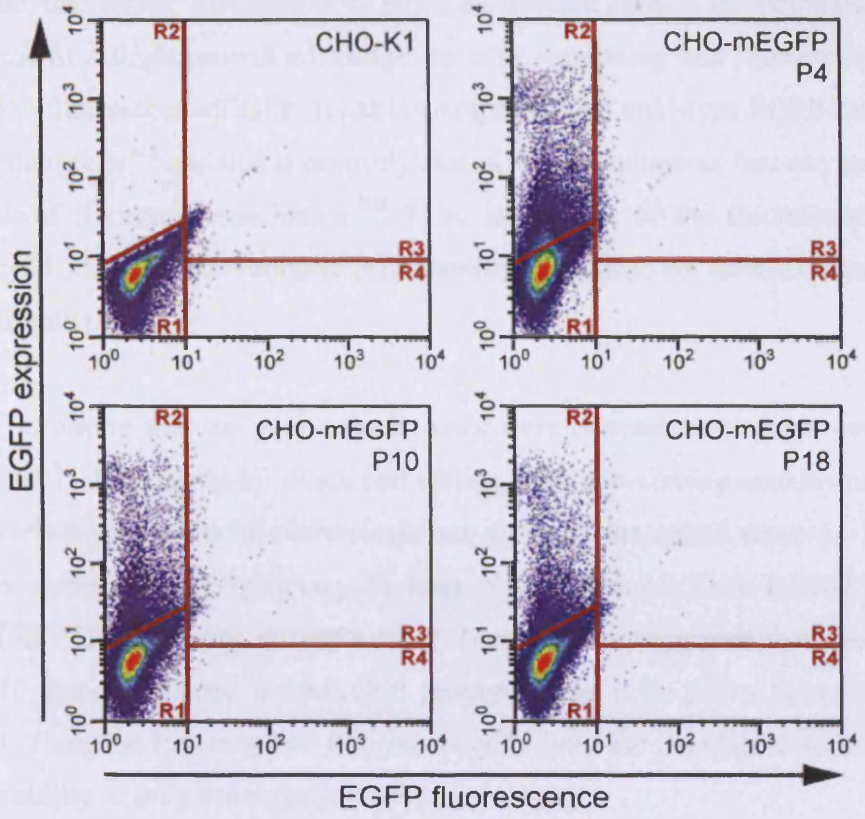


Figure 7-7 mEGFP expression in CHO-mEGFP cells over time in culture

Density plots of green fluorescence (x-axis) versus intracellular staining with an Alexa Fluor 647 labelled pan-antiEGFP antibody (y-axis). CHO-K1 cells were used as a control to determine non-expressing cells; mEGFP expression in CHO-mEGFP cells from passage 4, 10 and 18. Data are representative of two independent repeats.

7.1.3 Discussion

Here, I systematically investigated possible explanations for the loss of green fluorescent cells from the populations of gene-edited (sub-)clones that I observed during isolation of the gene-edited clones described in the previous chapter. At first, the most likely explanation seemed that the edited clones 11-B10 and 20-A11 were both derived from more than one cell.

In this case, decreasing percentages of green fluorescent cells in the populations could be the result of a slight growth advantage for cells expressing non-fluorescent mEGFP compared to fluorescent edEGFP. It has been reported that wild-type EGFP can be toxic in mammalian cells[303] and this is probably due to the production of free oxygen radicals as a result of fluorescent excitation.[304] Thus, expression of the fluorescent edEGFP variant could also lead to reduced cell viability compared to cells expressing non-fluorescent mEGFP.

However, to ensure that my gene-edited clones were derived from single cells, I sub-cloned clone 11-B10, again by single cell sorting. This sub-cloning demonstrated that a certain percentage of cells in every single sub-clone of the edited clone 11-B10 were green fluorescent, even though a large number of the sub-clones were derived from cells identified as EGFP-negative during sorting. These results suggested that every cell in the 11-B10 clone contained the edEGFP genotype even if no green fluorescence was expressed. Thus, the loss of green fluorescent cells from the population appeared to be due to instability of the phenotype and not the genotype.

Southern blot analysis of two 11-B10 sub-clones also showed that the specific sequence changes were still present in theses clones, more than 10 weeks of culture and two rounds of cloning after the original selection. However, because the CHO-mEGFP cells and their sub-clones contain multiple mEGFP copies and only one or two of these may be edited, Southern blotting alone could not completely exclude the possibility that 11-B10 and its sub-clones are mixed populations of edited and non-edited cells.

On the other hand, recovery of fluorescent EGFP expression in a significant proportion of cells in the sub-clones after treatment with sodium butyrate and Aza-dC established that the edited genotype was still present even in non-fluorescent cells. Therefore, 11-B10 and its sub-clones were not derived from more than one cell. In fact, this experiment demonstrated that epigenetic down-regulation was responsible for the loss of green fluorescent cells from the population of gene-edited clones. These results show that the genotype is stable, but the phenotype is not.

Although epigenetic down-regulation of transgene expression is a well-known phenomenon in transgenic mice[84] and mammalian cells,[302,305] I had to exclude the possibility that the gene editing event itself was responsible for the loss of fluorescent cells from the populations of gene-edited (sub-)clones. There is convincing evidence that the targeting oligonucleotide is physically integrated into the genome of edited cells[259] and, analogous to what happens after transgene integration the integration event itself could potentially cause specific down-regulation of the edited allele or general down-regulation of all EGFP alleles.

However, labelling of the gene-edited (sub-)clones with an EGFP antibody resulted in a clearly defined population expressing both non-fluorescent mEGFP and fluorescent edEGFP, while there were no cells expressing non-fluorescent EGFP only. Thus, the down-regulation was not specific to the gene-edited allele. Also, the direct proportionality of EGFP expression and green fluorescence seen in the gene-edited clone 11-B10 and its sub-clones suggests that the epigenetic modulation of expression of the different copies of the EGFP gene was coordinated. Further experiments demonstrated that the epigenetic down-regulation was already present in the parental CHO-mEGFP cells. Again, this established that gene editing itself was not the cause of the down-regulation.

In summary, oligonucleotide-mediated gene editing leads to a stable genotype and does not induce epigenetic down-regulation of the targeted gene. The unstable phenotype is an artefact of the EGFP transgene system and, at least in this system, a large majority of cells do not express the transgene. These results not only have implications for my research, but also for gene editing in general.

In my system, the lack of target gene expression in 80% of the cell population means that any editing in those cells would have been undetectable using the phenotypic assay. Thus, if any gene editing has occurred in that population, I have underestimated the efficiency. I also found that expression was unstable in viable edited cells.

These data suggest that calculating viability of gene-edited cells in pools of treated cells may result in underestimates unless corrected for instability of expression. To my knowledge, all published data in regards to viability of gene-edited clones resulted from exactly such experiments. In particular, viability data published by Olsen *et al.*[242,251] are based on the exact same system. Indeed, data from measuring viability of single gene-edited cells presented in a previous chapter (section 4.3.1) suggested a much higher viability (15.77% for cells treated with PTO-modified ssODN and 25.05% with unmodified ssODN) than that reported by Olsen *et al.*[251] (2% and 6% respectively).

Also, the observation by Olsen *et al.*[242] that the intensity of green fluorescence varied between the two daughter cells of gene-edited cells can no longer be taken as supporting evidence that gene editing occurs predominantly in the S or G2/M phase. However, this does not invalidate other data presented in the same paper demonstrating that transfection with targeting ssODN immediately after release from G1 arrest results in significantly higher gene editing efficiencies.

7.2 Gene editing detection is dependent on and positively associated with target gene expression

7.2.1 Results

Considering that successful gene editing as signified by green fluorescence would be undetectable in the 80% of cells not expressing any EGFP, I wondered what the gene editing efficiency would be if only EGFP-expressing cells were evaluated. Therefore, I transfected CHO-mEGFP cells with 4 μg of mEGFP-3A-NT-PTO ssODN as previously described and then looked at the percentage of green fluorescent cells compared to EGFP expression 24 h later. EGFP expression levels were determined by intracellular staining with a pan-EGFP antibody recognizing both the non-fluorescent mEGFP and fluorescent edEGFP variants (section 2.2.1.8).

For each sample, $3x10^5$ events were acquired on the flow cytometer, gated based on FSC versus SSC and EGFP fluorescence was plotted against total EGFP expression (Figure 7-8a). CHO-mEGFP cells transfected with 4 μg of mEGFP-Ctr-NT-PTO were used as negative control to determine positive green fluorescence (i.e. where to set region R1). In this way, I determined that 2.53% of cells in the population were green fluorescent and, therefore, gene-edited.

I then gated out cells that did not express any EGFP (Figure 7-8b) using non-expressing CHO-K1 cells transfected with the targeting ssODN as double negative control, as I would be unable to determine if those cells were gene-edited or not. I found that in the EGFP expressing subpopulation (Figure 7-8b, cells in region R2; 23.78% of the total population) derived in this manner, 10.39% of cells were green fluorescent. As expected, higher expressing cells were also stronger green fluorescent, but notably an elevated percentage of high expressing cells appeared to be gene-edited compared to low expressing cells.

To more accurately quantify this observation, I divided the EGFP expressing subpopulation into ten regions, each containing 10% of the cells (Figure 7-8c, regions R3-R12). The regions were set according to the level of total EGFP expression in those cells, so that R3 was the decile with the lowest expressing cells and R12 the decile with the highest expressing cells.

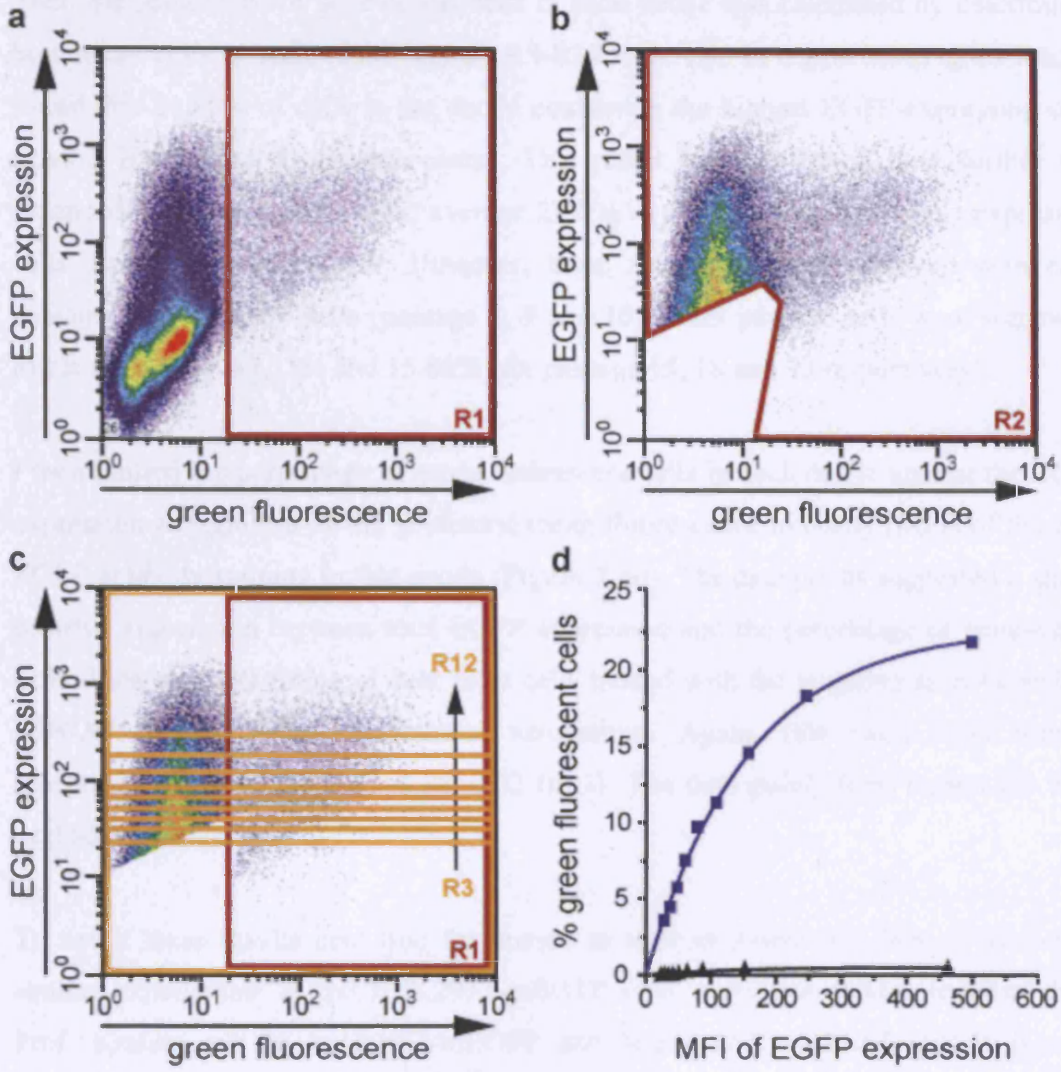


Figure 7-8 Percentage of edited cells and EGFP expression in CHO-mEGFP cells

a) Density plot of CHO-mEGFP cells treated with mEGFP-3A-NT-PTO 24 h after transfection. Green fluorescent gene-edited cells are located in region R1.

b) Non-expressing cells as determined by parental CHO-K1 control cells treated with targeting ssODN are gated out of the density plot, so only EGFP expressing cells are in region R2.

c) Cells were then divided into deciles of EGFP expressing cells (regions R3-R12) and the percentage of green fluorescent cells in each decile was determined (cells that were also in R1).

d) These percentages were plotted against the geometric mean fluorescent intensity (MFI) of the pan-EGFP antibody staining in each decile. CHO-mEGFP cells treated with mEGFP-3A-NT-PTO (■) or mEGFP-Ctr-NT-PTO ssODN (▲); one-phase exponential association curves of the data in the respective colours. Data are representative of three independent experiments.

Then the percentage of gene-edited cells in each decile was calculated by determining how many cells in each of the regions R3-R12 were also in region R2 (Figure 7-8c). I found that 21.80% of cells in the decile containing the highest EGFP-expressing cells (region R12) were green fluorescent. This result was confirmed in a further two independent experiments and on average 23.09±1.16% (n=3) of the highest expressing cells were green fluorescent. However, these results were all achieved with early passage CHO-mEGFP cells (passage 4, 9 and 10). Later passage cells were somewhat lower at 18.75%, 17.13% and 15.86% (for passage 15, 18 and 23 respectively).

I then plotted the percentage of green fluorescent cells in each decile against the EGFP expression as signified by the geometric mean fluorescence intensity (MFI) of the pan-EGFP antibody staining in that decile (Figure 7-8d). The data points suggested a strong positive association between total EGFP expression and the percentage of gene-edited cells. Indeed, curve fitting of data from cells treated with the targeting ssODN yielded R^2=0.99 for one-phase exponential association. Again, this value held true in independent repeats, with R^2=0.98±0.02 (n=3). The data points from region R3 were excluded in all cases.

To see if these results held true for human as well as rodent cell lines, I performed similar experiments in the HEK293T-mEGFP cells, which were also received from Prof. Krauss, and in a HepG2-mEGFP cell line I had produced myself (section 2.2.1.10). Cells were transfected with 4 µg mEGFP-3A-NT-PTO or mEGFP-Ctr-NT-PTO as described in section 2.2.1.4. After intracellular staining with the pan-EGFP antibody recognizing the mEGFP and edEGFP variants (section 2.2.1.8), $8x10^5$ events were acquired on the flow cytometer.

Events were gated based on FSC versus SSC and EGFP fluorescence was plotted against total EGFP expression (Figure 7-9a, b). HEK293T-mEGFP and HepG2-mEGFP cells transfected with the control ssODN were used as negative control to determine positive green fluorescence (i.e. where to set region R1). Staining with the pan-EGFP antibody showed that, in comparison to the CHO-mEGFP cells, a higher percentage of cells in these cell lines actually expressed mEGFP.

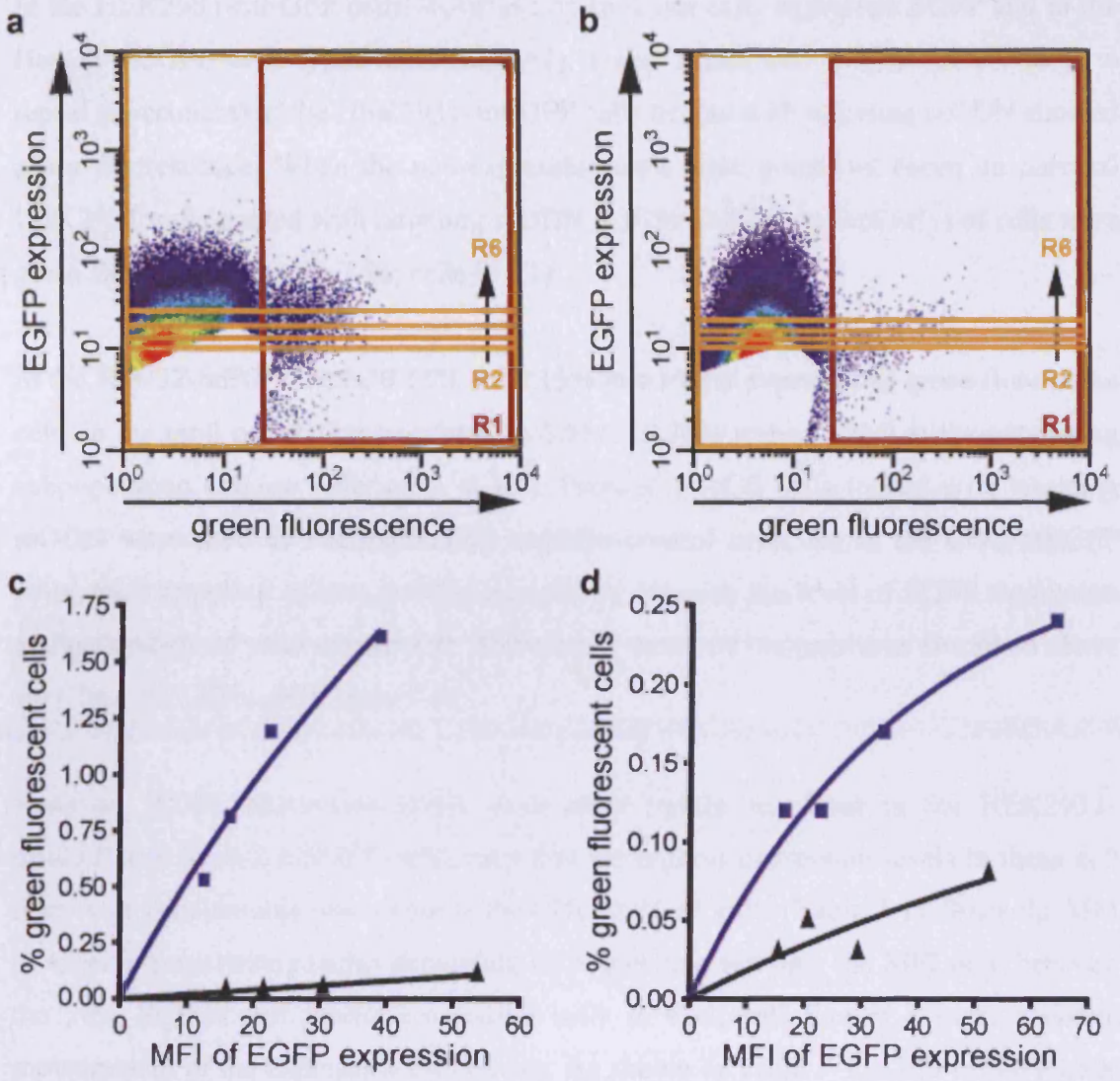


Figure 7-9 Percentage of edited cells and EGFP expression in HEK293T- and HepG2-mEGFP cells

Density plot of HEK293T-mEGFP (**a**) and HepG2-mEGFP cells (**b**) 24 h after treatment with targeting ssODN. Green fluorescent gene-edited cells are located in region R1. Non-expressing cells as determined by the parental HEK293T or HepG2 control cells treated with targeting ssODN were gated out of the respective density plots. Cells were divided into quintiles, each containing 20% of EGFP expressing cells (regions R2-R6) and the percentage of green fluorescent cells was determined (cells that were also in R1). These percentages were plotted against the MFI of the pan-EGFP antibody staining in each quintile of HEK293T-mEGFP (**c**) or HepG2-mEGFP (**d**) cells treated with targeting ssODN (■) or control ssODN (▲); one-phase exponential association curves of the data in the respective colours. Data are representative of two independent experiments each.

In the HEK293T-mEGFP cells 46.49%±1.85 (n=2) of cells expressed EGFP and in the HepG2-mEGFP cells 54.85%±13.62 (n=2). I also found that 0.49% (or 0.19% in a repeat experiment) of the HEK293T-mEGFP cells treated with targeting ssODN showed green fluorescence. When the non-expressing cells were gated out based on parental HEK293T cells treated with targeting ssODN, 1.02% (0.42% respectively) of cells were green fluorescent (Figure 7-9a; cells in R1).

In the HepG2-mEGFP cells, 0.11% (or 0.15% in a repeat experiment) green fluorescent cells in the total population translated to 0.15% (0.36% respectively) in the expressing subpopulation (Figure 7-9b; cells in R1). Parental HepG2 cells treated with targeting ssODN were used as non-expressing negative control cells. As in the CHO-mEGFP cells, there appeared to be a positive association between the level of EGFP expression and percentage of gene-edited cells. Therefore, I analyzed the results as described above for CHO-mEGFP cells (Figure 7-9).

However, EGFP expression levels were more tightly regulated in the HEK293T-mEGFP and HepG2-mEGFP cells, such that the highest expression levels in these cell lines were considerably lower than in the CHO-mEGFP cells (Table 7-1). Since the MFI of specific populations varies depending on acquisition settings, the MFI ratio between the 20% highest and lowest expressing cells in each cell line is a more accurate measurement of the expression differences. As shown in Table 7-1, CHO-mEGFP cells have a 4-fold higher MFI ratio (18.14±1.81; n=3) than either HEK293T-mEGFP (4.48±0.19; n=2) or HepG2-mEGFP cells (4.40±1.03; n=2).

Because of the tight regulation of EGFP expression in HEK293T- and HepG2-mEGFP cells, it was not practicable to divide these cells into deciles for further analysis. Instead they were separated into five regions containing 20% of the cells each (quintiles). The regions were set according to the level of total EGFP expression in the cells, so that region R2 was the quintile with the lowest expressing cells while region R6 contained the highest expressing cells (Figure 7-9a, b).

	MFI of the top and bottom quintile of EGFP-expressing cells in each experiment	MFI ratio
CHO-mEGFP	470.47/26.11; 373.40/23.35; 128.45/6.29	18.14±1.81
HEK293T-mEGFP	68.92/14.76; 38.61/9.00	4.48±0.19
HepG2-mEGFP	64.59/11.90; 33.76/10.03	4.40±1.03

Table 7-1 MFI ratio of the top and bottom quintiles of EGFP-expressing cells
To calculate the MFI ratio, the MFI of the highest 20% and lowest 20% expressing CHO-, HEK293T- and HepG2-mEGFP cells were divided. Then the mean average±standard deviation was determined for the values from the respective independent experiments for each cell line.

When the percentage of green fluorescent cells in each quintile was plotted against the MFI of the pan-EGFP antibody staining of the same quintile, again there was a direct positive association between them. Thus, in HEK293T-mEGFP cells 1.61% (or 0.45% in the repeat experiment) of cells in the highest expressing quintile were green fluorescent. In HepG2-mEGFP cells, the percentage was 0.20% (or 0.24% in the repeat experiment). Curve fitting of data from cells treated with the targeting ssODN yielded R^2=0.97±0.02 (n=2) for HEK293T and R^2=0.84±0.12 (n=2) for HepG2-mEGFP cells one-phase exponential association. The data points from region R2 were excluded in all cases.

I also attempted to further test the relationship between expression and gene editing by treating cells with targeting ssODN after reactivation of expression with Aza-dC and/or sodium butyrate. Unfortunately, the combined treatments were toxic for both CHO-mEGFP and HEK293T-mEGFP cells (data not shown).

7.2.2 Discussion

Here, I set out to determine the gene-editing efficiency of the expressing subpopulation in epigenetically modulated cell populations that for the most part did not express EGFP. As the read-out for gene editing is based on the phenotypic analysis of fluorescent EGFP expression, a large majority of cells that did not express EGFP would theoretically result in a significant underestimation. I found that this was indeed the case when the non-expressing cells were excluded from the analysis in all three cell lines examined. One might argue that by excluding non-expressing cells from the analysis, the percentages of green fluorescent and therefore gene-edited cells are artificially enhanced. However, in my opinion, the adjusted percentages give a much more accurate measurement of gene editing in these cell lines.

Similar to previously reported results for CHO-mEGFP cells (7.1.2.3) I observed variable EGFP expression levels in both HEK293T-mEGFP and HepG2-mEGFP cells. I exploited this fortuitous circumstance to study the relationship between levels of target gene expression and gene editing efficiencies in isogenic cell lines and found a highly significant correlation in CHO-mEGFP cells. I observed similar, if less pronounced correlations, in HEK293T-mEGFP and HepG2-mEGFP cells, although analyses were complicated by low gene editing efficiencies and less variable expression.

The positive association between expression and editing levels is consistent with abundantly expressed genes being better targets for gene editing than those that are less abundantly expressed. However, all three recombinant cell lines have multiple copy integrations of the transgene. Therefore, it is possible that such data could be obtained in the absence of a link between expression and editing efficiencies if different transgene copies were expressed at different levels and regulated independently of each other. On the other hand, multiple transgene copies in recombinant cell lines invariably integrate into a single locus[306,307] and evidence suggests that regulation of all copies is coordinated.[308]

My own results showing the direct proportionality of total EGFP expression and green fluorescence in the gene-edited clone 11-B10 and its sub-clones (Figure 7-6) confirm that there is coordinated epigenetic modulation of expression of all EGFP gene copies. If there was independent regulation of the individual EGFP copies, one would expect a subpopulation of cells in which the fluorescent edEGFP copy is down-regulated, but which are still expressing mEGFP from some of the other non-edited copies. Specifically, such cells should be located in region R2 analogous to where the parental CHO-mEGFP cells are found (Figure 7-6). However, no such cells can be seen on the flow cytometry plots of gene-edited (sub-)clones after intracellular staining with a pan-EGFP antibody (Figure 7-6).

Such subpopulations are clearly present on the density plots of the CHO-, HEK293T- and HepG2-mEGFP cells treated with editing ssODN (cells excluded from region R1; Figure 7-8 and Figure 7-9), but this is due to the fact that in contrast to the gene-edited (sub-) clones only a minority of cells were successfully edited. The majority of treated cells were non-edited and therefore showed mEGFP expression, but no green fluorescence.

Also, there are cells that show considerably lower green fluorescence than cells with the highest green fluorescent intensity for a particular level of EGFP expression. These cells are located between the cells excluded from region R1 and the leading right edge of cells in region R1. This observation could be explained by autonomous expression regulation of individual EGFP copies, that is, the edEGFP copy could be differentially expressed in each cell while the total protein expression remains similar.

However, I have already shown that this is not the case in the gene-edited (sub-)clones. More likely, the lower green fluorescence in these cells is explained by the fact that onset of fluorescent EGFP production was asynchronous depending on the time it took each cell to become gene-edited. Thus, a positive association between target gene expression and editing levels is the only remaining viable explanation. This raises the question of how target gene expression levels can be related to gene editing efficiency. In all probability, this association actually reflects an underlying association between target gene transcription and gene editing efficiency.

Indeed, a variety of published studies support an association between transcription and oligonucleotide-mediated gene editing. Evidence includes strand bias in the activities of oligonucleotides (transcribed vs non-transcribed strand),[240,241,253,264] studies in *S. cerevisiae*,[262] *E. coli*,[266] and of mammalian episomal targets,[267] as well as comparisons of editing frequencies between different cell lines with differential target gene expression.[263,268] Moreover, conventional gene targeting[309] and gene targeting by triplex-forming oligonucleotides[310] are also positively associated with transcription. Although the latter study reports a positive association with transcription, it also shows that gene targeting with triplex-forming oligonucleotides is not dependent on transcription.[310] This suggests that the positive association between transcription and gene targeting may be due to improved access for the oligonucleotide.

In an aside, curve fitting of the data derived from cells treated with the control ssODN suggests the appearance of green fluorescent cells in these samples as well as a direct association with the MFI of EGFP expression. However, I did not employ dead cell discrimination in these experiments because of problems with leakage of dead cell stains into the whole population after fixation. Unfortunately, the background fluorescence of dead HEK-293T and particularly HepG2 cells is increased by transfection with Lipofectamine 2000. Thus, low percentages of false positives were present in these samples. The apparent positive association between these false positives and the MFI of EGFP expression is due to the fact that the measurement error during acquisition on the flow cytometer increases with higher MFI.

Whatever the reason for the observed relationship between target gene expression and editing efficiency, the finding that in the highest expressing cells significantly increased editing rates are possible is very encouraging. Initially, the impressive gene editing efficiencies achieved in the highest expressing CHO-mEGFP cells (23.09±1.16%; n=3) might seem fanciful considering the problems many laboratories have reported with attaining any detectable gene editing.[229,232]

However, using the concept of complementary events in probability theory, the formula $1-(1-x)^n$ can be derived, where x is the chance that one specific copy of the target gene is gene-edited and n the number of target gene copies. Thus, the theoretically possible percentage of gene edited cells if each of the 38 mEGFP gene copies has a 1% chance of being edited is 31.74%. If each copy has a 0.5% chance of being edited, this percentage is reduced to 17.34%. Increased gene editing efficiencies in the highest expressing HEK293T-mEGFP and HepG2-mEGFP were not quite as impressive, but still apparent. This could reflect lower copy numbers and/or the reduced level of expression compared to CHO-mEGFP cells.

In conclusion, the data presented here suggest that there is a correlation between expression of the target gene and gene editing efficiency. Thus, highly expressed genes should be better targets. Direct proof of this hypothesis could easily be obtained by targeting mEGFP expressed from a tet-on/off adjustable expression system. However, it might be wiser to avoid the problems associated with unstable expression from recombinant cell lines altogether and target mammalian genes that can be up- or down-regulated with known stimuli. For example, transcription of ApoA-I can be up-regulated by PPARα agonists such as fibrates (1.2) or down-regulated by decreasing the glucose concentration.

CHAPTER 8: DETECTION OF APOA-I GENE EDITING USING FLOW CYTOMETRY

8.1 Introduction

As shown in the previous chapters (chapter 4-7), the combination of flow cytometry and targeting of mEGFP expressing cells was an extremely useful tool for quantification of gene editing efficiency. It allowed me to optimize gene editing conditions and study possible ways of increasing gene editing efficiencies or selecting for edited cells without the help of a marker gene. I was able to isolate gene-edited clones and demonstrate that oligonucleotide-mediated gene editing leads to a stable genotype and that it is positively associated with expression.

As the conclusion to my studies, I intended to use flow cytometry to examine gene editing of apolipoprotein A-I (ApoA-I) and isolate successfully edited ApoA-I_M expressing cells. Previously I was unable to isolate gene-edited ApoA-I_M expressing HepG2 cell clones by limiting dilution, even though PCR-RFLP of the total population suggested the presence of edited cells up to 10 d after treatment (section 3.2.1). However, a recent report by Favari *et al.*[70] described the use of a monoclonal mouse anti-human antibody that specifically recognizes only ApoA-I_M homodimers. This suggests that the antibody is specific to the conformation of the dimer, so I hoped that I would be able to use it in flow cytometry.

Although ApoA-I and ApoA-I_M are secreted proteins, it should be possible to capture these apolipoproteins by creating an antibody-based affinity matrix on the cell surface and subsequently labelling the captured antigen with a detecting fluorescent antibody (Manz *et al.*[311]). However, this is a technically challenging procedure. Intracellular staining on the other hand would almost certainly allow detection of both variants, but requires fixation and permeabilization, thus precluding isolation of live cells. Nevertheless, I decided to use intracellular staining for a pilot experiment as I was already familiar with the method.

Unfortunately, I could not detect ApoA-I_M in CHO-AIM cells by intracellular staining with the mouse anti-human ApoA-I_M homodimer antibody, even though Western blotting with the same antibody confirmed the presence of ApoA-I_M homodimers in cell lysates and supernatants. Intriguingly though, a polyclonal rabbit anti-human ApoA-I antibody gave a strong signal in CHO-AIM cells, but not CHO-AI cells, whilst Western blotting with the antibody demonstrated expression in both cell lines.

8.2 Results

In this pilot experiment I tested whether the mouse anti-human ApoA-I_M homodimer antibody could be used for flow cytometry. CHO-K1, CHO-AI and CHO-AIM cells were fixed, permeabilized and labelled as described in section 2.2.1.8. For each sample 1 µl polyclonal rabbit anti-human ApoA-I or 250 ng monoclonal mouse anti-human ApoA-I_M antibody were labelled with the rabbit Alexa Fluor 488 nm IgG or the mouse IgG2a Alexa Fluor 647 Zenon labelling kit. After staining, $5x10^4$ events were acquired and gated based on FSC versus SSC. To maximize the resulting information, the fluorescent intensities were displayed on density plots even though each sample was only stained with one antibody to avoid potential competition.

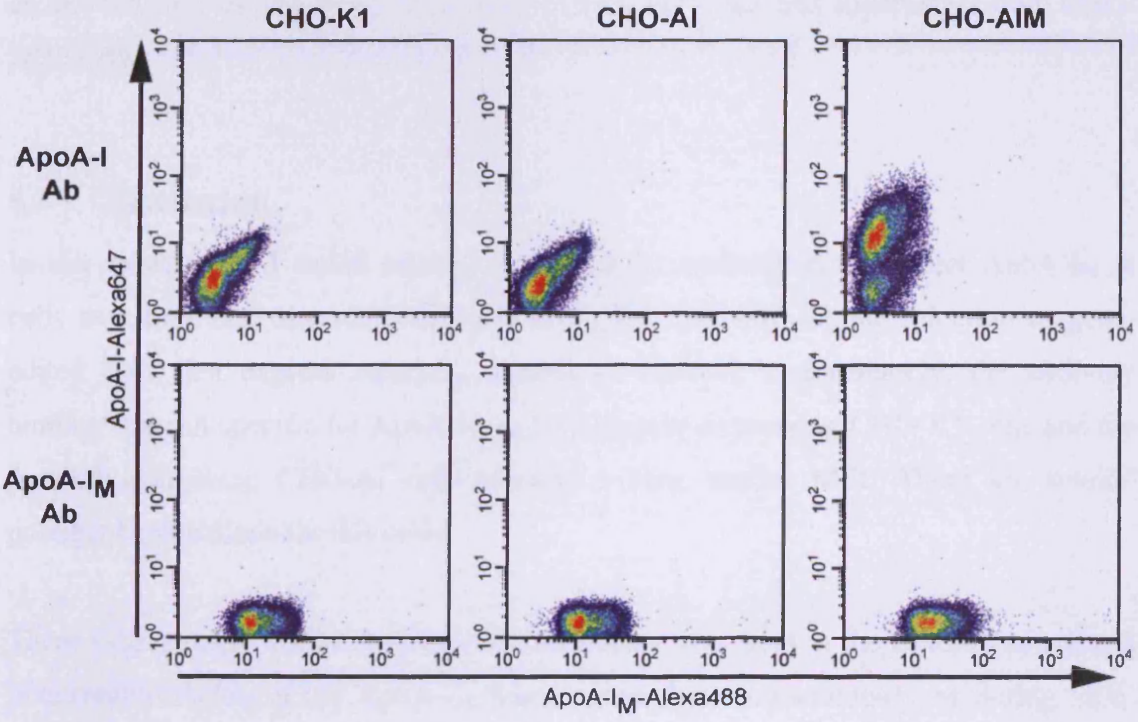


Figure 8-1 ApoA-I and ApoA-I_M detection by flow cytometry

CHO-K1, CHO-AI and CHO-AIM cells were stained with either the ApoA-I-Alexa647 or the ApoA-I_M-Alexa488 antibody. Data are representative of two independent experiments.

As shown in the lower panel of Figure 8-1, there was only a very slight increase in staining with the ApoA-I_M homodimer antibody in CHO-AIM cells (MFI 18.9) compared to CHO-K1 and CHO-AI control cells (MFI 15.0 and 14.1 respectively). Notably, the ApoA-I antibody did not give any signal over the background defined by CHO-K1 cells in the CHO-AI cells, but a distinct population of expressing cells was apparent in the CHO-AIM cells. The former observation suggested that there was no ApoA-I expression in the CHO-AI cells, while the latter was not surprising, since the ApoA-I antibody was selected against a peptide from the C-terminus and thus was not expected to distinguish between wild-type and mutant protein.

However, a Western blot performed by my colleague Dr. Ioannis Papaioannou on supernatant from the CHO-AI cells with the ApoA-I antibody demonstrated that these cells did express ApoA-I (personal communication). Also, the ApoA-I_M antibody clearly detected the ApoA-I_M homodimer in both the lysate and supernatant from CHO-AIM cells.

8.3 Discussion

In this experiment, I tested whether the ApoA-I_M antibody could detect ApoA-I_M in cells as a first step towards potentially using this antibody for the selection of gene-edited cells that express ApoA-I_M instead of ApoA-I. Unfortunately, the antibody binding was not specific for ApoA-I_M as both the non-expressing CHO-K1 cells and the ApoA-I expressing CHO-AI cells showed a very similar MFI. There are several possible explanations for this result.

There simply might not be any ApoA-I_M homodimers present in the CHO-AIM cells. It is currently unclear if the ApoA-I_M homodimers form intracellularly or during HDL maturation after secretion from the cells, though the Western blot performed by my colleague on cell lysate from CHO-AIM cells suggests the former. Also, intermolecular and intramolecular disulphide bonds of many secretory proteins, e.g. insulin, are known to form in the oxidative environment found in the lumen of the rough endoplasmatic reticulum. Thus, it seems likely that ApoA-I_M homodimers could also form in the endoplasmatic reticulum of CHO-AIM cells.

Alternatively, the fixation process could have masked the ApoA-I_M antibody epitope on the ApoA-I_M homodimers, since the formaldehyde in the fixation buffer cross-links amino side chains, particularly of lysine, with close-by nitrogen atoms. Other fixation reagents that immobilize proteins by precipitation, such as ethanol or methanol, might preserve the antibody epitope. Optimization of the fixation method should at least determine if the ApoA-I_M antibody can detect ApoA-I_M homodimers in the cells. However, the time might be better spent on establishing the cell surface capture method.

Interestingly, the ApoA-I antibody did not detect any ApoA-I in CHO-AI cells but did give a signal in CHO-AIM cells, although Western blotting of supernatant with the same antibody showed ApoA-I secretion from CHO-AI cells. This would suggest that ApoA-I is rapidly secreted from the CHO-AI cells, whereas ApoA-I_M is largely retained in CHO-AIM cells. Indeed, impaired secretion of ApoA-I_M has been described in hepatocytes from ApoA-I_M homo- and heterozygous knock-in mice.[312] Also, improperly folded proteins are retained in cells,[313] so perhaps the ApoA-I_M homodimer is recognized as such and not secreted as a consequence.

If intracellular ApoA-I is indeed negligible, intracellular staining with the ApoA-I antibody could be used to quantify ApoA-I_M gene editing efficiencies and optimize hepatic targeting *in vivo*. However, recovery of live gene-edited cells using the cell surface affinity-matrix capture method described by Manz *et al.*[311] requires a discriminatory antibody, which might still be possible with the ApoA-I_M antibody if its failure to bind specifically was due to the fixation method. These results are clearly encouraging, at least in regards to the possibility of quantifying gene editing of ApoA-I_M using flow cytometry, but the ultimate aim is to sort and clone gene-edited cells.

CHAPTER 9: GENERAL DISCUSSION

9.1 Summary

Atherosclerosis is the underlying reason for three quarters of deaths caused by cardiovascular disease equating to ~20% of deaths worldwide. Reverse cholesterol transport mediated by HDL and ApoA-I plays an important role in limiting the formation of atherosclerotic plaque in the body.[314] ApoA-I_M,[64] a point mutated form of ApoA-I, is particularly atheroprotective as demonstrated by injection of recombinant ApoA-I_M/ phospholipid complexes in a clinical phase II trial[58] as well as evidence from *in vitro*[68-70,72] and preclinical *in vivo* studies.[59-62] However, a recombinant protein therapy requiring tens of grams per treatment would be prohibitively expensive for widespread use, but gene therapy could be a cost effective alternative.

Here, I evaluated the use of oligonucleotide-mediated gene editing to introduce the ApoA-I_M mutation into the ApoA-I sequence in mammalian cells. After initial methodological problems with the PCR-RFLP analysis, I demonstrated successful gene editing in an ApoA-I expressing recombinant CHO cell line and also of the endogenous ApoA-I gene in HepG2 cells. I excluded the possibility that the diagnostic bands in the PCR-RFLP were due to analysis artefacts and confirmed by sequencing the presence of gene-edited cells in the population for up to 10 days after treatment.

However, I was unable to isolate viable gene-edited cell clones by limiting dilution and this, in addition to the problems with quantitative assessment of PCR-RFLP results, led me to employ the mutated EGFP reporter gene system in further experiments. In these experiments, I investigated the influence of the transfection reagents Lipofectamine 2000,[146] LPEI,[137] fully deacylated LPEI[140] and CDAN/DOPE[148] on gene editing efficiencies. Additional experiments assessed the newly-developed non-toxic biopolymer transfection reagents atelocollagen[144] and tri-methylated chitosan.[141,142] I found that Lipofectamine 2000 mediated significantly higher gene editing levels than any of the other reagents and that the biopolymers did not lead to any gene editing.

Furthermore, I examined the influence of phosphorothioate (PTO) and locked nucleic acid (LNA) oligonucleotide backbone modifications on editing efficiencies and viability of edited cells compared to unmodified (DNA) oligonucleotides. My results demonstrated that although PTO- and LNA-modified ssODNs produced higher levels of gene-edited cells initially, the viability of gene-edited cells treated with these protected oligonucleotides was significantly reduced. Thus, there was a trade-off between toxicity of the backbone modifications and viability of edited cells. Other experiments showed that 24 h after transfection was the optimal sampling time point and that an oligonucleotide with three mismatches performed better than a two mismatch ssODN.

I also used the mEGFP system to investigate ways of increasing gene editing efficiencies and selecting for gene-edited cells independent of a marker gene. Unfortunately, my findings showed that published data suggesting an increase with repeated targeting[184] were due to an artefact in sampling time and did not in fact lead to an increase in gene editing. On the other hand, co-transfection of targeting oligonucleotide with an RFP expressing plasmid seemed to increase gene editing efficiency, but curiously there were no double positive cells in the treated population. Thus, plasmid co-transfection could not be used to select for gene-edited cells independent of the targeted marker gene. Two other attempts at selecting gene-edited cells involved sorting for G2/M arrested cells and FRET between the targeting ssODN and genomic DNA. Disappointingly, both attempts failed due to technical difficulties.

In other experiments, I addressed the lack of incontrovertible proof that oligonucleotide-mediated gene editing can introduce specific mutations into the genomic sequence of mammalian cells. Rigorously following the four criteria proposed by Albuquerque-Silva *et al.*,[227] I demonstrated that three specific mismatches were introduced into the genomic sequence of gene-edited cells.

Moreover, the gene-edited genotype was stable even though green fluorescent cells were lost from the population of gene-edited clones. I showed that this was due to instability of the phenotype caused by epigenetic down-regulation of the multiple copy transgene array in the CHO-mEGFP cells. The observation that a large majority of CHO-mEGFP cells did not express mEGFP led me to believe that I had underestimated the gene editing efficiency in these cells. Thus, I investigated what percentage of the expressing cells was gene-edited.

Not surprisingly, I found that a much higher percentage of cells in the expressing subpopulation were gene-edited. In fact, gene editing efficiency was positively associated with expression of the target gene. This result was also confirmed in HEK293T-mEGFP and HepG2-mEGFP cells.

Coming to the end of my studies, I intended to return to targeting ApoA-I in HepG2 and recombinant CHO-AI cells with the aim of producing isogenic cell lines containing the ApoA-I_M mutation. Since my previous attempts using limited dilution were unsuccessful, I hoped to use flow cytometry for selection of gene-edited cells in these experiments. Unfortunately, the ApoA-I_M homodimer antibody did not appear to bind specifically to the protein in fixed cells.

Interestingly, a control ApoA-I antibody gave a specific signal in CHO-AIM, but not in CHO-AI cells even though both cell lines expressed the respective protein variants. This suggested that ApoA-I_M is retained to a much higher degree in cells than ApoA-I. Thus, flow cytometric quantitation of ApoA-I gene editing is possible. Thus, although I did not manage to establish gene-edited isogenic cell lines expressing the super-atheroprotective ApoA-I_M variant in this study, I made significant progress towards it.

9.2 Implications of my results for the chimeraplasty controversy

During my studies I encountered and overcame a number of problems that could potentially explain why other laboratories have reported negative results. For example, I observed heteroduplex formation during the evaluation of gene editing efficiencies with PCR-RFLP (section 3.3). It is a distinct possibility that only this heteroduplex formation enabled visual identification of the diagnostic bands in my experiments.

Since my PCR-RFLP was based on the destruction of a restriction site, the heteroduplex band co-located with the diagnostic band. In PCR-RFLP analyses based on the creation of a restriction site, the heteroduplex band would not be distinguishable from the bands characteristic for non-edited cells. Formation of heteroduplexes is dependent on PCR conditions[286] which can vary widely even within the same laboratory.

My experiments (section 4.1.1) also highlight the importance of using the optimal transfection reagent for oligonucleotide-mediated gene editing. A recently published study by Afifi *et al.*[233] reported negative results when skin fibroblasts were targeted with chimeraplasts and ssODNs using DMSO to permeabilize the cells. My results suggest that this failure was probably due to insufficient transfection efficiency. Unfortunately, no data studying the intracellular location or concentration of the oligonucleotides were presented in this publication, so no definite conclusion can be drawn.

I also found that PTO and LNA oligonucleotide backbone modifications initially mediate higher gene editing levels, but reduce viability of gene-edited cells significantly compared to unmodified ssODNs (section 4.3.1). Location of the backbone modifications seems to influence gene editing efficiencies as well (I. Papaioannou, paper submitted). Moreover, the number of mismatches between the gene editing ssODN and the target sequence significantly influenced editing efficiencies (section 4.4.1). The importance of the mismatch location[246] in the oligonucleotide and strand-bias[267] has already been reviewed in section 1.4.5.3.

In other experiments I showed that gene editing efficiency is positively associated with expression of the target gene. Thus, endogenous genes that are expressed at low levels are likely to be edited less efficiently. For example, correction of the L444P mutation in the acid β-glucosidase gene of skin fibroblasts from a Gaucher disease patient was unsuccessful using chimeraplasts,[229] but this gene is transcribed at 6- fold lower levels in skin fibroblasts than in a glioblastoma cell line and 8-fold lower than in astrocytoma cells.[315]

Also, unstable expression in the CHO-mEGFP cells was due to epigenetic down-regulation of the transgene array and led to a significant underestimation of gene editing efficiencies in these cells. This finding was confirmed in two other recombinant mEGFP cells lines. Although continuous application of selection pressure can ameliorate the problem of loss of expression,[305] it is common practice to withdraw selection pressure after isogenic cells have been isolated.

Furthermore, the large majority of reporter gene cell lines used in oligonucleotide-mediated gene editing research contain multiple copies of the respective transgenes. Thus, differing levels of epigenetic down-regulation in such cell lines could account for some of the variation in reported gene editing efficiencies. Previous research in our laboratory has also shown that the quality and stability of chimeraplasts has to be taken into account when negative results are obtained.[237] Taken together, these points could well explain the difficulties of reproducibility associated with chimeraplasty.

Notably, all published negative results described the use of chimeraplasts with the sole exception of the study by Afifi *et al.*[233] already discussed above. So far, no other negative data has been published on using single-stranded oligonucleotides for gene editing. On the contrary, a number of independent research groups from the United Kingdom,[183,249] Norway,[242,251] France,[184,254] The Netherlands,[243,260] Germany,[241,247,259] USA[248,256] and China[246,277] have published positive results when using ssODNs instead of chimeraplasts. Furthermore, rigorously applying the four criteria suggested by Albuquerque-Silva *et al.*[227] for gene repair validation, I demonstrated that oligonucleotide-mediated gene editing with ssODNs can introduce three specific nucleotide changes into the genomic sequence of mammalian cells (section 6.2) and that these alterations lead to a stable inheritable genotype (section 7.1).

Thus, I propose the following guidelines for oligonucleotide-mediated gene editing research. First, ssODNs should be used instead of chimeraplasts. Second, PCR-RFLP analysis should be avoided because of heteroduplex formation and problematic quantitation. Potential alternatives include flow cytometry if mutation specific antibodies are available, or allele specific quantitative RT-PCR. Third, transfection efficiency and in particular nuclear localization of the targeting ssODN must be evaluated and optimized if necessary. Fourth, the design of the targeting oligonucleotide such as backbone modifications, mismatch number and location and length must be optimized, as only a few firm rules are available so far. Fifth, the expression level of the target gene should be controlled for especially in recombinant cell lines, but also when targeting endogenous genes. If possible the expression levels should be maximized.

9.3 Further Research

Due to time constraints, I was unable to follow up on several interesting research avenues. For example, I intended to return to targeting the endogenous ApoA-I gene in CHO-AI and HepG2 cells and to isolate gene-edited ApoA-I_M expressing cell clones using flow cytometry, but was hampered by the non-specific binding of the ApoA-I_M homodimer antibody (section 8.2). However, the data suggest that the formaldehyde-based fixation I used may have hidden the antibody epitope and consequently other fixation methods should be evaluated.

Use of the antibody in unfixed cells employing the cell surface capture method described by Manz *et al.*[311] should also be attempted. Furthermore, I showed that the ApoA-I antibody gives a specific signal in formaldehyde-fixed CHO-AIM cells. Thus, gene editing efficiencies of ApoA-I to ApoA-I_M can be quantified by flow cytometry, avoiding the problems with PCR-RFLP analysis described above.

On a completely different note, I showed that I could introduce three specific nucleotide alterations with the same targeting oligonucleotide (section 6.2) and other groups have shown similar data.[247,255] These results raise the question of how specific oligonucleotide-mediated gene editing is when targeting endogenous genes. The answer could potentially limit the application of this method for gene therapy, but even a negative result would still allow generation of gene-edited cells for research into genetic diseases and SNPs. The specificity of gene editing could perhaps be evaluated by recovering biotin-labelled oligonucleotides from the genomic DNA of gene-edited cells, in experiments similar to how Radecke *et al.*[259] showed physical integration of the ssODN. Random priming could be used to amplify the selected sequences or alternatively, the sequences could be subcloned into bacterial plasmids.

Further research into ways of selecting or enriching for gene-edited cells independent of the target gene could be particularly rewarding. Two of my attempts in this direction failed due to technical difficulties (chapter 5), but further optimization might overcome these difficulties. Also, results reported by Ferrara *et al.*[279] suggest that CHK1 and CHK2 are phosphorylated only in the gene-edited subpopulation of treated cells.

Confocal microscopy and flow cytometry could be employed to analyse activation of the ATM/ATR DNA damage and repair pathway in gene-edited cells with an emphasis on CHK1 and CHK2 phosphorylation. If the results by Ferrara *et al.*[279] are confirmed, an intracellular kinase assay which is currently being developed by Invitrogen could be used for fluorescence activated cell sorting of cells according to their phosphorylated CHK1 or CHK2 status. Thus, the need for mutation specific antibodies would be circumvented.

Temporary down-regulating of the expression of certain proteins in the DNA damage and repair pathways by siRNA could also increase gene editing efficiencies. Of particular interest are the mismatch repair pathway proteins, MSH2 and MSH3,[243,260] which have been implicated in the suppression of gene editing. It has been suggested that this is due to MSH2 and MSH3 preventing homologous recombination between sequences containing non-homologous nucleotides.[260,261] However, research in our laboratory suggests that MSH2 is activated at a later stage of gene editing, after the oligonucleotide is integrated into the genomic sequence (I. Papaioannou, paper submitted). In particular, there seems to be a positive influence of MSH2 suppression on viability of gene-edited cells, but these findings need to be studied in more detail.

A better understanding of the DNA damage response pathways involved in oligonucleotide-mediated gene editing will hopefully lead to enhanced editing efficiencies and thus more wide-spread application of this method not only in gene therapy, but also in the generation of transgenic cell and animal systems.

BIBLIOGRAPHY

1. [BHF] British Heart Foundation (9-5-2006) 2006 Coronary Heart Disease statistics. <http://www.heartstats.org/datapage.asp?id=5739>. Accessed 28-4-2007.

2. [WHO] World Health Organisation (2007) Factsheet: Cardiovascular diseases. <http://www.who.int/mediacentre/factsheets/fs317/en/index.html>. Accessed 28-4-2007.

3. Lusis AJ (2000) Atherosclerosis. *Nature* 407(6801):233-41.

4. Ross R (1999) Atherosclerosis--an inflammatory disease. *N Engl J Med* 340(2):115-26.

5. Cunningham KS, Gotlieb AI (2005) The role of shear stress in the pathogenesis of atherosclerosis. *Lab Invest* 85(1):9-23.

6. Wilcox JN *et al.* (1997) Expression of multiple isoforms of nitric oxide synthase in normal and atherosclerotic vessels. *Arterioscler Thromb Vasc Biol* 17(11):2479-88.

7. Napoli C *et al.* (1997) Fatty streak formation occurs in human fetal aortas and is greatly enhanced by maternal hypercholesterolemia. Intimal accumulation of low density lipoprotein and its oxidation precede monocyte recruitment into early atherosclerotic lesions. *J Clin Invest* 100(11):2680-90.

8. Osterud B, Bjorklid E (2003) Role of monocytes in atherogenesis. *Physiol Rev* 83(4):1069-112.

9. Rouis M (2005) Matrix metalloproteinases: a potential therapeutic target in atherosclerosis. *Curr Drug Targets Cardiovasc Haematol Disord* 5(6):541-8.

10. Gordon T *et al.* (1977) High-density lipoprotein as a protective factor against coronary heart-disease - Framingham study. *Am J Med* 62(5):707-14.

11. Barter PJ *et al.* (2004) Antiinflammatory properties of HDL. *Circ Res* 95(8):764-72.

12. Glomset JA (1968) The plasma lecithins:cholesterol acyltransferase reaction. *J Lipid Res* 9(2):155-67.

13. Schaefer EJ *et al.* (1982) Plasma apolipoprotein A-1 absence associated with a marked reduction of high density lipoproteins and premature coronary artery disease. *Arteriosclerosis* 2(1):16-26.

14. Brooks-Wilson A *et al.* (1999) Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat Genet* 22(4):336-45.

15. Bojanovski D *et al.* (1987) In vivo metabolism of proapolipoprotein A-I in Tangier disease. *J Clin Invest* 80(6):1742-7.

16. Serfaty-Lacrosniere C *et al.* (1994) Homozygous Tangier disease and cardiovascular disease. *Atherosclerosis* 107(1):85-98.

17. Hovingh GK *et al.* (2005) Compromised LCAT function is associated with increased atherosclerosis. *Circulation* 112(6):879-84.

18. Hovingh GK *et al.* (2005) Inherited disorders of HDL metabolism and atherosclerosis. *Curr Opin Lipidol* 16(2):139-45.

19. Martin DD *et al.* (2006) Apolipoprotein A-I assumes a "looped belt" conformation on reconstituted high density lipoprotein. *J Biol Chem* 281(29):20418-26.

20. Barter P *et al.* (2003) High density lipoproteins (HDLs) and atherosclerosis; the unanswered questions. *Atherosclerosis* 168(2):195-211.

21. Wang N *et al.* (2004) ATP-binding cassette transporters G1 and G4 mediate cellular cholesterol efflux to high-density lipoproteins. *Proc Natl Acad Sci USA* 101(26):9774-9.

22. Rader DJ (2006) Molecular regulation of HDL metabolism and function: implications for novel therapies. *J Clin Invest* 116(12):3090-100.

23. Rye KA, Barter PJ (2004) Formation and metabolism of prebeta-migrating, lipid-poor apolipoprotein A-I. *Arterioscler Thromb Vasc Biol* 24(3):421-8.

24. Acton S *et al.* (1996) Identification of scavenger receptor SR-BI as a high density lipoprotein receptor. *Science* 271(5248):518-20.

25. Gordon DJ *et al.* (1989) High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies. *Circulation* 79(1):8-15.

26. Barter PJ, Rye KA (1996) High density lipoproteins and coronary heart disease. *Atherosclerosis* 121(1):1-12.

27. Badimon JJ *et al.* (1989) High density lipoprotein plasma fractions inhibit aortic fatty streaks in cholesterol-fed rabbits. *Lab Invest* 60(3):455-61.

28. Badimon JJ, Badimon L, Fuster V (1990) Regression of atherosclerotic lesions by high density lipoprotein plasma fraction in the cholesterol-fed rabbit. *J Clin Invest* 85(4):1234-41.

29. Duverger N *et al.* (1996) Inhibition of atherosclerosis development in cholesterol-fed human apolipoprotein A-I-transgenic rabbits. *Circulation* 94(4):713-7.

30. Rubin EM *et al.* (1991) Inhibition of early atherogenesis in transgenic mice by human apolipoprotein AI. *Nature* 353(6341):265-7.

31. Plump AS, Scott CJ, Breslow JL (1994) Human apolipoprotein A-I gene expression increases high density lipoprotein and suppresses atherosclerosis in the apolipoprotein E-deficient mouse. *Proc Natl Acad Sci USA* 91(20):9607-11.

32. Huttunen JK *et al.* (1991) The Helsinki heart-study - central findings and clinical implications. *Ann Med* 23(2):155-9.

33. Rubins HB *et al.* (1999) Gemfibrozil for the secondary prevention of coronary heart disease in men with low levels of high-density lipoprotein cholesterol. *New Engl J Med* 341(6):410-8.

34. Schlesinger Z *et al.* (2000) Secondary prevention by raising HDL cholesterol and reducing triglycerides in patients with coronary artery disease - The bezafibrate infarction prevention (BIP) study. *Circulation* 102(1):21-7.

35. Torra IP *et al.* (2001) Peroxisome proliferator-activated receptors: from transcriptional control to clinical practice. *Curr Opin Lipidol* 12(3):245-54.

36. Inazu A *et al.* (1990) Increased high-density-lipoprotein levels caused by a common cholesteryl-ester transfer protein gene mutation. *New Engl J Med* 323(18):1234-8.

37. Ikewaki K *et al.* (1993) Delayed catabolism of high-density-lipoprotein apolipoprotein-A-I and apolipoprotein-A-Ii in human cholesteryl ester transfer protein-deficiency. *J Clin Invest* 92(4):1650-8.

38. Shah PK (2007) Inhibition of CETP as a novel therapeutic strategy for reducing the risk of atherosclerotic disease. *Eur Heart J* 28(1):5-12.

39. Morehouse LA *et al.* (2007) Inhibition of CETP activity by torcetrapib reduces susceptibility to diet-induced atherosclerosis in New Zealand White rabbits. *J Lipid Res* 48(6):1263-72.

40. Brousseau ME *et al.* (2004) Effects of an inhibitor of cholesteryl ester transfer protein on HDL cholesterol. *New Engl J Med* 350(15):1505-15.

41. Clark RW *et al.* (2004) Raising high-density lipoprotein in humans through inhibition of cholesteryl ester transfer protein: An initial multidose study of torcetrapib. *Arterioscler Thromb Vasc Biol* 24(3):490-7.

42. Tall AR, Yvan-Charvet L, Wang N (2007) The failure of torcetrapib - Was it the molecule or the mechanism? *Arterioscler Thromb Vasc Biol* 27(2):257-60.

43. Barter PJ *et al.* (2007) Effects of torcetrapib in patients at high risk for coronary events. *New Engl J Med* 357(21):2109-22.

44. Nissen SE *et al.* (2007) Effect of torcetrapib on the progression of coronary atherosclerosis. *New Engl J Med* 356(13):1304-16.

45. Bots ML *et al.* (2007) Torcetrapib and carotid intima-media thickness in mixed dyslipidaemia (RADIANCE 2 study): a randomised, double-blind trial. *Lancet* 370(9582):153-60.

46. Kastelein JJP *et al.* (2007) Effect of torcetrapib on carotid atherosclerosis in familial hypercholesterolemia. *New Engl J Med* 356(16):1620-30.

47. Krishna R *et al.* (2007) Effect of the cholesteryl ester transfer protein inhibitor, anacetrapib, on lipoproteins in patients with dyslipidaemia and on 24-h ambulatory blood pressure in healthy individuals: two double-blind, randomised placebo-controlled phase I studies. *Lancet* 370(9603):1907-14.

48. Bisoendial RJ *et al.* (2005) Consequences of cholesteryl ester transfer protein inhibition in patients with familial hypoalphalipoproteinemia. *Arterioscler Thromb Vasc Biol* 25(9):E133-E134.

49. Kuivenhoven JA *et al.* (2005) Effectiveness of inhibition of cholesteryl ester transfer protein by JTT-705 in combination with Pravastatin in type II dyslipidemia. *Am J Cardiol* 95(9):1085-8.

50. Carlson LA (1995) Effect of a single infusion of recombinant human proapolipoprotein A-I liposomes (synthetic HDL) on plasma-lipoproteins in patients with low high-density-lipoprotein cholesterol. *Nutr Metab Cardiovas* 5(2):85-91.

51. Eriksson M *et al.* (1999) Stimulation of fecal steroid excretion after infusion of recombinant proapolipoprotein A-I. Potential reverse cholesterol transport in humans. *Circulation* 100(6):594-8.

52. Nanjee MN *et al.* (2001) Intravenous apoA-I/lecithin discs increase pre-beta-HDL concentration in tissue fluid and stimulate reverse cholesterol transport in humans. *J Lipid Res* 42(10):1586-93.

53. Nanjee MN *et al.* (1999) Acute effects of intravenous infusion of ApoA1/phosphatidylcholine discs on plasma lipoproteins in humans. *Arterioscler Thromb Vasc Biol* 19(4):979-89.

54. Nanjee MN *et al.* (1996) Effects of intravenous infusion of lipid-free apo A-I in humans. *Arterioscler Thromb Vasc Biol* 16(9):1203-14.

55. Spieker LE *et al.* (2002) High-density lipoprotein restores endothelial function in hypercholesterolemic men. *Circulation* 105(12):1399-402.

56. Bisoendial RJ *et al.* (2003) Restoration of endothelial function by increasing high-density lipoprotein in subjects with isolated low high-density lipoprotein. *Circulation* 107(23):2944-8.

57. Tardif JC *et al.* (2007) Effects of reconstituted high-density lipoprotein infusions on coronary atherosclerosis: a randomized controlled trial. *JAMA* 297(15):1675-82.

58. Nissen SE *et al.* (2003) Effect of recombinant ApoA-I Milano on coronary atherosclerosis in patients with acute coronary syndromes: a randomized controlled trial. *JAMA* 290(17):2292-300.

59. Ibanez B *et al.* (2008) Rapid change in plaque size, composition, and molecular footprint after recombinant apolipoprotein A-I_{Milano} (ETC-216) administration - Magnetic resonance imaging study in an experimental model of atherosclerosis. *J Am Coll Cardiol* 51(11):1104-9.

60. Parolini C *et al.* (2008) Dose-related effects of repeated ETC-216 (recombinant apolipoprotein A-I_{Milano}/1-palmitoyl-2-oleoyl phosphatidylcholine complexes) administrations on rabbit lipid-rich soft plaques - In vivo assessment by intravascular ultrasound and magnetic resonance imaging. *J Am Coll Cardiol* 51(11):1098-103.

61. Shah PK *et al.* (2001) High-dose recombinant apolipoprotein A-I_{Milano} mobilizes tissue cholesterol and rapidly reduces plaque lipid and macrophage content in apolipoprotein E-deficient mice - Potential implications for acute plaque stabilization. *Circulation* 103(25):3047-50.

62. Kaul S *et al.* (2004) Rapid reversal of endothelial dysfunction in hypercholesterolemic apolipoprotein E-null mice by recombinant apolipoprotein A-I-Milano-phospholipid complex. *J Am Coll Cardiol* 44(6):1311-9.

63. Franceschini G *et al.* (1980) A-I_{Milano} apoprotein. Decreased high density lipoprotein cholesterol levels with significant lipoprotein modifications and without clinical atherosclerosis in an Italian family. *J Clin Invest* 66(5):892-900.

64. Weisgraber KH *et al.* (1980) A-I_{Milano} apoprotein. Isolation and characterization of a cysteine-containing variant of the A-I apoprotein from human high density lipoproteins. *J Clin Invest* 66(5):901-7.

65. Weisgraber KH *et al.* (1983) Apolipoprotein A-I_{Milano}. Detection of normal A-I in affected subjects and evidence for a cysteine for arginine substitution in the variant A-I. *J Biol Chem* 258(4):2508-13.

66. Gualandri V *et al.* (1985) AI_{Milano} apoprotein identification of the complete kindred and evidence of a dominant genetic transmission. *Am J Hum Genet* 37(6):1083-97.

67. Sirtori CR *et al.* (2001) Cardiovascular status of carriers of the apolipoprotein A-I_{Milano} mutant: the Limone sul Garda study. *Circulation* 103(15):1949-54.

68. Franceschini G *et al.* (1999) Increased cholesterol efflux potential of sera from ApoA-I_{Milano} carriers and transgenic mice. *Arterioscler Thromb Vasc Biol* 19(5):1257-62.

69. Calabresi L *et al.* (1999) Cell cholesterol efflux to reconstituted high-density lipoproteins containing the apolipoprotein A-I_{Milano} dimer. *Biochemistry* 38(49):16307-14.

70. Favari E *et al.* (2007) A unique protease-sensitive high density lipoprotein particle containing the apolipoprotein A-I_{Milano} dimer effectively promotes ATP-binding Cassette A1-mediated cell cholesterol efflux. *J Biol Chem* 282(8):5125-32.

71. Bielicki JK *et al.* (1997) Evidence that apolipoprotein A-I_{Milano} has reduced capacity, compared with wild-type apolipoprotein A-I, to recruit membrane cholesterol. *Arterioscler Thromb Vasc Biol* 17(9):1637-43.

72. Weibel GL *et al.* (2007) Wild-type ApoA-I and the Milano variant have similar abilities to stimulate cellular lipid mobilization and efflux. *Arterioscler Thromb Vasc Biol* 27(9):2022-9.

73. Wang WQ, Moses AS, Francis GA (2001) Cholesterol mobilization by free and lipid-bound apoAI_{Milano} and apoAI_{Milano}-apoAII heterodimers. *Biochemistry* 40(12):3666-73.

74. Westman J *et al.* (1998) Sterol 27-hydroxylase- and ApoAI/phospholipid-mediated efflux of cholesterol from cholesterol-laden macrophages - Evidence for an inverse relation between the two mechanisms. *Arterioscler Thromb Vasc Biol* 18(4):554-61.

75. Zhu X *et al.* (2005) Cysteine mutants of human apolipoprotein A-I: a study of secondary structural and functional properties. *J Lipid Res* 46(6):1303-11.

76. Perez-Mendez O *et al.* (2000) Metabolism of apolipoproteins AI and AII in subjects carrying similar apoAI mutations, apoAI Milano and apoAI Paris. *Atherosclerosis* 148(2):317-26.

77. Roma P *et al.* (1993) In vivo metabolism of a mutant form of apolipoprotein A-I, apo A-I_{Milano}, associated with familial hypoalphalipoproteinemia. *J Clin Invest* 91(4):1445-52.

78. Calabresi L *et al.* (1997) Activation of lecithin cholesterol acyltransferase by a disulfide-linked apolipoprotein A-I dimer. *Biochem Bioph Res Co* 232(2):345-9.

79. Bielicki JK, Oda MN (2002) Apolipoprotein A-I_{Milano} and apolipoprotein A-I_{Paris} exhibit an antioxidant activity distinct from that of wild-type apolipoprotein A-I. *Biochemistry* 41(6):2089-96.

80. Gomaraschi M *et al.* (2007) Normal vascular function despite low levels of high-density lipoprotein cholesterol in carriers of the apolipoprotein A-I_{Milano} mutant. *Circulation* 116(19):2165-72.

81. Friedmann T, Roblin R (1972) Gene therapy for human genetic disease. *Science* 175(4025):949-55.

82. Edelstein ML, Abedi MR, Wixon J (2007) Gene therapy clinical trials worldwide to 2007-an update. *J Gene Med* 9(10):833-42.

83. Gragoudas ES *et al.* (2004) Pegaptanib for neovascular age-related macular degeneration. *N Engl J Med* 351(27):2805-16.

84. Garrick D *et al.* (1998) Repeat-induced gene silencing in mammals. *Nat Genet* 18(1):56-9.

85. Ellis J (2005) Silencing and variegation of gammaretrovirus and lentivirus vectors. *Hum Gene Ther* 16(11):1241-6.

86. Chen ZY *et al.* (2004) Silencing of episomal transgene expression by plasmid bacterial DNA elements in vivo. *Gene Ther* 11(10):856-64.

87. Hong K, Sherley J, Lauffenburger DA (2001) Methylation of episomal plasmids as a barrier to transient gene expression via a synthetic delivery vector. *Biomol Eng* 18(4):185-92.

88. Weeratna RD *et al.* (2001) Designing gene therapy vectors: avoiding immune responses by using tissue-specific promoters. *Gene Ther* 8(24):1872-8.

89. Scrable H, Stambrook PJ (1997) Activation of the lac repressor in the transgenic mouse. *Genetics* 147(1):297-304.

90. Riu E *et al.* (2005) Increased maintenance and persistence of transgenes by excision of expression cassettes from plasmid sequences in vivo. *Hum Gene Ther* 16(5):558-70.

91. Brown BD *et al.* (2007) A microRNA-regulated lentiviral vector mediates stable correction of hemophilia B mice. *Blood* 110(13):4144-52.

92. Thomas CE, Ehrhardt A, Kay MA (2003) Progress and problems with the use of viral vectors for gene therapy. *Nat Rev Genet* 4(5):346-58.

93. Kay MA, Glorioso JC, Naldini L (2001) Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nat Med* 7(1):33-40.

94. Cavazzana-Calvo M *et al.* (2000) Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* 288(5466):669-72.

95. Hacein-Bey-Abina S *et al.* (2002) Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *New Engl J Med* 346(16):1185-93.

96. Hacein-Bey-Abina S *et al.* (2003) A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *New Engl J Med* 348(3):255-6.

97. Hacein-Bey-Abina S *et al.* (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302(5644):415-9.

98. Blesch A (2004) Lentiviral and MLV based retroviral vectors for ex vivo and in vivo gene transfer. *Methods* 33(2):164-72.

99. Carlotti F *et al.* (2004) Lentiviral vectors efficiently transduce quiescent mature 3T3-L1 adipocytes. *Mol Ther* 9(2):209-17.

100. Bukrinsky MI, Haffar OK (1999) HIV-1 nuclear import: in search of a leader. *Front Biosci* 4:D772-D781.

101. Levine BL *et al.* (2006) Gene transfer in humans using a conditionally replicating lentiviral vector. *Proc Natl Acad Sci USA* 103(46):17372-7.

102. Palmer JA *et al.* (2000) Development and optimization of herpes simplex virus vectors for multiple long-term gene delivery to the peripheral nervous system. *J Virol* 74(12):5604-18.

103. Krisky DM *et al.* (1998) Development of herpes simplex virus replication-defective multigene vectors for combination gene therapy applications. *Gene Ther* 5(11):1517-30.

104. Samaniego LA, Neiderhiser L, Deluca NA (1998) Persistence and expression of the herpes simplex virus genome in the absence of immediate-early proteins. *J Virol* 72(4):3307-20.

105. Hu JCC *et al.* (2006) A phase I study of OncoVEX(GM-CSF), a second-generation oncolytic herpes simplex virus expressing granulocyte macrophage colony-stimulating factor. *Clin Cancer Res* 12(22):6737-47.

106. Senzer, N. N. *et al.* (2008) Phase II clinical trial with a second generation, GM-CSF encoding, oncolytic herpesvirus in unresectable metastatic melanoma. <http://www.asco.org/ASCO/Abstracts+%26+Virtual+Meeting/Abstracts?&vmview=abst_detail_view&confID=55&abstractID=33352>. Accessed 5-7-2008.

107. [Anon] (2002) Assessment of adenoviral vector safety and toxicity: Report of the National Institutes of Health Recombinant DNA Advisory Committee. *Hum Gene Ther* 13(1):3-13.

108. Amalfitano A *et al.* (1998) Production and characterization of improved adenovirus vectors with the E1, E2b, and E3 genes deleted. *J Virol* 72(2):926-33.

109. Alba R, Bosch A, Chillon M (2005) Gutless adenovirus: last-generation adenovirus for gene therapy. *Gene Ther* 12:S18-S27.

110. Nakai H *et al.* (2003) AAV serotype 2 vectors preferentially integrate into active genes in mice. *Nat Genet* 34(3):297-302.

111. Nakai H *et al.* (2005) Large-scale molecular characterization of adeno-associated virus vector integration in mouse liver. *J Virol* 79(6):3606-14.

112. Gao GP, Vandenberghe LH, Wilson JM (2005) New recombinant serotypes of AAV vectors. *Curr Gene Ther* 5(3):285-97.

113. Rabinowitz JE *et al.* (2002) Cross-packaging of a single adeno-associated virus (AAV) type 2 vector genome into multiple AAV serotypes enables transduction with broad specificity. *J Virol* 76(2):791-801.

114. Carter BJ (2005) Adeno-associated virus vectors in clinical trials. *Hum Gene Ther* 16(5):541-50.

115. Piechaczek C *et al.* (1999) A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells. *Nucleic Acids Res* 27(2):426-8.

116. Jenke BHC *et al.* (2002) An episomally replicating vector binds to the nuclear matrix protein SAF-A in vivo. *EMBO Rep* 3(4):349-54.

117. Jenke AC *et al.* (2004) Expression of a transgene encoded on a non-viral episomal vector is not subject to epigenetic silencing by cytosine methylation. *Mol Biol Rep* 31(2):85-90.

118. Liu HZ, Visner GA (2007) Applications of Sleeping Beauty transposons for nonviral gene therapy. *IUBMB Life* 59(6):374-9.

119. Liu L, Mah C, Fletcher BS (2006) Sustained FVIII expression and phenotypic correction of hemophilia A in neonatal mice using an endothelial-targeted Sleeping Beauty transposon. *Mol Ther* 13(5):1006-15.

120. Fischer SEJ, Wienholds E, Plasterk RHA (2001) Regulated transposition of a fish transposon in the mouse germ line. *Proc Natl Acad Sci USA* 98(12):6759-64.

121. Wilber A *et al.* (2007) Efficient and stable transgene expression in human embryonic stem cells using transposon-mediated gene transfer. *Stem Cells* 25(11):2919-27.

122. Ivics Z *et al.* (1997) Molecular reconstruction of Sleeping beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 91(4):501-10.

123. Yant SR *et al.* (2007) Site-directed transposon integration in human cells. *Nucleic Acids Res* 35(7):e50.

124. Patil SD, Rhodes DG, Burgess DJ (2005) DNA-based therapeutics and DNA delivery systems: A comprehensive review. *Aaps J* 7(1):E61-E77.

125. Pouton CW, Seymour LW (2001) Key issues in non-viral gene delivery. *Adv Drug Deliv Rev* 46(1-3):187-203.

126. Pack DW *et al.* (2005) Design and development of polymers for gene delivery. *Nat Rev Drug Discov* 4(7):581-93.

127. Niidome T, Huang L (2002) Gene therapy progress and prospects: nonviral vectors. *Gene Ther* 9(24):1647-52.

128. Lin MTS *et al.* (2000) The gene gun: current applications in cutaneous gene therapy. *Int J Dermatol* 39(3):161-70.

129. Neumann E *et al.* (1982) Gene-transfer into mouse lyoma cells by electroporation in high electric-fields. *EMBO J* 1(7):841-5.

130. Wells DJ (2004) Gene therapy progress and prospects: Electroporation and other physical methods. *Gene Ther* 11(18):1363-9.

131. Gordon JW *et al.* (1980) Genetic-transformation of mouse embryos by micro-injection of purified DNA. *Proc Natl Acad Sci USA* 77(12):7380-4.

132. Wang W *et al.* (2007) A fully automated robotic system for microinjection of zebrafish embryos. *PLoS ONE* 2(9):e862.

133. Zhang G *et al.* (2004) Hydroporation as the mechanism of hydrodynamic delivery. *Gene Ther* 11(8):675-82.

134. Zhang G *et al.* (2001) Efficient expression of naked DNA delivered intraarterially to limb muscles of nonhuman primates. *Hum Gene Ther* 12(4):427-38.

135. Fabre JW *et al.* (2008) Hydrodynamic gene delivery to the pig liver via an isolated segment of the inferior vena cava. *Gene Ther* 15(6):452-62.

136. Jordan M, Schallhorn A, Wurm FM (1996) Transfecting mammalian cells: Optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic Acids Res* 24(4):596-601.

137. Boussif O *et al.* (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in-vivo - polyethylenimine. *Proc Natl Acad Sci USA* 92(16):7297-301.

138. Akinc A *et al.* (2005) Exploring polyethylenimine-mediated DNA transfection and the proton sponge hypothesis. *J Gene Med* 7(5):657-63.

139. Lv HT *et al.* (2006) Toxicity of cationic lipids and cationic polymers in gene delivery. *J Control Release* 114(1):100-9.

140. Thomas M *et al.* (2005) Full deacylation of polyethylenimine dramatically boosts its gene delivery efficiency and specificity to mouse lung. *Proc Natl Acad Sci USA* 102(16):5679-84.

141. Borchard G (2001) Chitosans for gene delivery. *Adv Drug Deliv Rev* 52(2):145-50.

142. Thanou M *et al.* (2002) Quaternized chitosan oligomers as novel gene delivery vectors in epithelial cell lines. *Biomaterials* 23(1):153-9.

143. Murata J, Ohya Y, Ouchi T (1997) Design of quaternary chitosan conjugate having antennary galactose residues as a gene delivery tool. *Carbohyd Polym* 32:105-9.

144. Sano A *et al.* (2003) Atelocollagen for protein and gene delivery. *Adv Drug Deliv Rev* 55(12):1651-77.

145. Nakamura M *et al.* (2004) Targeted conversion of the transthyretin gene in vitro and in vivo. *Gene Ther* 11(10):838-46.

146. Dalby B *et al.* (2004) Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. *Methods* 33(2):95-103.

147. Keller M *et al.* (2003) Thermodynamic aspects and biological profile of CDAN/DOPE and DC-Chol/DOPE lipoplexes. *Biochemistry* 42(20):6067-77.

148. Spagnou S, Miller AD, Keller M (2004) Lipidic carriers of siRNA: differences in the formulation, cellular uptake, and delivery with plasmid DNA. *Biochemistry* 43(42):13348-56.

149. Dominski Z, Kole R (1993) Restoration of Correct Splicing in Thalassemic Premessenger Rna by Antisense Oligonucleotides. *Proc Natl Acad Sci USA* 90(18):8673-7.

150. Stein CA *et al.* (1988) Physicochemical Properties of Phosphorothioate Oligodeoxynucleotides. *Nucleic Acids Res* 16(8):3209-21.

151. Wahlestedt C *et al.* (2000) Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc Natl Acad Sci USA* 97(10):5633-8.

152. Summerton J, Weller D (1997) Morpholino antisense oligomers: Design, preparation, and properties. *Antisense Nucleic A* 7(3):187-95.

153. Bunka DHJ, Stockley PG (2006) Aptamers come of age - at last. *Nat Rev Microbiol* 4(8):588-96.

154. Plasterk RHA (2002) RNA silencing: The genome's immune system. *Science* 296(5571):1263-5.

155. Jackson AL *et al.* (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* 12(7):1179-87.

156. Crooke ST (1998) Vitravene (TM) - another piece in the mosaic. *Antisense Nucleic A* 8(4):VII-VIII.

157. Medical News TODAY (3-6-2008) Clinical Proof Of Concept For TGF-beta 2-Inhibitor AP 12009 In Phase IIb / EMEA Provides Guidance On Phase III Design And Approval Requirements. <http://www.medicalnewstoday.com/articles/109606.php>. Accessed 29-6-2008.

158. Daneholt, B. (2008) The Nobel Prize in Physiology or Medicine 2006. <http://nobelprize.org/nobel_prizes/medicine/laureates/2006/adv.html>. Accessed 15-7-2008.

159. [NIH] National Institutes of Health (8-5-2008) Safety and Efficacy Study of Antisense Oligonucleotides in Duchenne Muscular Dystrophy. <http://clinicaltrials.gov/ct2/show/NCT00159250?show_desc=Y#desc>. Accessed 10-7-2008.

160. Rhie A *et al.* (2003) Characterization of 2 '-fluoro-RNA aptamers that bind preferentially to disease-associated conformations of prion protein and inhibit conversion. *J Biol Chem* 278(41):39697-705.

161. Dey AK *et al.* (2005) An aptamer that neutralizes R5 strains of human immunodeficiency virus type 1 blocks gp120-CCR5 interaction. *J Virol* 79(21):13806-10.

162. Bhindi R *et al.* (2007) Brothers in arms - DNA enzymes, short interfering RNA, and the emerging wave of small-molecule nucleic acid-based gene-silencing strategies. *Am J Pathol* 171(4):1079-88.

163. Vasquez KM *et al.* (2001) Manipulating the mammalian genome by homologous recombination. *Proc Natl Acad Sci USA* 98(15):8403-10.

164. Olsen PA. Targeted sequence alterations in the genomes of mammalian cells mediated by oligonucleotides [dissertation]. University of Oslo; 2005.

165. Jiricny J (2006) The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* 7(5):335-46.

166. Fleck O, Nielsen O (2004) DNA repair. *J Cell Sci* 117(Pt 4):515-7.

167. Sancar A *et al.* (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem* 73:39-85.

168. Takata M *et al.* (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J* 17(18):5497-508.

169. Lombard DB *et al.* (2005) DNA repair, genome stability, and aging. *Cell* 120(4):497-512.

170. Bakkenist CJ, Kastan MB (2004) Initiating cellular stress responses. *Cell* 118(1):9-17.

171. Capecchi MR (2005) Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet* 6(6):507-12.

172. The Nobel Foundation (2008) The Nobel Prize in Physiology or Medicine 2007. <nobelprize.org>. Accessed 20-7-0008.

173. Thomas KR, Capecchi MR (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem-cells. *Cell* 51(3):503-12.

174. LePage DF, Conlon RA (2006) Animal models for disease: knockout, knock-in, and conditional mutant mice. *Methods Mol Med* 129:41-67.

175. Kunzelmann K *et al.* (1996) Gene targeting of CFTR DNA in CF epithelial cells. *Gene Ther* 3(10):859-67.

176. Goncz KK *et al.* (1998) Targeted replacement of normal and mutant CFTR sequences in human airway epithelial cells using DNA fragments. *Hum Mol Genet* 7(12):1913-9.

177. Kapsa R *et al.* (2001) In vivo and in vitro correction of the mdx dystrophin gene nonsense mutation by short-fragment homologous replacement. *Hum Gene Ther* 12(6):629-42.

178. Goncz KK *et al.* (2002) Application of SFHR to gene therapy of monogenic disorders. *Gene Ther* 9(11):691-4.

179. Sangiuolo F *et al.* (2005) In vitro restoration of functional SMN protein in human trophoblast cells affected by spinal muscular atrophy by small fragment homologous replacement. *Hum Gene Ther* 16(7):869-80.

180. Tsuchiya H, Harashima H, Kamiya H (2005) Factors affecting SFHR gene correction efficiency with single-stranded DNA fragment. *Biochem Bioph Res Co* 336(4):1194-200.

181. de Semir D, Aran JM (2003) Misleading gene conversion frequencies due to a PCR artifact using small fragment homologous replacement. *Oligonucleotides* 13(4):261-9.

182. Thorpe P, Stevenson BJ, Porteous DJ (2002) Optimising gene repair strategies in cell culture. *Gene Ther* 9(11):700-2.

183. Nickerson HD, Colledge WH (2003) A comparison of gene repair strategies in cell culture using a lacZ reporter system. *Gene Ther* 10(18):1584-91.

184. Andrieu-Soler C *et al.* (2005) Stable transmission of targeted gene modification using single-stranded oligonucleotides with flanking LNAs. *Nucleic Acids Res* 33(12):3733-42.

185. Colosimo A *et al.* (2001) Targeted correction of a defective selectable marker gene in human epithelial cells by small DNA fragments. *Mol Ther* 3(2):178-85.

186. Ellis J, Bernstein A (1989) Gene targeting with retroviral vectors - recombination by gene conversion into regions of nonhomology. *Mol Cell Biol* 9(4):1621-7.

187. Mitani K *et al.* (1995) Gene targeting in mouse embryonic stem-cells with an adenoviral vector. *Somat Cell Molec Gen* 21(4):221-31.

188. Wang Q, Taylor MW (1993) Correction of a deletion mutant by gene targeting with an adenovirus vector. *Mol Cell Biol* 13(2):918-27.

189. Russell DW, Hirata RK (1998) Human gene targeting by viral vectors. *Nat Genet* 18(4):325-30.

190. Hendrie PC, Russell DW (2005) Gene targeting with viral vectors. *Mol Ther* 12(1):9-17.

191. Inoue N, Hirata RK, Russell DW (1999) High-fidelity correction of mutations at multiple chromosomal positions by adeno-associated virus vectors. *J Virol* 73(9):7376-80.

192. Miller DG *et al.* (2006) Gene targeting in vivo by adeno-associated virus vectors. *Nat Biotechnol* 24(8):1022-6.

193. Vasileva A, Linden RM, Jessberger R (2006) Homologous recombination is required for AAV-mediated gene targeting. *Nucleic Acids Res* 34(11):3345-60.

194. Schultz BR, Chamberlain JS (2008) Recombinant adeno-associated virus transduction and integration. *Mol Ther* 16(7):1189-99.

195. Rouet P, Smih F, Jasin M (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* 14(12):8096-106.

196. Rouet P, Smih F, Jasin M (1994) Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci USA* 91(13):6064-8.

197. Smih F *et al.* (1995) Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res* 23(24):5012-9.

198. Kim YG, Chandrasegaran S (1994) Chimeric restriction-endonuclease. *Proc Natl Acad Sci USA* 91(3):883-7.

199. Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci USA* 93(3):1156-60.

200. Wu J, Kandavelou K, Chandrasegaran S (2007) Custom-designed zinc finger nucleases: what is next? *Cell Mol Life Sci* 64(22):2933-44.

201. Porteus MH, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science* 300(5620):763.

202. Urnov FD *et al.* (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 435(7042):646-51.

203. Cathomen T, Joung JK (2008) Zinc-finger nucleases: The next generation emerges. *Mol Ther* 16(7):1200-7.

204. Radecke F *et al.* (2006) Targeted chromosomal gene modification in human cells by single-stranded oligodeoxynucleotides in the presence of a DNA double-strand break. *Mol Ther* 14(6):798-808.

205. Cooney M *et al.* (1988) Site-specific oligonucleotide binding represses transcription of the human c-myc gene in vitro. *Science* 241(4864):456-9.

206. Wang G, Seidman MM, Glazer PM (1996) Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science* 271(5250):802-5.

207. Faruqi AF *et al.* (2000) Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway. *Mol Cell Biol* 20(3):990-1000.

208. Faruqi AF *et al.* (1996) Recombination induced by triple-helix-targeted DNA damage in mammalian cells. *Mol Cell Biol* 16(12):6820-8.

209. Knauert MP *et al.* (2006) Triplex-stimulated intermolecular recombination at a single-copy genomic target. *Mol Ther* 14(3):392-400.

210. Cannata F *et al.* (2008) Triplex-forming oligonucleotide-orthophenanthroline conjugates for efficient targeted genome modification. *Proc Natl Acad Sci USA* 105(28):9576-81.

211. Simon P *et al.* (2008) Targeting DNA with triplex-forming oligonucleotides to modify gene sequence. *Biochimie* [Epub ahead of print].

212. Yoon K, ColeStrauss A, Kmiec EB (1996) Targeted gene correction of episomal DNA in mammalian cells mediated by a chimeric RNA:DNA oligonucleotide. *Proc Natl Acad Sci USA* 93(5):2071-6.

213. ColeStrauss A *et al.* (1996) Correction of the mutation responsible for sickle cell anemia by an RNA-DNA oligonucleotide. *Science* 273(5280):1386-9.

214. Thomas KR, Capecchi MR (1997) Recombinant DNA technique and sickle cell anemia research. *Science* 275(5305):1404-5.

215. Stasiak A, West SC, Egelman EH (1997) Sickle cell anemia research and a recombinant DNA technique. *Science* 277(5325):460-2.

216. Kren BT *et al.* (1997) Targeted nucleotide exchange in the alkaline phosphatase gene of HuH-7 cells mediated by a chimeric RNA/DNA oligonucleotide. *Hepatology* 25(6):1462-8.

217. Kren BT, Bandyopadhyay P, Steer CJ (1998) In vivo site-directed mutagenesis of the factor IX gene by chimeric RNA/DNA oligonucleotides. *Nat Med* 4(3):285-90.

218. Zhang ZP *et al.* (1998) Failure to achieve gene conversion with chimeric circular oligonucleotides: Potentially misleading PCR artifacts observed. *Antisense Nucleic A* 8(6):531-6.

219. Strauss M (1998) The site-specific correction of genetic defects. *Nat Med* 4(3):274-5.

220. Alexeev V, Yoon K (1998) Stable and inheritable changes in genotype and phenotype of albino melanocytes induced by an RNA-DNA oligonucleotide. *Nat Biotechnol* 16(13):1343-6.

221. Alexeev V *et al.* (2000) Localized in vivo genotypic and phenotypic correction of the albino mutation in skin by RNA-DNA oligonucleotide. *Nat Biotechnol* 18(1):43-7.

222. Rando TA, Disatnik MH, Zhou LZH (2000) Rescue of dystrophin expression in mdx mouse muscle by RNA/DNA oligonucleotides. *Proc Natl Acad Sci USA* 97(10):5363-8.

223. Bertoni C, Rando TA (2002) Dystrophin gene repair in mdx muscle precursor cells in vitro and in vivo mediated by RNA-DNA chimeric oligonucleotides. *Hum Gene Ther* 13(6):707-18.

224. Suzuki T, Murai A, Muramatsu T (2003) Low-dose bleomycin induces targeted gene repair frequency in cultured melan-c cells using chimeric RNA/DNA oligonucleotide transfection. *Int J Mol Med* 12(1):109-14.

225. Li ZH *et al.* (2001) Targeted correction of the point mutations of beta-thalassemia and targeted mutagenesis of the nucleotide associated with HPFH by RNA/DNA oligonucleotides: Potential for beta-thalassemia gene therapy. *Blood Cell Mol Dis* 27(2):530-8.

226. van der Steege G *et al.* (2001) Persistent failures in gene repair. *Nat Biotechnol* 19(4):305-6.

227. Albuquerque-Silva J *et al.* (2001) Chimeraplasty validation. *Nat Biotechnol* 19(11):1011.

228. Tran ND *et al.* (2003) Efficiency of chimeraplast gene targeting by direct nuclear injection using a GFP recovery assay. *Mol Ther* 7(2):248-53.

229. Diaz-Font A *et al.* (2003) Unsuccessful chimeraplast strategy for the correction of a mutation causing Gaucher disease. *Blood Cell Mol Dis* 31(2):183-6.

230. Nickerson HD, Colledge WH (2004) A LacZ-based transgenic mouse for detection of somatic gene repair events in vivo. *Gene Ther* 11(17):1351-7.

231. Ino A *et al.* (2004) Somatic gene targeting with RNA/DNA chimeric oligonucleotides: an analysis with a sensitive reporter mouse system. *J Gene Med* 6(11):1272-80.

232. De Meyer SF *et al.* (2007) False positive results in chimeraplasty for von Willebrand Disease. *Thromb Res* 119(1):93-104.

233. Afifi A *et al.* (2008) Failure to repair the C.338C > T mutation in carnitine palmitoyl transferase 2 deficient skin fibroblasts using chimeraplasty. *Mol Genet Metab* 93(3):347-9.

234. Tagalakis AD *et al.* (2001) Gene correction of the apolipoprotein (Apo) E2 phenotype to wild-type ApoE3 by in situ chimeraplasty. *J Biol Chem* 276(16):13226-30.

235. Tagalakis AD *et al.* (2005) Correction of the neuropathogenic human apolipoprotein E4 (APOE4) gene to APOE3 in vitro using synthetic RNA/DNA oligonucleotides (chimeraplasts). *J Mol Neurosci* 25(1):95-103.

236. Mohri Z *et al.* (2005) Gene editing of the wild-type ApoE3 gene to the dysfunctional variants ApoE2 or ApoE4 using synthetic RNA-DNA oligonucleotides (chimeraplasts). *Gene Ther Mol Biol* 9(B):143-52.

237. Manzano A *et al.* (2003) Failure to generate atheroprotective apolipoprotein AI phenotypes using synthetic RNA/DNA oligonucleotides (chimeraplasts). *J Gene Med* 5(9):795-802.

238. Campbell CR *et al.* (1989) Homologous recombination involving small single-stranded oligonucleotides in human cells. *New Biol* 1(2):223-7.

239. Gamper HB *et al.* (2000) The DNA strand of chimeric RNA/DNA oligonucleotides can direct gene repair/conversion activity in mammalian and plant cell-free extracts. *Nucleic Acids Res* 28(21):4332-9.

240. Igoucheva O, Alexeev V, Yoon K (2001) Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. *Gene Ther* 8(5):391-9.

241. Radecke F, Radecke S, Schwarz K (2004) Unmodified oligodeoxynucleotides require single-strandedness to induce targeted repair of a chromosomal EGFP gene. *J Gene Med* 6(11):1257-71.

242. Olsen PA, Randol M, Krauss S (2005) Implications of cell cycle progression on functional sequence correction by short single-stranded DNA oligonucleotides. *Gene Ther* 12(6):546-51.

243. Dekker M, Brouwers C, te RH (2003) Targeted gene modification in mismatch-repair-deficient embryonic stem cells by single-stranded DNA oligonucleotides. *Nucleic Acids Res* 31(6):e27.

244. Murphy BR *et al.* (2006) Delivery and mechanistic considerations for the production of knock-in mice by single-stranded oligonucleotide gene targeting. *Gene Ther* 14(4):304-15.

245. Pierce EA *et al.* (2003) Oligonucleotide-directed single-base DNA alterations in mouse embryonic stem cells. *Gene Ther* 10(1):24-33.

246. Yin WX *et al.* (2005) Targeted correction of a chromosomal point mutation by modified single-stranded oligonucleotides in a GFP recovery system. *Biochem Bioph Res Co* 334(4):1032-41.

247. Hegele H *et al.* (2008) Simultaneous targeted exchange of two nucleotides by single-stranded oligonucleotides clusters within a region of about fourteen nucleotides. *BMC Mol Biol* 9:14.

248. Morozov V, Wawrousek EF (2007) Single-strand DNA-mediated targeted mutagenesis of genomic DNA in early mouse embryos is stimulated by Rad51/54 and by Ku70/86 inhibition. *Gene Ther* 15(6):468-72.

249. Thorpe PH, Stevenson BJ, Porteous DJ (2002) Functional correction of episomal mutations with short DNA fragments and RNA-DNA oligonucleotides. *J Gene Med* 4(2):195-204.

250. Olsen PA, McKeen C, Krauss S (2003) Branched oligonucleotides induce in vivo gene conversion of a mutated EGFP reporter. *Gene Ther* 10(21):1830-40.

251. Olsen PA *et al.* (2005) Genomic sequence correction by single-stranded DNA oligonucleotides: role of DNA synthesis and chemical modifications of the oligonucleotide ends. *The Journal of Gene Medicine* 7(12):1534-44.

252. Seidman MM, Glazer PM (2004) Setting standards in gene repair. *Oligonucleotides* 14(2):79.

253. Sorensen CB *et al.* (2005) Site-specific strand bias in gene correction using single-stranded oligonucleotides. *J Mol Med* 83(1):39-49.

254. De Piedoue G *et al.* (2007) Targeted gene correction with 5 ' acridine-oligonucleotide conjugates. *Oligonucleotides* 17(2):258-63.

255. Agarwal S, Gamper HB, Kmiec EB (2003) Nucleotide replacement at two sites can be directed by modified single-stranded oligonucleotides in vitro and in vivo. *Biomol Eng* 20(1):7-20.

256. Flagler K *et al.* (2008) Site-specific gene modification by oligodeoxynucleotides in mouse bone marrow-derived mesenchymal stem cells. *Gene Ther* 15(14):1035-48.

257. Drury MD, Skogen MJ, Kmiec EB (2005) A tolerance of DNA heterology in the mammalian targeted gene repair reaction. *Oligonucleotides* 15(3):155-71.

258. Takahashi N, Dawid IB (2005) Characterization of zebrafish Rad52 and replication protein A for oligonucleotide-mediated mutagenesis. *Nucleic Acids Res* 33(13):e120.

259. Radecke S *et al.* (2006) Physical incorporation of a single-stranded oligodeoxynucleotide during targeted repair of a human chromosomal locus. *The Journal of Gene Medicine* 8(2):217-28.

260. Dekker M *et al.* (2006) Effective oligonucleotide-mediated gene disruption in ES cells lacking the mismatch repair protein MSH3. *Gene Ther* 13(8):686-94.

261. Aarts M *et al.* (2006) Generation of a mouse mutant by oligonucleotide-mediated gene modification in ES cells. *Nucleic Acids Res* 34(21):e147.

262. Liu L *et al.* (2002) Strand bias in targeted gene repair is influenced by transcriptional activity. *Mol Cell Biol* 22(11):3852-63.

263. Igoucheva O *et al.* (2003) Transcription affects formation and processing of intermediates in oligonucleotide-mediated gene alteration. *Nucleic Acids Res* 31(10):2659-70.

264. Lu IL *et al.* (2003) Correction/mutation of acid alpha-D-glucosidase gene by modified single-stranded oligonucleotides: in vitro and in vivo studies. *Gene Ther* 10(22):1910-6.

265. Korzheva N *et al.* (2000) A structural model of transcription elongation. *Science* 289(5479):619-25.

266. Huen MSY *et al.* (2007) Active transcription promotes single-stranded oligonucleotide mediated gene repair. *Biochem Bioph Res Co* 353(1):33-9.

267. Brachman EE, Kmiec EB (2004) DNA replication and transcription direct a DNA strand bias in the process of targeted gene repair in mammalian cells. *J Cell Sci* 117(Pt 17):3867-74.

268. Hu YL *et al.* (2005) Reaction parameters of targeted gene repair in mammalian cells. *Mol Biotechnol* 29(3):197-210.

269. Huen MSY *et al.* (2006) The involvement of replication in single stranded oligonucleotide-mediated gene repair. *Nucleic Acids Res* 34(21):6183-94.

270. Ferrara L, Parekh-Olmedo H, Kmiec EB (2004) Enhanced oligonucleotide-directed gene targeting in mammalian cells following treatment with DNA damaging agents. *Exp Cell Res* 300(1):170-9.

271. Wu XS *et al.* (2005) Increased efficiency of oligonucleotide-mediated gene repair through slowing replication fork progression. *Proc Natl Acad Sci USA* 102(7):2508-13.

272. Brachman EE, Kmiec EB (2005) Gene repair in mammalian cells is stimulated by the elongation of S phase and transient stalling of replication forks. *DNA Repair (Amst)* 4(4):445-57.

273. Ferrara L, Kmiec EB (2004) Camptothecin enhances the frequency of oligonucleotide-directed gene repair in mammalian cells by inducing DNA damage and activating homologous recombination. *Nucleic Acids Res* 32(17):5239-48.

274. Suzuki T, Murai A, Muramatsu T (2005) Bleomycin stimulates targeted gene repair directed by single-stranded oligodeoxynucleotides in BHK-21 cells. *Int J Mol Med* 16(4):615-20.

275. Schwartz TR, Kmiec EB (2006) Using methyl methanesulfonate (MMS) to stimulate targeted gene repair activity in mammalian cells. *Gene Ther Mol Biol* 9:193-202.

276. Saleh-Gohari N, Helleday T (2004) Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Res* 32(12):3683-8.

277. Wang Z *et al.* (2006) Single-stranded oligonucleotide-mediated gene repair in mammalian cells has a mechanism distinct from homologous recombination repair. *Biochem Bioph Res Co* 350(3):568-73.

278. Kawabe T (2004) G2 checkpoint abrogators as anticancer drugs. *Mol Cancer Ther* 3(4):513-9.

279. Ferrara L, Kmiec EB (2006) Targeted gene repair activates Chk1 and Chk2 and stalls replication in corrected cells. *DNA Repair* 5(4):422-31.

280. Liu L, Maguire KK, Kmiec EB (2004) Genetic re-engineering of Saccharomyces cerevisiae RAD51 leads to a significant increase in the frequency of gene repair in vivo. *Nucleic Acids Res* 32(7):2093-101.

281. Trojan J *et al.* (2002) Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology* 122(1):211-9.

282. Parekh-Olmedo H, Kmiec EB (2007) Progress and Prospects: targeted gene alteration (TGA). *Gene Ther* 14(24):1675-80.

283. Igoucheva O *et al.* (2006) Involvement of ERCC1/XPF and XPG in Oligodeoxynucleotide-directed Gene Modification. *Oligonucleotides* 16(1):94-104.

284. Jarvis R (2005) Optimizing siRNA transfection for RNAi. *Ambion TechNotes* 12(1):18-20.

285. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132:365-86.

286. Thompson JR, Marcelino LA, Polz MF (2002) Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nucleic Acids Res* 30(9):2083-8.

287. Engstrom JU, Kmiec EB (2007) Manipulation of cell cycle progression can counteract the apparent loss of correction frequency following oligonucleotide-directed gene repair. *BMC Mol Biol* 8(1):9.

288. Kovala AT *et al.* (2000) High-efficiency transient transfection of endothelial cells for functional analysis. *FASEB J* 14(15):2486-94.

289. Chen R *et al.* (1999) Enrichment of transiently transfected mesangial cells by cell sorting after cotransfection with GFP. *Am J Physiol* 276(5 Pt 2):F777-F785.

290. Igoucheva O, Alexeev V, Yoon K (2006) Differential cellular responses to exogenous DNA in mammalian cells and its effect on oligonucleotide-directed gene modification. *Gene Ther* 13(3):266-75.

291. Yanez RJ, Porter AC (1999) Influence of DNA delivery method on gene targeting frequencies in human cells. *Somat Cell Molec Gen* 25(1):27-31.

292. Khazanov E, Simberg D, Barenholz Y (2006) Lipoplexes prepared from cationic liposomes and mammalian DNA induce CpG-independent, direct cytotoxic effects in cell cultures and in mice. *J Gene Med* 8(8):998-1007.

293. Levin AA (1999) A review of the issues in the pharmacokinetics and toxicology of phosphorothioate antisense oligonucleotides. *Biochim Biophys Acta* 1489(1):69-84.

294. Koziolkiewicz M *et al.* (2001) The mononucleotide-dependent, nonantisense mechanism of action of phosphodiester and phosphorothioate oligonucleotides depends upon the activity of an ecto-5'-nucleotidase. *Blood* 98(4):995-1002.

295. Selinger CI, Day CJ, Morrison NA (2005) Optimized transfection of diced siRNA into mature primary human osteoclasts: Inhibition of cathepsin K mediated bone resorption by siRNA. *J Cell Biochem* 96(5):996-1002.

296. Escriou V *et al.* (2001) Critical assessment of the nuclear import of plasmid during cationic lipid-mediated gene transfer. *J Gene Med* 3(2):179-87.

297. Nur-E-Kamal *et al.* (2003) Single-stranded DNA induces ataxia telangiectasia mutant (ATM)/p53-dependent DNA damage and apoptotic signals. *J Biol Chem* 278(14):12475-81.

298. Bradford, J. A *et al.* (2007) Cell Cycle Analysis in Live Cells Using Novel Vybrant DyeCycle Stains. <http://probes.invitrogen.com/media/publications/611.pdf>. Accessed 15-1-2008.

299. Manders EM, Kimura H, Cook PR (1999) Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. *J Cell Biol* 144(5):813-21.

300. Bruckmann E, Wojcik A, Obe G (1999) Sister chromatid differentiation with biotin-dUTP. *Chromosome Res* 7(3):185-9.

301. Choi KH *et al.* (2005) Activation of CMV promoter-controlled glycosyltransferase and beta - galactosidase glycogenes by butyrate, tricostatin A, and 5-aza-2'-deoxycytidine. *Glycoconj J* 22(1-2):63-9.

302. McBurney MW *et al.* (2002) Evidence for repeat-induced gene silencing in cultured mammalian cells: Inactivation of tandem repeats of transfected genes. *Exp Cell Res* 274(1):1-8.

303. Liu HS *et al.* (1999) Is green fluorescent protein toxic to the living cells? *Biochem Bioph Res Co* 260(3):712-7.

304. Dixit R, Cyr R (2003) Cell damage and reactive oxygen species production induced by fluorescence microscopy: effect on mitosis and guidelines for non-invasive fluorescence microscopy. *Plant J* 36(2):280-90.

305. Liu WM, Xiong YZ, Gossen M (2006) Stability and homogeneity of transgene expression in isogenic cells. *J Mol Med* 84(1):57-64.

306. Robins DM *et al.* (1981) Transforming DNA integrates into the host chromosome. *Cell* 23(1):29-39.

307. Huttner KM *et al.* (1981) DNA-mediated gene-transfer without carrier DNA. *J Cell Biol* 91(1):153-6.

308. Pikaart MI, 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):2852-62.

309. Thyagarajan B, Johnson BL, Campbell C (1995) The effect of target site transcription on gene targeting in human cells in-vitro. *Nucleic Acids Res* 23(14):2784-90.

310. Macris MA, Glazer PM (2003) Transcription dependence of chromosomal gene targeting by triplex-forming oligonucleotides. *J Biol Chem* 278(5):3357-62.

311. Manz R *et al.* (1995) Analysis and sorting of live cells according to secreted molecules, relocated to a cell-surface affinity matrix. *Proc Natl Acad Sci USA* 92(6):1921-5.

312. Parolini C *et al.* (2003) Targeted replacement of mouse apolipoprotein A-I with human ApoA-I or the mutant ApoA-I_{Milano}. Evidence of APOA-I_M impaired hepatic secretion. *J Biol Chem* 278(7):4740-6.

313. Hendershot LM (2000) Giving protein traffic the green light. *Nat Cell Biol* 2(6):E105-E106.

314. Curtiss LK *et al.* (2006) What is so special about apolipoprotein AI in reverse cholesterol transport? *Arterioscler Thromb Vasc Biol* 26(1):12-9.

315. Doll RF, Smith FI (1993) Regulation of Expression of the Gene Encoding Human Acid Beta-Glucosidase in Different Cell-Types. *Gene* 127(2):255-60.