CONTRIBUTION OF DISTAL PROMOTER ELEMENTS TO TRANSCRIPTIONAL REGULATION OF GLUCOSE-DEPENDENT INSULINOTROPIC POLYPEPTIDE IN INTESTINAL STC-1 CELLS

by

IRENE LOK YAN YU

B.Sc. (Hons), The University of British Columbia, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Physiology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2009

ABSTRACT

Diabetes mellitus is a group of chronic metabolic disorders that are characterized by high blood glucose resulting from a lack of or insufficient secretion of insulin, which is a source of medical and financial burden to more than 285 million people worldwide. Current treatments for diabetes include lifestyle modifications, medication, and insulin therapy, but these treatments do not save patients from diabetic complications including blindness, limb amputations, circulatory disorders, and increased risk of developing kidney failure, cardiovascular diseases, and neuropathies. Glucose-dependent insulinotropic polypeptide (GIP) is a gastrointestinal hormone that plays an integral role in the finely-tuned secretion of insulin following a meal, and the cells that express GIP have demonstrated potential for being a target for insulin gene therapy. Understanding how the GIP gene is regulated will provide insights into the defining characteristics of GIP-expressing cells and how these can be harnessed for therapy. In the present study, two enhancer *cis*-regulatory elements which accounted for 40-65% of GIP promoter activity were identified in a previously uncharacterized well-conserved region of the distal 5' upstream rat GIP promoter by a series of luciferase reporter studies. Pax6 and Pdx1, two transcription factors that have been previously shown to be important for GIP expression, were shown to bind at these sites using electrophoretic mobility shift assays, mutational analysis, and chromatin immunoprecipitation. The development of a fluorescence-based isolation technique for primary GIP-expressing cells was documented. Cell numbers (20,000 – 35,000) were purified for the isolation of RNA in sufficient quantity and quality (80-140 ng, and RNA integrity number = 6.8-7.9, respectively) for microarray. The feasibility of isolating primary

GIP-expressing cells presents a model which would allow for non-biased screening for the identification of additional *trans*-regulatory elements which may act at well-established and newly characterized *cis*-regulatory elements.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

$\%Hb_{A1c}$	percent glycosylated haemoglobin A1c
β-gal	β-galactosidase (gene product of *LacZ*)
Ab	antibody
ANOVA	analysis of variance
bHLH	basic helix-loop-helix
BP	biotinylated probe
BSA	bovine serum albumin
cAMP	cyclic adenosine monophosphate
cDNA	complementary deoxyribonucleic acid
ChIP	chromatin immunoprecipitation
CP	cold probe
DAPI	4', 6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
DPPIV	dipeptidyl peptidase IV
dsRed2	rapid maturing variant of red fluorescent protein from *Discosoma sp.*
dsRed2+	expresses dsRed2
dsRed2-	does not express dsRed2
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EMSA	electrophoretic mobility shift assay
FACS	fluorescence-activated cell sorting
GATA4	GATA-binding protein 4
GATA6	GATA-binding protein 6
GIP	glucose-dependent insulinotropic polypeptide
GLP-1	glucagon-like peptide 1
HBSS	Hank's buffered saline solution
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hes1	hairy and enhancer of split 1

IgG	immunoglobulin G
Isl1	LIM homeobox islet-1
KCl	potassium chloride
KH_2PO_4	potassium dihydrogen phosphate
LacZ	gene encoding β-galactosidase
Math1	atonal homolog 1
$MgCl_2$	magnesium chloride
Na_2HPO_4	sodium phosphate
NE	nuclear extract
Ngn3	neurogenin-3
NOD	non-obese diabetic
Pax6	paired homeobox 6
PBS	phosphate buffered saline
PC1/3	prohormone convertase 1/3
PCR	polymerase chain reaction
Pdx1	pancreatic and duodenal homeobox 1
RNA	ribonucleic acid
S-MEM	minimum essential medium for suspension cultures
SV40	simian virus 40
TBE	Tris-Borate-EDTA buffer
TCF4	T-cell specific, HMG-box transcription factor 4
WT	wild type

ACKNOWLEDGEMENTS

The writing of this thesis marks the end of five memorable years as a student in the Physiology program at UBC. I am grateful to my teachers, mentors, and friends, who have helped make this journey what it was. Thanks to:

Dr. Honglin Luo and Dr. Doris Doudet, for letting me take my first stab at research in your laboratories. The foundations that I developed under your guidance were instrumental in my successes as a graduate student.

Dr. Tim Kieffer, for taking me on as a student in your lab and for allowing me to explore my varied interests there.

Drs. Jim Johnson, Bruce Verchere, and Chris McIntosh, for your expert advice and for helping me stay on track throughout my numerous project changes.

This work would not have been possible without the patient technical assistance of the UBC and TFL Flow Cytometry Core Staff.

To my mentors who have seen me through school, thank you for your advice and confidence in me.

Dr. Jochen Seufert, thank you for welcoming me so warmly to your lab and giving me the opportunity to see a little bit more of the world.

Dr. Sally Osborne, grad school would not have been the same without PHYL 303.

Dr. Lucy Marzban and Dr. Majid Mojibian, thank you for being such excellent role models for me. Your encouragement during tough times and your perspectives on science as a career and balancing career and family were extremely precious to me.

To the past and present members of the Kieffer lab, thank you for making the lab the dynamic place it is. In particular, I thank present and former bench neighbours and occupants of the grad office, especially Gary, Blair, Ada, Frank, Mike, Maddy, and Anastasia for your camaraderie. Jeannie, for sharing your secrets to good lab practices, stories of mishaps and weirdos, and your photography with me; Rhonda, for your guidance that first summer and beyond; Travis, for all your behind-the-scenes work; Rob "Calvin Klone" Baker, for taking me on as your little apprentice, and for four leaf clovers, horseshoes, and rabbit feet.

My graduate school journey would not have been complete without the support of friends and family. Yalin and Kelrie, thank you for always being there to lend an ear even though we weren't even in the same city most of time. Tatiana, Teodora, Nicole, Bartek, and Susanne, Freiburg would never have felt like home without you. Felix, thank you for giving us a chance – here's to the future!

I dedicate this work to Mom and Dad, who made sacrifices to help me find my place in the world.

INTRODUCTION

Diabetes mellitus is a chronic group of metabolic disorders that are characterized by hyperglycaemia resulting from a lack of insulin (Type 1) or insufficient insulin secretion (Type 2). First described as a disease of excessive urination, diabetes mellitus has been documented as early as 1500 B.C. in ancient Egyptian scrolls, and its incidence since ancient times has steadily increased to the point that it is now a global epidemic afflicting more than 285 million people worldwide (1, 2). By 2025, the number of people living with diabetes mellitus is projected to increase to 400 million (2). The impact of health care costs and resources incurred by diabetes mellitus and its corollary conditions on the global health care budget is enormous. The costs associated with managing diabetes mellitus and diabetes-mellitus-related complications have increased every year in the past decades and continue to increase at faster-than-expected rates. By 2007, diabetes mellitus and diabetes mellitus-related complications were already costing Americans $174 billion, an amount that was initially predicted to not be reached until 2010 (3). The immense burden that diabetes mellitus and diabetes mellitus-related complications place on governments and health care systems and the alarming rate of increase in diabetes mellitus incidence prompted the establishment of a United Nations Resolution in 2006, which strives to support governments in implementing preventative programs that would slow the rate of increase of diabetes mellitus incidence. The devastating medical and economic toll of diabetes mellitus has become an impetus for physicians and scientists alike to pursue better treatments and hopes of finding a cure, long before the mechanisms of the disease were even formally recognized.

The Incretin Concept

In his treatise on the history of diabetes, Papaspyros pointed out that the relationship between diet and health has been recognized since ancient times, and a connection between

diabetes mellitus and diet was suspected long before a formal understanding of metabolism (1). As early as the 17th century, the first account of treating diabetes by undernourishment was documented (1). In the late 1800's, the pancreas and pancreatic islets were anatomically characterized, but it was not until the early 1900s that an understanding encompassing a link between the gut and the pancreas emerged (1).

This understanding coincided with the discovery of the first hormone, secretin, by William Bayliss and Ernest Starling in 1902 (4). Prior to the discovery of secretin, communication between organs was accredited almost exclusively to nervous interactions, and the concept of a hormone, or chemical messenger capable of mediating communication between organs, introduced a new way of perceiving how signalling between distant organs can occur to produce integrated responses to external stimuli. The idea that the pancreas secreted an internal substance responsible for carbohydrate metabolism, and that deficits in glucose metabolism resulted in diabetes mellitus was already established based on complete and partial pancreatectomy studies (5), and the hypothesis that secretin-containing duodenal extracts could stimulate this internal secretion from the pancreas led to the first attempts at treating diabetes by administering crude duodenal extracts to newly diagnosed diabetic patients by the early 1900's (5). A retrospective review noted that since the extracts were administered orally, any effect due to incretins would have been nullified since the hormones would be degraded in the digestive tract prior to arriving at their site of action (6). Furthermore, the improvements observed in some of the patients following crude extract administration was accredited to the honeymoon phase common in newly-diagnosed type 1 diabetes, characterized by a temporary remission period of normoglycaemia (6). In 1932, the concept of gut-pancreas interaction was refined, and the criteria for these yet unknown gut-derived circulating factors that potentiated insulin secretion called "incretins" was established (6). In order to qualify as an incretin, the candidate hormone

must be released in response to nutrients, and be able to stimulate insulin release at physiologic levels in a glucose-dependent manner. Although no incretin hormones were identified for the next three decades, a number of regulatory peptides in the intestinal tract were discovered in this time that led to the current appreciation of the gut one of the largest endocrine organs (7). With the discovery of insulin and characterization of the pathophysiology of diabetes, the early notion that diet and diabetes were linked was consolidated into what became known as the enteroinsular axis. The term, coined by Unger and Eisentraut in 1969 (8), describes the crosstalk between intestinal factors and the pancreas, including but not limited to, hormonal interactions.

It was not until 1960, when the radioimmunoassay for insulin was invented, that the idea of incretin hormones was further developed. With the ability to directly measure insulin, McIntyre and colleagues made the seminal observation that orally-administered glucose could elicit a larger insulin response than intravenously-administered glucose of the same dose, even though the former route of administration led to a more modest increase in blood glucose levels (9). Multiple studies since (10, 11, 12) have suggested that incretins account for as much as 50-60% of the insulin response. This potentiation became known as the “incretin effect”.

Glucose-dependent Insulinotropic Polypeptide: the First Incretin

The first hormone discovered that met the criteria for classification as an incretin is glucose-dependent insulinotropic polypeptide (GIP). GIP was purified from crude porcine intestinal extracts based on its ability to inhibit gastric acid secretion, and originally named “Gastric Inhibitory Polypeptide” based on this function. Subsequent analysis of GIP action demonstrated its ability to potentiate insulin secretion, and the acronym was modified to reflect its major physiologic role (13, 14, 15, 16, 17).

Structure and Metabolism of GIP

GIP is a 42-amino acid peptide that belongs to the glucagon/secretin superfamily of hormones, although truncated forms of GIP (GIP_{1-39} and GIP_{1-30}), which share comparable biological activity with GIP_{1-42}, have been identified (18). These peptides are derived from a precursor preproGIP, which is encoded by the GIP gene on chromosome 17 in humans (19) and chromosome 10 in rats (20). The human and rat GIP gene share a common structure in that they consist of six exons and five introns (19, 20). Exons 1 and 6 encode for a 5'-untranslated region and 3'-polyadenosine tail, respectively, while exons 2-5 encode for a GIP precursor consisting of a signal peptide and GIP_{1-42} flanked by a N-terminus peptide and a C-terminus peptide (19, 20). Post-translational cleavage by the enzyme PC1/3, releases GIP_{1-42} in its active form (18). Meal-regulated transcription that parallels GIP secretion has been demonstrated using both glucose and fat (21, 22), the two main stimulants of GIP secretion. In addition to direct nutrient stimulation, GIP secretion is regulated by hormonal feedback via insulin and even C-peptide alone reduces glucose-stimulated GIP secretion (18). The overall effect of nutrient stimulation and hormonal feedback results in a rapid release of GIP within 15-30 min of a meal and return to fasting levels after a few hours, a secretion profile that mirrors circulating glucose and insulin levels (18). A return to basal GIP levels is facilitated by the degradation of GIP_{1-42} by the enzyme dipeptidyl peptidase IV (DPPIV), which cleaves away the first two amino acids on the N-terminus end of biologically active GIP, rendering it inactive (23). Degradation by DPPIV results in the short half life of biologically active GIP of 5-7 min (24, 25, 26), a property that allows GIP levels to quickly return to basal levels following stimulation. Since the degradation of GIP by DPPIV is not a meal-regulated process, changes in the levels of biologically active GIP in the circulation following a meal depend primarily on the rate of secretion, which in turn is

dependent on the rate of GIP transcription. It is therefore, not surprising, that there is glucose-dependent and closely coupled meal-stimulated GIP transcription and secretion (21, 22, 27, 28).

Tissue Distribution of GIP

GIP is expressed predominantly in open-type enteroendocrine cells that are scattered in the mucosa of the duodenum and upper jejunum, called K-cells (29, 30, 31), but has also been reported to be expressed in the eye (32), brain (33), and salivary glands (34). K-cells represent one of at least 15 types of enteroendocrine cells in the intestine (35). Enteroendocrine cells make up <1% of the total intestinal epithelial cell population, which also consists of goblet cells, Paneth cells, and enterocytes (35). In the human, rat, and pig intestine, K-cells are most abundant in the duodenum, and decrease in number from the proximal to distal intestine, although GIP-expressing cells are still detected in the proximal ileum (29, 30, 36). Notably, a large subpopulation of GIP-expressing cells, called K/L-cells, simultaneously expresses the other known incretin to date, glucagon-like peptide-1 (GLP-1) (29, 30, 36).

The observation that injection of glucose to the upper intestine can generate an incretin effect equivalent to oral administration of glucose (37) suggests that nutrient-responsive secretion of GIP may occur through direct exposure to nutrients, although the mechanism of glucose sensing is still debated. While some studies suggest that there are similarities between the glucose-sensing mechanisms in the GIP-expressing cells and those of the insulin-expressing pancreatic β-cell (36, 38, 39, 40), a recent study demonstrated that GIP-expressing cells do not utilize glucokinase as a glucose sensor as do pancreatic cells based on the lack of a difference in GIP response in patients carrying a heterozygous mutation in the gene encoding glucokinase compared to normal patients (41), despite a 420-fold expression of the gene encoding glucokinase in an enriched primary GIP-expressing cell population compared to a primary

non-GIP-expressing cell population (42). It must be noted, however, that glucokinase may be haplosufficient, suggested by the minor deficits seen in insulin secretion reported for patients with the glucokinase gene mutations (41), in which case the role of glucokinase as a glucose sensor for GIP secretion cannot be definitively refuted. Another hypothesis was that GIP-expressing cells may detect glucose in a similar process as the tasting of flavours on the tongue (43, 44). The co-localization of α-gustducin, a G-protein which is activated upon binding of the sweet taste receptor made from the T1R2 and T1R3 subunits, with GIP-expressing cells (44) and the observation of an impaired incretin effect in α-gustducin null mice (45) generated by targeted replacement of the α-gustducin exon 1 with a neomycin resistance cassette (46), suggested that glucose sensing in K-cells may occur by activation of the sweet receptors. However, quantitative RT-PCR failed to detect any expression of T1R2 and α-gustducin in an enriched primary GIP-expressing cell population (42), and oral administration of five different artificial sweeteners failed to improve glucose excursion after an intraperitoneal glucose tolerance test (47). A recent study by Moriya *et al.* (37) demonstrated that activation of sodium glucose cotransporter-1, regardless of whether the substrate was a metabolizable or not, was sufficient for initiating GIP and GLP-1 secretion and an incretin effect (37), suggesting that this may be another candidate glucose sensor. In support of this finding is the observation that an enriched primary GIP-expressing cell population expresses high levels of the gene encoding the sodium glucose cotransporter-1 (42) and sodium glucose cotransporter-1 immunoreactivity was reported to be located on the apical surface of cells lining the villi (42). Other candidate glucose sensors include the K_{ATP} subunit Kir6.2 and sulphonylurea receptor Sur1, which are expressed 190 and 470-fold more in an enriched primary GIP-expressing cell population compared to a primary non-GIP-expressing cell population (42).

Given the major physiologic role of GIP in regulating insulin secretion in the β-cell, the observation of GIP expression in glucagon expressing islet cells, is particularly interesting. GIP expression in the pancreatic islet was first reported in 1978 by Alumets *et al.* (48), an observation that was dismissed in subsequent years as cross-immunoreactivity of the antibodies used (49). However, recent studies in our laboratory have demonstrated that the initial observations of GIP localization in the pancreas should be not discredited, since processing of GIP in the α-cell results in the GIP_{1-30} truncated form, which is only recognized by some antibodies but not others (unpublished data). From an evolutionary perspective, the localization of GIP in the islet is not surprising. In invertebrates and prochordates, endocrine cells that secrete pancreatic hormones are located in the gut mucosa, and the existence of the pancreas and intestine as distinct organs coincided with the existence of intestinal incretin hormones (50, 51). There is some suggestion that GIP plays a role in invertebrate and early vertebrates distinct from its mammalian incretin role (52, 53). Whether this function is the same as or different from the function of mammalian pancreatic GIP remains to be elucidated.

Incretin-based Therapies for Diabetes

There are numerous structural and functional similarities between GIP and GLP-1, the two incretin hormones identified to date. Like GIP, GLP-1 belongs to the glucagon/secretin superfamily and has a precursor encoded by six exons and five introns (54). Processing of the GLP-1 precursor also occurs through PC enzymes and GLP-1 is also degraded by DPPIV (54). The nutrient stimuli and secretion profile of GLP-1 closely matches that of GIP, and both hormones exert their effects on target organs via G-protein coupled receptors that share downstream pathways (18, 54). However, the functional similarities appear to diverge in the pathological state. In human type 2 diabetic subjects, GLP-1 secretion is impaired (55, 56); in contrast, GIP secretion is left intact but the insulinotropic effect of GIP is blunted (57). This

critical difference in function in the pathological state has resulted in incretin therapies that heavily emphasize the therapeutic potential of GLP-1.

A Focus on GLP-1

GLP-1-based therapies can be grouped based on how they increase the level of active circulating GLP-1. DPPIV inhibitors increase the half-life of existing active GLP-1 molecules, exogenous administration of DPPIV-resistant GLP-1 analogues increases the total active GLP-1 level in the circulation, and GLP-1 receptor agonists interact directly with the receptor to increase its activity. All three classes aim to increase the overall incretin effect mediated by GLP-1, and there is a common consensus that the benefits of incretin mimetic therapies goes beyond maintaining adequate glucose control (58, 59). GLP-1's effects on enhancing β-cells regeneration and promoting β-cell survival have been documented in numerous studies (60, 61, 62). Functionally, GLP-1 also increases the glucose competence of β-cells (60, 63). Since the insufficiency of insulin secretion that characterizes type 2 diabetes is due to a progressive loss of β-cell mass and function (64), these effects of GLP-1 on the β-cells have been shown to reverse diabetes induced by the β-cell toxin, streptozotocin (62), and may slow the progression of the disease (60). Additionally, GLP-1 agonists have been shown to decrease food intake and improve cardiovascular functions (58), which may reduce obesity-related stresses on β-cell function and decrease mortality and morbidity due to cardiovascular complications associated with type 2 diabetes.

A Role for GIP?

Like GLP-1, GIP also has beneficial effects independent of its insulinotropic effects. These properties include the ability to stimulate insulin biosynthesis (65, 66); stimulate GLP-1 synthesis and secretion (67, 68, 69); and promote β-cell survival (70, 71, 72), although it has

been suggested that the relative contribution of the direct activation of the GIP receptor to these protective effects are less than the direct activation of the GLP-1 receptor (73). The presence of a blunted GIP incretin effect in type 2 diabetes (57) does not necessarily mean that these non-insulinotropic effects of GIP are also reduced. Indeed, since the degradation of GIP also occurs by dipeptide cleavage by DPPIV, DPPIV inhibitors also increase the half-life of GIP, such that the benefits of DPPIV inhibition do not only act through GLP-1 alone. Furthermore, the defect that causes the reduction of the GIP incretin effect may be reversible, as suggested by the ability of DPPIV-resistant GIP analogues to augment insulin secretion (74).

GIP can also be used as an early diagnostic tool. Type 2 diabetes has a strong genetic component, and a large proportion of normoglycaemic first-degree relatives of type 2 diabetics have a reduced insulinotropic effect of GIP (75), raising the possibility that a diminished GIP incretin effect may be a marker of diabetes potential. Should this be the case, lifestyle interventions may be sufficient to curb the development and progression of diabetes mellitus before more aggressive treatments are required. Further investigation into whether the defect leading to the loss of GIP's insulinotropic effect is a cause or effect of type 2 diabetes would provide a better picture of the therapeutic potential of GIP, as well as possibly, the contribution of this defect to the pathogenesis of type 2 diabetes.

GIP Antagonism

Unlike GLP-1, GIP appears to promote fat accumulation and obesity (74, 76). This property may account for the observation that DPPIV inhibition, which increases both GLP-1 and GIP levels as opposed to DPPIV-resistant GLP-1 mimetics, which only increases GLP-1 levels, does not induce weight loss (77). Several lines of evidence point towards a direct effect of GIP on fat accumulation. GIP expression and secretion is stimulated by fat intake (18, 21),

GIP receptors are found to be expressed in adipocytes (78), and disruption of GIP signalling either by genetic means (79, 80) or by use of GIP antagonists (81, 82, 83) have been shown to protect against the development of obesity. Although it has been suggested that GIP antagonism may have therapeutic potential for reducing metabolic disorders associated with obesity, including type 2 diabetes, the current evidence suggesting a role for GIP in protecting against diet-induced obesity relies only on the removal of GIP signalling in rodents fed on a high-fat diet. The lack of gain-of-function evidence suggesting a direct role for GIP in promoting obesity suggests that the role of GIP in human obesity requires further investigation (18).

A Role for GIP in Insulin Therapy

Central to the aetiology of diabetes mellitus is the lack of or insufficiency of insulin secretion. Insulin is a hormone released from the pancreas following a meal and is required to facilitate the uptake of glucose into the tissues for metabolism. Type 1 diabetes, representing 2-3% of all diabetic cases worldwide (2), is characterized by the autoimmune destruction of the insulin-producing β-cells (84), while type 2 diabetes, is characterized by the progressive loss of β-cell mass and function (64) which eventually leads to insufficient insulin secretion. Insufficient or absence of insulin signalling leads to accumulation of glucose in the blood and results in constantly elevated blood glucose levels. The high concentration of plasma glucose results in an accumulation of glycosylated haemoglobin (85). Since the percentage of glycosylated haemoglobin reaches steady state over time, while absolute blood glucose levels vary with feeding, the percentage of a form of haemoglobin called A1c which is glycosylated ($\%Hb_{A1c}$) is the standard measure of long term glucose control. Rigorous regimens of multiple daily insulin injections are required to achieve the $\%Hb_{A1c}$ target of treatment of < 7%, although this value tends to be higher than the average 4-6 $\%Hb_{A1c}$ in non-diabetic subjects (85).

Current Approaches and Limitations of Insulin Replacement Therapy

The most common and widely used therapy for maintaining blood glucose control for diabetic patients involves a combination of lifestyle modifications and medication. The battery of medications include DPPIV inhibitors and incretin mimetics, whose effects are already discussed, as well as through a variety of other mechanisms, such as lowering glucogenesis in the liver, activating peroxisome proliferators-activated receptor γ, preventing carbohydrate metabolism, and sensitising insulin release (77). In severe cases of type 2 diabetes, in which the insulin-secreting ability of β-cells have been exhausted, or in type 1 diabetes, in which β-cells have been destroyed, insulin replacement therapy is required for effective glucose management.

The discovery of insulin for the use of treating diabetes saved diabetic patients from what the ancient Greek physician Aretaeus aptly described as a "short, disgusting, and painful life as the bone and flesh melted into urine" (1). With the benefit of insulin therapy, patient survival rates drastically increased. Leonard Thompson, the first diabetic to be treated with insulin injections, who was diagnosed at the age of 14, lived to the age of 27, when he died of pneumonia (a cause unrelated to his diabetes). Another prominent patient was Elizabeth Hughes, daughter of U.S. Secretary of State Charles Evans Hughes, who was diagnosed at 13 and lived to the age of 73 with the insulin therapy (86). However, insulin replacement therapy has its limitations. Since insulin injections are administered based on blood glucose monitoring at discrete time points, even the best timing and dosage regimes are unable to mimic the finely tuned physiologic release of insulin following meals. The resulting blood glucose fluctuations present both long term and short term health risks for the patient. In the long term, hyperglycaemic episodes damage microvasculatures, placing the diabetic patient at increased risk of cardiovascular and renal disease, retinopathies, neuropathies, and circulation problems (85). An even more urgent consequence of inadequate blood glucose control is the

hypoglycaemic episodes that result when insulin is administered in too large of a dose or too soon before an expected glucose load. A sudden depletion of glucose in the blood can have fatal consequences. In the last 30 years, rapid-acting and long-acting basal insulin analogues, continuous subcutaneous insulin infusion pumps, and continuous glucose monitoring systems, have been developed for improved precision and lifestyle flexibility. For the most part, studies show that insulin analogues and continuous insulin infusion pumps are capable of delivering tighter glucose control as measured by Hb_{A1c} glycosylation and the number of hypoglycaemic episodes (87, 88).

As an alternative to exogenous insulin administration, whole pancreas or islet transplantation is emerging as a promising treatment for diabetes. In the field of transplantation science, transplantation of the pancreas is relatively new. Since diabetes mellitus is responsible for approximately 40% of end-stage renal failure, the procedure was initially performed in 1966 as a joint pancreas-kidney transplant in a type 1 diabetic patient with end-stage renal failure (89). This initial case demonstrated that normoglycaemia can be achieved in a type 1 diabetic patient without the need for exogenous insulin by transplantation of cadaveric pancreata. At around the same time, Paul Lacy began to investigate the feasibility of transplanting pancreatic islets, demonstrating in rodents that transplantation of the endocrine cells of the pancreas alone was sufficient to restore normoglycaemia in diabetic rats (90, 91, 92). The procedure was less invasive than that of a whole organ transplant and the first islet transplantation to a human diabetic patient in 1990 (93) marked the beginning of a promising treatment alternative to exogenous insulin administration. Further improvements to the islet isolation procedure led to the development of the Edmonton protocol in 2000, which currently serves as a basis for islet transplant protocol for transplant centers worldwide (94). Of the patients who achieved insulin-independence following transplantation in the original Edmonton cohort, only 10% remained so

after five years, but even in patients who have reverted to insulin, lower doses of insulin were required to maintain normoglycaemia and glucose fluctuations occurred in much narrower range following transplantation (95). Indeed, islet transplantation shows great promise as a treatment option for patients who experience frequent hypoglycaemic episodes with conventional insulin replacement therapy, despite the requirement for life-long immunosuppression (96). However, islet transplantation is not available to all those who are eligible because the demand for islets still exceeds supply. Islets from multiple donors for each transplant and multiple transplants are often required in order for a single patient to achieve insulin independence, since transplanted β-cell mass may decrease over time due to poor β-cell survival resulting from poor revascularization, allograft rejection, or in the case of a type 1 diabetic, recurrence of autoimmune destruction of the replacement β-cells may occur (95, 96). The lack of cadaveric islet donors has served as the impetus for efforts to increase graft survival, preserve islet function, and increase the supply of insulin-producing cells.

The K-cell as a Target for Insulin Gene Therapy

The unaffected secretion of GIP in the diabetic state indicates that the transcriptional machinery in GIP-expressing cells is functional in the diabetic state, making these cells an attractive target for gene therapy. The large imbalance between supply of cadaveric donors and demand for islets for transplantation have spurred a field of research in genetically engineering different types of somatic cells into insulin-producing cells (97, 98, 99, 100, 101, 102, 103, 104, 105). One limitation of such approaches is the lack of regulation of insulin secretion from these β-cell surrogates. However, since GIP-expressing cells share a similar secretion profile with insulin-secreting β-cells, targeting these cells should result in glucose-dependent insulin secretion. Indeed, transgenic mice expressing human insulin under the control of a 2.5 kb fragment of the rat GIP promoter were able to maintain normal glucose excursion following

treatment with the β-cell toxin streptozotocin, which rendered control mice diabetic (40). Since the autoimmune destruction of β-cells causing type 1 diabetes may potentially be triggered by antigens present in the β-cell itself, a regulated, endogenous source of insulin from a non-β-cell source may possibly provide the added benefit of circumventing recurrence of autoimmune destruction following transplantation. Utilizing the GIP-promoter to direct insulin expression in insulin gene therapy relies heavily on a thorough understanding of the factors regulating GIP expression at the level of the promoter.

Transcriptional Regulation in the K-cell

An Overview of Transcription

Transcription, the process of converting DNA to mRNA, is a key step in gene regulation, and is itself regulated by the primary, secondary, and tertiary structure of DNA. At the primary or genome sequence level, transcription is regulated by promoters, the DNA regions that flank the gene to be transcribed. Usually, DNA sequences that play a role in controlling transcriptional activity are located upstream of the gene, but there is also evidence for promoter regions located in introns and downstream of the gene (106, 107). The location at which the DNA nucleotides are used as a template for the synthesis of mRNA is called the transcription start site, and the location of promoter regions are designated as the number of nucleotides upstream (-) or downstream (+) of this site. The process of transcription consists of transcription initiation, elongation, and termination. The canonical understanding of transcription regulation is that initiation, the rate-limiting step, is the process upon which transcriptional activity is regulated, but mechanisms of transcriptional regulation acting to control the elongation and termination process are beginning to be appreciated (108). Regulation of transcriptional initiation depends on the concerted actions of factors located on the same molecule of DNA as

the transcribed gene, or *cis*-regulatory elements, such as binding sites for transcription factors; as well as factors external to the molecule, or *trans*-regulatory elements, such as transcription factors. Furthermore, *cis*-regulatory elements may be enhancer elements, which acts to increase promoter activity, or repressor elements, which decreases promoter activity.

The core promoter is the minimal segment of flanking sequences around a gene that is required for initiation to occur. Typically, it is found in the region between -35 bp and +35 bp (107) and consists of binding sites for general transcription factors, including the presence of CCAAT boxes (consensus sequence: GGNCAATCT), which are required to form the pre-initiation complex. In the case of eukaryotic transcription of protein-coding genes, the enzymatic unit of the initiation complex is RNA polymerase II, which binds at the TATA box (consensus sequence: TATA(A/T)AA(G/A)), a promoter element that is found on many eukaryotic promoters (107) required for RNA polymerase II recruitment. Even though the same components of the pre-initiation complex are required between cell types and between species, the core promoter region has been shown to display structural and functional diversity (107). For example, the well-characterized and well-conserved proximal sequence element in the core promoter of small nuclear RNA genes have a species-specific consensus sequence and within the same organism, facilitates binding of RNA polymerase II at certain gene promoters and binding of RNA polymerase III at other promoters (107).

In addition to the core promoter, gene-specific regulatory sequences are found in the proximal promoter, which is typically found within 200 bp upstream of the transcription start site. These sequences contain binding sites for specific transcription factors, which modulate the rate of initiation of complex formation upon binding (106). Generally, additional regulation occurs through the sequence upstream of the proximal promoter (106). The extent of the distal promoter is likely gene-specific, but interaction of response elements have been reported with

separation of as many as 5 to 200 kb (106, 109). The importance of the distal regulatory elements is particularly important for the regulation of transcriptional activity of genes whose promoters are found within CpG islands, regions of the genome ranging from 500 to 2000 bp with a GC content of more than 50% (107). These promoters make up approximately 50% of all mammalian promoters and lack the canonical TATA box, therefore recruitment of the general transcription factors and RNA polymerase is hypothesized to result from recruitment by enhancer proteins bound to distal promoter elements (107).

Transcriptional Regulation in the Development of the Enteroendocrine Cell Lineage

Transcriptional regulation plays a major role in determining intestinal cell fate during development. The intestinal epithelium is structurally organized into villi and crypts of Lieberkühn. Located in the crypts are intestinal stem cells which are the common progenitor of the four major cell types of the intestinal epithelium (35, 110). All the cells derived from a single crypt are derived from a single stem cell, so that each crypt represents a monoclonal population of cells (35). Each villus is surrounded by six to ten crypts and the cell population in a given villus represents a polyclonal population of the cells generated by the crypts that surround it (35). As cells differentiate, they migrate from the base of the crypt to the tip of the villus and are extruded into the lumen (110, 111, 112, 113, 114). This process lasts 3- 5 days and corresponds to the life span of intestinal epithelial cells (110, 111, 112, 113, 114). During this migration from crypt to villus tip, the expression of several transcription factors determine if a particular cell will develop into an enteroendocrine cells, and which type of enteroendocrine cell it develops into. Many of these factors also play a role in determining endocrine cell identity in the pancreas, once again reflecting the closeness of the evolutionary and embryologic origins between the intestine and pancreas.

The commitment to the secretory cell lineage, which gives rise to the mucous-secreting goblet cells, the lysozyme-secreting Paneth cells, and the endocrine-secreting enteroendocrine cells as opposed to the absorptive enterocyte cell lineage is largely determined by basic-helix-loop-helix (bHLH) transcription factor Math1 (115). Math1-deficient transgenic mice heterozygous for an allele where the Math1 coding region was replaced by β-galactosidase (βgal), were shown to have no goblet, Paneth, nor enteroendocrine cells, while the numbers of enterocytes in these mutant mice were comparable to wild type mice (116). Further commitment to an enteroendocrine cell fate within the secretory cell population requires the expression of another bHLH transcription factor, Ngn3 (117). Ngn3-deficient mice generated by homologous recombination using a construct consisting of the Ngn3 promoter driving *LacZ* (118) fail to develop endocrine cells and endocrine progenitors both in the pancreas (118) and intestine (117, 119).

As bHLH transcription factors, both Math1 and Ngn3 are repressed by a third bHLH transcription factor, Hes1, which binds the N-box regulatory element (CACNAG) to antagonize the activity of bHLH transcription factors bound to the E-box regulatory element (CANNTG) (120). Analysis of embryos of Hes-1-null mice generated by gene replacement of the bHLH domain with a neomycin-resistance cassette at E14 (120) demonstrated 3 to 6.5-fold increase in the number of enteroendocrine cells immunoreactive for cholecystokinin, somatostatin, proglucagon, and GIP compared to wild type mice (121). *Hes1* is upregulated by Notch signalling, hence, low Notch signalling favours enteroendocrine cell development (115). Other transcription factor known to be important for pancreatic endocrine cell development such as bHLH transcription factors NeuroD/BETA2 (122, 123), paired box transcription factors Pax4 (124), and Pax6 (124), homeobox transcription factors Isl-1 (125) and Pdx1 (126, 127) , have also been suggested to be critical for the determination of certain enteroendocrine cell lineages.

Transcriptional Regulation of GIP Expression

Of the transcription factors that direct endocrine cell fate, Pax6 and Pdx1 have been shown to be particularly important in GIP expression. Pax6 is necessary for proglucagon expression in the pancreatic glucagon-secreting α-cell (124, 128, 129) as well as in the enteroendocrine GLP-1 secreting L-cell (130, 131, 132, 133). In Pax6-deficient mice generated by homologous recombination using a construct consisting of the Pax6 promoter driving expression of *LacZ* (124) , animals die in the prenatal period and analyses performed on fixed embryonic tissue show both pancreatic and intestinal abnormalities. Pax6-deficient mice present a phenotype characterized by a lack of glucagon-producing α-cells in the pancreas and an approximately 90% reduction in the number of GIP-expressing cells (124). Whether this reduction in GIP-expressing cells reflects the reduction of the subpopulation of GIP-expressing cells that simultaneously express GLP-1 is not known, although there is evidence that the subpopulation of GLP-1-expressing cells that express Pdx1 in addition to Pax6 simultaneously express GIP (132).

Pdx-1 null mice generated by targeted mutagenesis by replacement of the second exon of Pdx-1 with a neomycin resistance cassette completely fail to develop a pancreas (134) and display a 98% reduction of GIP-expressing cells (135) in an anatomically intact intestinal tract with sparser and shorter villi (127, 136). Interestingly, rescue of the Pdx1-null phenotype using transgenic animals with a transgene encoding Pdx1 driven by a segment of the Pdx1 promoter region demonstrated recovery of normal pancreas morphology and insulin expression comparable to wild type animals but only weak recovery of pdx1 expression in intensity and cell number in the duodenum (136). There were 50% and 90% fewer GIP-expressing cells in the duodenum of a Pdx1null animal possessing 5 and 3 copies of the rescue transgene, respectively, compared to the number of GIP-expressing cells in the duodenum of a wild type animal (136).

The Pdx1 promoter segment in the rescue transgene lacked the most distal of four highly conserved regulatory elements (137, 138, 139), suggesting that this distal element may be important for directing expression of Pdx1 to intestine (136). The importance of Pdx1 in the determination of a GIP-expressing cell fate is also demonstrated by the colocalization of Pdx1 with GIP in the developing intestine. Pdx1 is required for the normal structural organization of the duodenum and is highly expressed in epithelial cells in the rostral duodenum at E16.5 (135). Following this critical period, expression of Pdx1 is downregulated in cells that will not develop into endocrine cells (135). Analysis of mouse duodenal tissue harvested at E18.5 and 10 weeks of age demonstrated Pdx1/GIP colocalization suggesting that Pdx1 is necessary for GIP expression in that there was no instances where a GIP-immunoreactive cell was not immunoreactive for Pdx1 (135). Furthermore, an *in vitro* report showed that overexpression of Pdx1 in the undifferentiated rat intestinal IEC-6 cell line is sufficient to coax cells to become GIP-producing cells (140) and overexpression of Pdx1 in the plurihormonal mouse intestinal STC-1 cell line, which already expresses GIP, further increases GIP promoter activity (135).

A recent study attempted to segregate the effects of Pdx1 on pancreatic and intestinal development by characterizing transgenic mice with a targeted deletion of Pdx1 in the intestine using a *Cre/lox* system (141). The authors reported that while the number of GIP immunoreactive cells from targeted Pdx1 knockout animals did not differ from wild type animals, the level of GIP mRNA was decreased (141). However, this conclusion that there was no difference in cell number was based on the lack of statistical significance. The actual percent decrease of GIP-immunoreactive cell count in the Pdx1 knockout animals compared to controls was 32%, a decrease that would sufficiently explain the decrease in GIP mRNA levels, which were detected by real-time PCR using total RNA isolated from whole intestinal segments.

Transcriptional elements at which transcription factors bind to activate GIP expression have been identified through a series of *in vitro* studies using the human and rat GIP promoters. Cloning of the 5'-sequence upstream of the human GIP gene by Inagaki *et al.* (19) resulted in the identification of several potential regulatory elements based on consensus sequences. These include typical features of core promoters, such as a TATA motif (-28 bp to -23 bp), a CCAAT box (-152 bp to -156 bp), an enhancer core sequence (-137 bp to -130 bp), and a binding site for the transcriptional enhancer Sp1 (-383 bp to -378 bp) (19). Further upstream of these elements, putative binding sites at which protein kinases A and C and cAMP may exert a regulatory effect were identified, namely an AP-1 site (-344 bp to -338 bp), an AP-2 site (-368 bp to -360 bp), and two cAMP response elements (CREs) (-349 bp to -342 bp and -376 bp to -368 bp) (19). No attempts were made to test the functionality of these putative binding sites in this study.

A functional study of regulatory elements in the human GIP promoter was performed a few years later by Someya *et al.* (142) using the hamster insulinoma cell line HIT T15 as a model. Various lengths of the human GIP promoter were used to drive the expression of a reporter gene expressing chloramphenicol acetyltransferase (142). Shortening of a reporter construct containing -1200 bp of the human GIP promoter to -386 bp and -334 bp resulted in a decrease in promoter activity, but further truncation of the promoter to -258 bp and -180 bp resulted in augmentation of promoter activity compared to the basal level established by the longest reporter construct containing -1200 bp of the human GIP promoter (142). Further truncation to -134 bp resulted in a complete loss of promoter activity (142). These observations led the authors to conclude that the basal promoter for the human GIP gene spans -180 bp to +14 bp (142). An additional experimental condition in which the activity of each of the reporter constructs was measured following treatment with the cell-permeable cAMP analogue dibutyryl cAMP demonstrated that cAMP-induced promoter activity in all the constructs except for the

shortest construct driven by -134 bp of the human GIP promoter, which led to the speculation that a CRE may be located between -180 bp and -134 bp (142). Indeed, two such elements were identified by DNase I footprinting spanning -350 bp to -339 bp and -164 bp to -149 bp (142). Note that the first corresponds to the approximate location of one of the putative CREs identified by Inagaki *et al.* (19), while the second corresponds to the region containing the CCAAT box in the Inagaki *et al.* study (19). Someya and colleagues (142) note that while the first CRE displays a high homology with CRE consensus sequence (TGACGTCA), the second CRE contain only a portion which appear similar to the first half of the CRE consensus sequence (142). Despite this, mutation of the second CRE generated a larger decrease in promoter activity (88%) than mutation of the first CRE (50%) and a simultaneous mutation of both CRE sites resulted in a further decrease in promoter activity to 85% of the wild type promoter (142).

A second functional study of the human GIP promoter was performed by Fujita *et al.* (132) in the plurihormonal mouse enteroendocrine STC-1 cell line provided further evidence that -184 bp of the human GIP promoter was required for basal activity. Shortening a -2.9 kb human GIP promoter to -1.9 kb, -1.0 kb, and -210 bp resulted in a progressive increase in promoter activity peaking at the construct containing -210 bp of the human GIP promoter driving the luciferase reporter (132). Truncation to -184 bp yielded a promoter activity level comparable to the longest -2.9 kb construct and further shortening of the promoter to -145 bp and -93 bp decreased promoter activity by 90% relative to the longest -2.9 kb construct (132). Binding of Pax6, Pdx1, and GATA4 to the human GIP promoter was demonstrated using electrophoretic mobility shift assays (EMSAs) using a probe that spanned -193 bp to -138 bp of the human GIP promoter and mutations to five putative regulatory elements within this region using the -210 bp construct as a template resulted in decreases in GIP promoter activity ranging from 20% to 90% compared to wild type promoter activity (132). Overexpression of Pax6 and Pdx1 in the

undifferentiated rat intestinal IEC-6 cell line demonstrated dose-dependent induction of human GIP promoter activity and conversely, expression of dominant-negative mutants of Pax6 and Pdx1, which contains an intact DNA-binding domain but lacks a transactivation domain, reduced human GIP promoter activity in a dose-dependent manner (132). Furthermore, overexpression of Pdx1 and Pax6 simultaneously produced an augmentation of GIP promoter activity that exceeded overexpression of one or the other factor alone (132).

Cloning of the 5'-sequence upstream of the rat GIP gene by Higashimoto *et al.* (20) did for the rat GIP promoter what Inagaki *et al.* (19) did for the human GIP promoter, namely, generate a list of putative *cis*-regulatory elements based on consensus sites for various elements. These putative sites include: TATA boxes at -115 bp, -27 bp, and +755 bp; CCAAT boxes at -171 bp, -158 bp, and +599 bp; AP-1 binding sites (consensus sequence: TGA(G/C)TCA) at -842 bp, -802 bp, -411 bp, -53 bp, and +624 bp; AP-2 binding sites (consensus sequence: ((T/C)C(C/G)CC(A/C)NC(G/C)(C/G)(G/C)) at -813 bp, -751 bp, +153 bp, +688 bp, and +745 bp; and CREs (consensus sequence: TGACGTCA) at -379 bp, +35 bp, and +226 bp (20). Noting that a number of these sites are located downstream of the transcriptional start site, transcriptional regulation by intron 1 was hypothesized, and data demonstrating a possible role for intron 1 in promoter activity was presented based on RNase protection experiments (20).

Characterization of the rat GIP promoter has largely been done using the plurihormonal mouse enteroendocrine STC-1 cell line as a model (135, 143, 144, 145). Boylan *et al.* (143) demonstrated using luciferase reporter plasmids driven by -2.5 kb to -173 bp of the rat GIP promoter that -193 bp of the rat GIP promoter was the minimum length required to achieve 200-fold activity from the promoterless luciferase backbone (143). A further reduction of the promoter length from the distal end to -182 bp resulted in 30-fold activity and reduction to -173 bp resulted in 11-fold activity from background (143). The requirement for minimal

activity and the identification of a TATA box (-27 bp to -24 bp), enhancer core element (-133 bp to -136 bp) and two CCAAT boxes (-158 bp to -154 bp and -171 bp to -167 bp) led to the definition of the region between -173 bp and +19 bp as the core promoter (143). While intron 1 appears to direct minimal promoter activity, as seen with a construct containing +1 bp to +111 bp of the rat GIP gene driving the luciferase gene (143), no differences between promoter activity was observed when a luciferase reporter series driven by GIP promoter fragments ending with +19 bp at the 3' end of the insert and an identical series ending with +111 bp to include intron one were observed (143). Furthermore, a deletion of the region between -177 bp and +111 bp in two of the intron-containing reporter constructs reduced promoter activity to that seen with the +1 bp to +111 bp fragment alone, suggesting that there is no effect on proximal promoter activity mediated by the intronic sequence (143). Note that this region does not contain the intronic TATA box proposed by Higashimoto *et al.* (20) to be important in regulating GIP gene transcription. Two putative GATA-binding sites (-191 bp to -186 bp and -173 bp to -178 bp) were located within the proximal promoter spanning -193 bp to -182 bp based on the GATA-binding consensus sequence (A/T)GATA(A/G) (143). Putative AP-1 (-413 to -405) and CRE sites (-380 to -373) were identified upstream of this proximal promoter (143), which occupy similar positions to the corresponding sites identified on the human GIP promoter in the Inagaki *et al.* study (19).

Functional characterization of the two GATA binding elements was carried out by a single nucleotide change from G to A and from T to A, such that the core GATA consensus sequence was disrupted into AAAA (143). Utilizing the luciferase reporter construct driven by -193 bp to +19 bp of the rat GIP promoter as a template, disruption of the distal (-191 bp to -186 bp) GATA site in this manner resulted in a 90% decrease in activity from the wild type, compared to 35% when the proximal (-173 bp to -178 bp) site was mutated (143). Mutation of

both sites simultaneously resulted in a 90% decrease in activity, equivalent to mutating the distal site only, demonstrating that it is functionally more important of the two sites (143). Six years later, a publication from the same group (144) identified GATA4 as the *trans*-regulatory factor binding to this site using EMSAs with a DNA probe spanning -193 bp to -168 bp. Five years after that initial finding that GATA4 binds the rat GIP promoter in a region that activates GIP gene expression, the role of GATA4 in GIP gene expression was further clarified by demonstration of GATA4 immunoreactivity in GIP-expressing K-cells as well as comparison studies of GIP promoter activity between the STC-1 cells and mouse pancreatic β-cell tumour cells, βTC3 (145). The GIP promoter in βTC3 cells, which unlike STC-1 cells, do not express GATA4 as shown by reverse-transcriptase PCR, Northern blot, and Western blot, was approximately 13-fold less active than in STC-1 cells (145). Overexpression of GATA4 in a βTC3 cell culture resulted in increased endogenous GIP mRNA levels as measured by Northern blot, which corresponded to increased GIP secretion as measured by a functional GIP bioassay (145). Conversely, suppression of GATA4 expression by co-transfection of a GIP promoter-driven luciferase reporter with a short-hairpin RNA construct specific for GATA4 resulted in attenuation of GIP promoter activity by 70-80% (145).

In addition to GATA4, an Isl1 motif (CATTAG) identified between -156 bp to -151 bp was revealed to account for 85% of rat GIP promoter activity based on mutation of this site to GAAAAG using the -193 bp to +19 bp rat GIP promoter as a template (144). In another set of EMSAs with a probe spanning -160 bp to -141 bp of the rat GIP gene, Isl1 but not GATA4 was identified as the *trans*-regulatory element acting at this *cis*-element (144). Interestingly, the consensus binding site for Isl1 is the same for Pdx1. Using the same EMSA probe for demonstrating Isl1 binding to the CATTAG site, Pdx1 was also shown to participate in the transcription factor complex acting on the CATTAG element (135). Binding of Pdx1 to the

endogenous GIP gene in STC-1 cells was demonstrated by immunoprecipitation of chromatin isolated from these cells (144). The manuscript text suggests that the same was shown for Isl1, but these data were not presented in the paper.

As seen from this brief literature review, functional promoter studies in the human and rat GIP promoter have focused largely on the proximal promoter extending as far as -200 bp of the promoter. However, reporter genes driven by the GIP promoter indicate that elements even more distal to this region have a role in regulating GIP expression. Notably, transgenic mice expressed thymidine kinase driven by 1.2 kb of the human GIP promoter in the stomach, and islets, but not in the intestine (146). In a different line, a 2.5 kb fragment of the rat GIP promoter directed insulin expression in the stomach and intestine, but not the islets (40), while in a third line, a longer piece of the rat promoter consisting of 3.1 kb of the 5' upstream promoter and a part of intron 1 of the rat GIP gene (80) drove expression of dsRed2 in the intestine but not in the islets. Whether the restriction of GIP expression from intestine and stomach to intestine only is attributed to regions of the 3.1 kb rat promoter that were absent in the 2.5 kb promoter or regions in intron 1 remains to be determined.

Model Systems for Studying K-cell Biology

An overview of the literature concerning GIP gene regulation shows that much of what we know about K-cells arises from the use of mutant animals and cell lines. The diffuse location and rarity of K-cells in the intestinal epithelium and the lack of identifying markers which could be used to isolate a live K-cell population make studying GIP gene transcription as well as K-cell biology in general challenging. While some insight into the contribution of transcription factors to regulating GIP gene expression could be gained from transgenic animals in which genes encoding for transcription factors have been knocked out or knocked down, this approach is

limited, since disruption of the expression of a single transcription factor always affects more than one system by virtue of transcription factors acting in multiple tissue types and at different stages of development. Tissue-specific transgenic mice, such as those generated by the *Cre-lox* system, may better isolate the effects of a particular transcription factor to a certain tissue, but the temporal removal of the gene of interest depends on the promoter element driving *Cre* recombinase, which may or may not recapitulate the temporal expression of the transcription factor in question. In the end, the mechanisms by which various transcription factors regulate GIP gene expression cannot be delineated by mutant mouse studies.

In an attempt to investigate the mechanisms by which the GIP gene is transcriptionally regulated, tumour cell lines have been widely-used convenient models for studying the sites and interactions of transcription regulatory elements (135, 142, 143, 144, 145). Before embarking on identifying regulatory elements on the rat GIP promoter, Boylan *et al.* found that of a number of cell lines, including two cell lines derived from the salivary gland and six endocrine lines, only the human embryonic intestinal 407 line and the plurihormonal mouse intestinal STC-1 line expressed rat GIP (143). One of the cell lines which did not express GIP was the Syrian hamster β-cell insulinoma line HIT T15 line, the same model used by Someya *et al.* (142) to provide the first functional characterization of the human GIP promoter, indicating that differences in promoter activity of the same gene from different species may occur in the same cell line model. Also, despite human GIP promoter activity in a hamster β-cell-derived cell line (142) and rat GIP promoter activity in a mouse β-cell-derived cell line (145), GIP is not known to be localized in the mature β-cell, suggesting that *in vitro* promoter studies may not always accurately predict *in vivo* outcomes.

The STC-1 cell line is a mixed endocrine intestinal cell line derived from transgenic mice carrying the rat insulin promoter driving the SV40 large T antigen (147) and has been used as a

model for studying transcriptional regulation of incretins (135, 143, 144, 145, 148). The GIP-expressing cell population consists of a subpopulation that expresses GLP-1 in addition to GIP, similar to *in vivo* observations (132). However, other hormones including secretin, gastrin, pancreatic polypeptide, neurotensin, and somatostatin are also expressed in the cell line (147). Therefore, claims of "cell-specific expression" demonstrated in STC-1 cell lines generally mean specific to STC-1 cells as opposed to constitutive expression in cell lines derived from other tissues rather than specificity to the cell type that is modelled by STC-1 in that particular study (135, 143, 144). Furthermore, promoter requirements for gene expression *in vitro* may not necessarily be the same *in vivo*. When a -2.5 kb region of the of the gene encoding myogenic regulatory factor myoD shown to drive chloramphenicol acetyltransferase expression in the mouse myogenic cell line 23A2 was used to drive βgal expression in transgenic mice, none of the transgenic mouse embryos expressed βgal in myogenic cells despite successful integration of transgene indicated by ectopic βgal expression (149). Another example comes from a study of the vasoactive intestinal peptide gene promoter, in which a 5.2 kb promoter fragment was able to drive chloramphenicol acetyltransferase activity in the human SK-N-SH neuroblastoma cell line directed chloramphenicol acetyltransferase expression in the intestine but not the brain of transgenic mice (150, 151). In the latter example, the extrapolation of *in vitro* to *in vivo* data is further complicated by the difference in species origins between the *in vitro* and *in vivo* model systems.

To address these limitations of current model systems in studying K-cell biology, Parker and colleagues (42) attempted to generate a primary K-cell line from transgenic mice expressing a fluorescent marker. The strategy involved expressing the Venus yellow fluorescent protein under the control of a bacterial artificial chromosome construct containing the approximately 200 kb of the rat GIP promoter (42). Characterization of the resulting transgenic mouse tissue

indicated the expression of Venus was restricted to GIP-expressing cells, although some cells isolated based on fluorescence also expressed GLP-1 and polypeptide tyrosine tyrosine at 9- and 780-fold compared to non-GIP-expressing cells (42). This finding is consistent with prior observation that a subpopulation of K-cells expresses GLP-1 in addition to GIP (29, 30, 36), but no polypeptide YY immunoreactivity was seen to occur in GIP-immunoreactive cells in either foetal human pancreas (152) or in adult pig, rat, and human pancreata (29). Sufficient cell numbers were purified by this method to compare the expression levels of various candidate genes between the GIP-expressing and non-GIP-expressing cells by quantitative RT-PCR to allow for a better understanding of the glucose and nutrient sensing machinery present on the K-cell (42). Theoretically, if enough genes are compared between the GIP-expressing population and the non-GIP-expressing population, this approach may eventually facilitate the identification of a K-cell-specific surface marker, which would enable purification techniques for K-cells without the need of fluorescent markers. Although secretion experiments were performed in this study (42), cultures of mixed primary cells isolated from the adult mouse intestine were used as opposed to an enriched culture of K-cells. An enriched primary culture of GIP-expressing cells would be an ideal model for studying K-cell biology, as it would be a compromise between a system that is more physiologically relevant than tumour cell lines yet is suitable for the type of controlled manipulation required to study mechanisms.

THESIS INVESTIGATION

Through a candidate gene and unbiased screening approach, the **aim** of this thesis was to identify transcriptional regulatory elements and transcription factors critical for the transcriptional regulation of GIP in a segment of the GIP promoter distal to the region that has been previously characterized. The longest segment of the rat GIP promoter characterized to date using reporter assays extended from -2.5 kb to +19 bp as documented by Boylan *et al.* (143) and Jepeal *et al.* (135) and identification of *cis*-regulatory elements for the GIP gene have largely focused on the proximal promoter (-193 bp to +19 bp) region. Transgenic animal studies where reporter genes are driven by segments of the GIP promoter extending beyond the distal limits of the proximal promoter show that regulatory elements that contribute to restricting GIP localization to specific tissues may be located in the yet uncharacterized promoter upstream of -2.5 kb. These observations led to the **hypothesis** that transcription factors, including those known to be important in regulating GIP expression at the proximal promoter, are acting on elements in the distal promoter and contribute to regulate overall GIP promoter activity. Luciferase-reporter-based screening for promoter elements important for GIP expression was carried out using the STC-1 cell line as a model for K-cells and the *trans*-regulatory factors that bind to these elements were identified.

The rarity and diffuse location of GIP-expressing enteroendocrine cells across the intestinal mucosa have presented a significant challenge in obtaining data on transcriptional regulation in primary cells. The second objective of this thesis was to develop a fluorescence-based method for obtaining highly-enriched populations of primary GIP-expressing cells, which would allow for a high-throughput microarray-based comparison of genes expression in GIP-expressing cells versus non-GIP-expressing cells. Since transcription factors play a pivotal role in determining K-cell fate, it is expected that this comparison will yield a list of candidate

transcription factors based on their high expression level in GIP-expressing cells compared to the rest of the intestinal cell population. The factors can then be subjected to functional testing to assess their role as potential *trans*-regulatory elements on the GIP promoter. Related to this goal, attempts were made to maintain enriched primary GIP-expressing cells in culture to create a more physiologically relevant model for studying transcriptional regulation among other aspects of K-cell biology.

MATERIALS AND METHODS

Alignment of GIP Promoter Sequences from Multiple Vertebrate Species and Prediction of Putative Transcription Factor Binding Sites

Genomic DNA sequences immediately distal to the transcriptional start site of the GIP gene were obtained for nine species (*species, Ensembl ID number:* human, ENSG00000159224; chimpanzee, ENSPTRG00000009367; macaque, ENSMMUG00000021088; mouse, ENSMUSG00000014351; rat, ENSRNOG00000006306; cow, ENSBTAG00000005045; dog, ENSCAFG00000016899; elephant, ENSLAFG00000004536) from the Ensembl Genome Browser (http://www.ensembl.org). The sequence of 4.5 kb of the pig distal GIP promoter was obtained from enGene, Inc. Percent identity plots were generated using mVISTA (http://genome.lbl.gov/vista/) and nucleotide-level multiple alignments were obtained by inputting sequences into MultiPipMaker (153, 154) to determine the location of the distal well-conserved promoter elements in the GIP promoter sequences driving the expression of luciferase in our reporter constructs. In addition to literature review of known transcription factor consensus sequences, putative transcription factor binding sites were identified using ALGGEN-PROMO (155, 156). Default settings in ALGGEN-PROMO were used in putative binding sites using the SearchSites function.

Plasmids

Generation of GIP Promoter-Luciferase Truncation Plasmid Series for Investigating the Distal GIP Promoter

A series of constructs consisting of luciferase driven by varying lengths of the rat GIP promoter (pGL3795, pGL3075, pGL2574, pGL2033) was generated by Dr. Robert Baker

(Centre for Human Islet Transplant and Beta-Cell Regeneration, University of British Columbia, Vancouver, BC). pGL2574 was first generated by linearizing the commercial vector pGL4.10 [luc2] (Promega, Madison, WI) with HindIII (New England Biolabs, Ipswich, MA; all subsequent restriction enzymes are from this source, unless otherwise indicated) and simultaneously treated with alkaline phosphatase to prevent autoligation. A 2.6 kb rat GIP promoter fragment generously provided by Dr. Michael Wolfe (Boston University School of Medicine and Boston Medical Centre, Boston) was excised from the pcDNA3.1(-2.6)rGIP-CreER plasmid with HindIII, and ligated into the linearized pGL4.10 [luc2] vector (Promega, Madison, WI). Vectors containing insert were identified by electrophoresis; insert orientation was determined by diagnostic NheI (Fermentas, Burlington, ON) and XhoI digests.

The longest GIP promoter-luciferase construct (pGL3795) was generated by cloning the distal rat GIP promoter (-3795 bp to -2515 bp) by PCR from rat genomic DNA using primers rGIP-3795 (5'- CCACACTCGAGCTCTCTCCCCAAAACCAAACAAGCCAGT-3') and rGIP-2515 (5'-GTGAGGTTTCTTGGGGTTTGAGGCTG-3') (Integrated DNA Technologies, Coralville, IA). The 5' primer introduces a XhoI site (underlined) at the promoter's distal end and the 3' primer anneals just downstream of an endogenous HindIII site. The PCR product and vector pGL4.10 [luc2] (Promega, Madison, WI) were both digested with XhoI and HindIII, and subsequently ligated together, producing the vector pGL4.10 rGIP (-3.7/-2.5). This vector was then re-opened with HindIII and ligated to the HindIII-excised 2.6 kb rat GIP promoter fragment from pcDNA3.1(-2.6)rGIP-CreER used to generate the pGL2574 plasmid (described above), effectively extending the -2.6 kb rat GIP promoter driving the luciferase gene to a -3.7 kb promoter. Correct insert orientation was determined by diagnostic NheI (Fermentas, Burlington, ON) and XhoI digests and the sequence of the distal promoter was confirmed by cloning the distal promoter into the pBlueScriptSK+ vector (Stratagene, La Jolla, CA) and sequencing with

T3 and T7 primers. As the 2.6 kb promoter fragment was previously sequenced, only the extended sequence between -3.7 kb and -2.6 kb was sequenced. The two remaining plasmids in the series, pGL3075 and pGL2033, were generated using the pGL3795 plasmid as a template. pGL3075 was generated by digestion of pGL3795 with KpnI to remove 700 bp from the distal end of the -3.7 kb promoter, and subsequent autoligation of the remaining 7.3 kb fragment Similarly, digestion of the pGL3795 plasmid with NheI (Fermentas, Burlington, ON) removed 1.7 kb from the distal end of the -3.7 kb promoter, resulting in a 6.2 kb fragment for autoligation to create pGL2033. DNA preparations of the above plasmids were made by transforming One Shot TOP10 Chemically Competent *E. coli* (Invitrogen, Carlsbad, CA) for 30 s at 42°C followed by DNA isolation from bacteria using the GeneJET Plasmid MiniPrep Kit (Fermentas, Burlington, ON). The sequence of the entire -3.7 kb of the rat GIP promoter, which was used as a template for creating these plasmids had been previously sequenced, and clones were digested with KpnI and XhoI to verify separation of the 3.1 kb and 2.0 kb promoter from the 4.2 kb pGL4.10 [luc2] backbone. After verification of correct ligation by restriction enzyme digests, larger quantities of DNA were isolated for transfection using the EndoFree Plasmid MaxiPrep Kit (QIAGEN, Valencia, CA).

A second truncation series with GIP promoter lengths spanning the region between the promoter segments in pGL3075 and pGL2574 was generated based on the locations of several restriction enzyme sites in the DNA sequence between -2574 bp and -3075 bp of the rat GIP promoter. Since some of the restriction enzymes used in the cloning strategy were sensitive to Dam and Dcm methylation, a DNA preparation of pGL3075 was made by transforming dam$^-$/dcm$^-$ competent *E. coli* (New England Biolabs, Ipswich, MA) for 30 s at 42°C, to use as a template. Limit digestion of the pGL3075 template with PvuII (Fermentas, Burlington, ON) for pGL2918 and with XbaI for pGL2894 resulted in a linearized pGL3075 fragment. Limit digests

were carried out using 10 μg DNA in a 100 μL reaction volume with 1 μL of the respectively restriction enzymes. The DNA was allowed to digest for 30 s at room temperature for PvuII and 5 min at 37°C. These digestion conditions were empirically determined and found to yield the most linearized product out of all the conditions tested. The digestion product was cleaned up and ran on a 0.7% agarose gel in 1 x TAE to separate the linearized product from the other digestion products.

The DNA cleanup and gel extraction protocols are described as follows: A five-fold volume relative to the digestion reaction volume of binding buffer PB (QIAGEN, Valencia, CA), was mixed thoroughly with the restriction digest reaction and allowed to bind to a DNA clean-up column from the GeneJET Plasmid MiniPrep Kit (Fermentas, Burlington, ON) by centrifugation at 3000 g for 20 s. The flow-through was allowed to pass through the column again to maximize binding. The column was then washed with 750 μL of wash buffer PE (QIAGEN, Valencia, CA). After discarding the flow-through, the column was dried by centrifuging at 13,000 g for 2 min, and the cleaned DNA was eluted with 50 μL of TE Buffer (QIAGEN, Valencia, CA). After separation, the gel was visualized using a UV-C transilluminator (Spectronics, Westbury, NY) and the thin piece of gel containing linear fragment was excised into a pre-weighed microcentrifuge tube. The mass of the excised gel piece was determined by subtraction and 3 μL/mg of gel of solubilization and binding buffer QG (QIAGEN, Valencia, CA) was added to the microcentrifuge tube. The excised gel piece was melted by incubation at 55°C for 15 min and DNA was extracted by passing the QG-gel solution through a DNA clean-up column from the GeneJET Plasmid MiniPrep Kit (Fermentas, Burlington, ON). The wash and elution steps are the same as described for DNA cleanup.

Subsequent digestion of the linearized products by PvuII (Fermentas, Burlington, ON) and XbaI digestion with KpnI removed 157 bp and 181 bp from the distal end of the -3.1 kb rat

GIP promoter driving the luciferase gene, respectively, to generate a linear fragment consisting of 2918 bp and 2894 bp of the rat GIP promoter followed by the luciferase gene. This linear fragment was incubated with the large Klenow fragment isolated from *E. coli* DNA polymerase I (New England Biolabs, Ipswich, MA) to produce compatible blunt ends, gel-purified as described above, and autoligated overnight at 12°C. After transformation into One Shot TOP10 Chemically Competent *E. coli* (Invitrogen, Carlsbad, CA) for 30 s at 42°C, colonies were screened using a PCR-based method. Colonies were touched to the surface of a PCR reaction mix (as described in the protocol for Accuprime *Taq* DNA Polymerase System, Invitrogen, Carlsbad, CA) containing forward primer RVprimer3 (5' CTAGCAAAATAGGCTGTCCCC 3'; Promega, Madison, WI) and reverse primer truncVAR (5' CTTTGGCAGTCCTGGGAGG 3'; Integrated DNA Technologies, Coralville, IA) before the preparation of a starter culture. The 519 bp PCR product from the uncut pGL3075 vector was easily distinguishable from the 341 bp and 317 bp products generated by the successful truncation at the PvuII (pGL2918) and XbaI (pGL2894) sites, respectively, making this method a simple high-resolution way to screen many colonies effectively.

pGL2646 was generated using a similar cloning strategy using the restriction enzyme StuI. Since StuI was a unique cutter for the pGL3075 template plasmid, no limit digestion was required. Double digestion of pGL3075 with KpnI and StuI resulted in the desired -2646 bp of the rat GIP promoter sequence followed by the luciferase gene, which was blunted, gel-purified, and autoligated, as described above. Clones were digested with NcoI/KpnI and SacII/NheI (NheI from Fermentas, Burlington, ON) to ensure that the KpnI restriction site was destroyed and to confirm that the -3075 bp GIP promoter was shortened by 450 bp. The sequences of all three plasmids generated by autoligation were confirmed by sequencing using RVprimer3 (Promega, Madison, WI).

Generation of GIP Promoter-Luciferase Plasmids Containing Mutated Putative Transcription Factor Binding Sites

Two constructs, each containing a mutation of two candidate distal binding sites for Pax6 and Pdx1, were cloned using a three-way ligation method of site-directed mutagenesis (Figure 1). Briefly, forward and reverse PCR primers (Pax6mutF/R, Pdx1mutF/R; Integrated DNA Technologies, Coralville, IA) containing the mutated candidate binding site were designed and used in combination with a reverse (SacI-R; Integrated DNA Technologies, Coralville, IA) and forward (KpnI-F; Integrated DNA Technologies, Coralville, IA) primer, respectively, in two separate PCR reactions. The reverse and forward primers, SacI-R and KpnI-F, contain the nearest unique restriction site downstream and upstream, of the mutation site, respectively, effectively generating two PCR products (as per protocol, Accuprime *Taq* DNA Polymerase System, Invitrogen, Carlsbad, CA) when the pGL3075 plasmid was used as a template (Figure 1). These products were digested to completion with restriction enzymes (ApaI to generate pGL3075Pax6mut and EcoRI to generate pGL3075Pdx1mut) and ligated to each other and a linearized vector generated by doubly-digesting the pGL3075 vector with KpnI and SacI to remove the fragment to be replaced by the PCR products (indicated in blue in Figure 1). Following bacterial transformation, colonies were screened by PCR using primers flanking the KpnI and SacI restriction sites (rGIPmutSeqFwd: GCTGTCCCCAGTGCAAGTGCAGGTGCC, rGIPmutSeqRev: GCTGTGATCCTGTCACCACGCCGTCCCA; Integrated DNA Technologies, Coralville, IA) and the presence of the mutated binding site was further demonstrated by restriction digest of the PCR products by ApaI or EcoRI. The insert was subsequently sequenced to ensure that the only base changes were those intended to occur at the mutated putative binding sites.

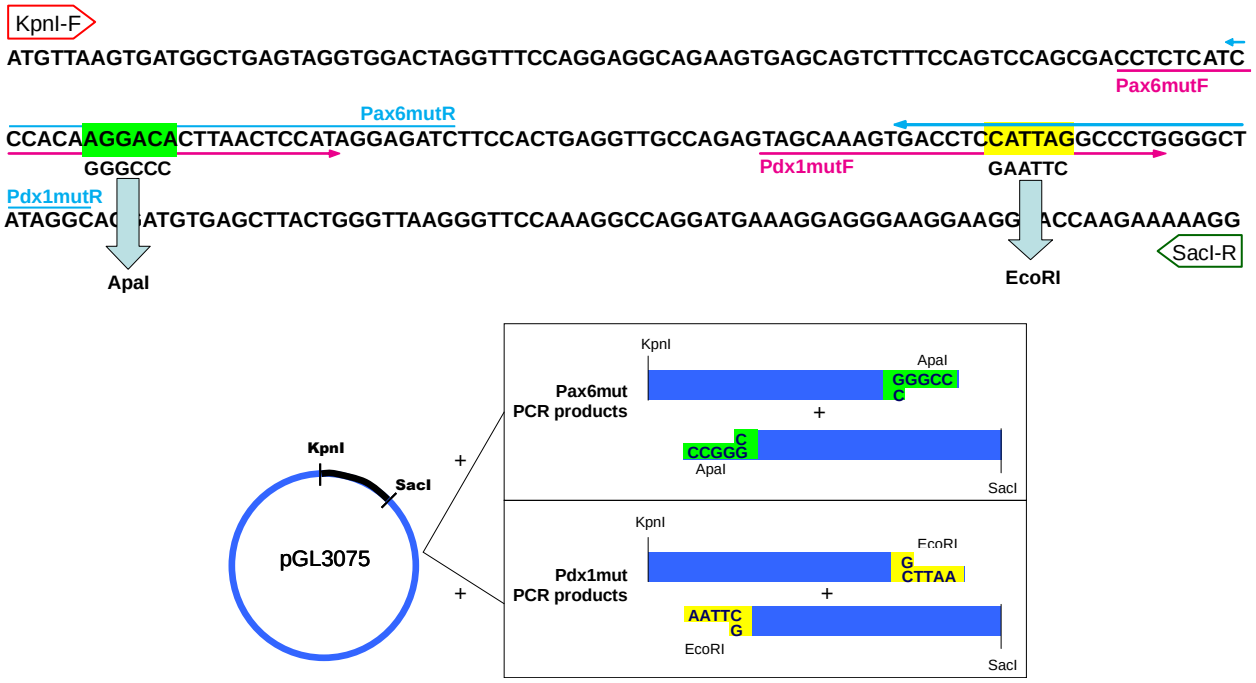


Figure 1. Schematic of site-directed mutagenesis by three-way ligation. A putative Pax6 binding site (AGGACA) was mutated to a ApaI restriction site to generate the pGL3075Pax6mut vector and a putative Pdx1 binding site (CATTAG) was mutated to a EcoRI restriction site to generate the pGL3075Pdx1mut vector. To generate pGL3075pax6mut, two PCR reactions were carried out with the pGL3075 vector as a PCR template, one using the primer set KpnI-F/Pax6mutR and one using the primer set Pax6mutF/SacI-R. These products were subsequently digested with ApaI and ligated to each other and a vector (blue circle) generated by doubly-digesting the pGL3075 vector with KpnI and SacI to remove the fragment to be replaced by the PCR products. To generate pGL3075pdx1mut, a similar approach was used, except the two PCR reactions were performed using KpnI-F/Pdx1mutR and Pdx1mutF/SacI-R and the products from these reactions were digested with EcoRI.

Human and Porcine GIP Promoter-Luciferase Plasmids

A series of luciferase reporter plasmids from Dr. Yukihiro Fujita (132) consisting of the luciferase gene driven by -93 bp to -2.9 kb of the human GIP promoter relative to the transcriptional start site and a series of luciferase reporter plasmids consisting of the luciferase gene driven by different lengths of the porcine GIP promoter-luciferase plasmids, generated by Dr. Robert Baker (Centre for Human Islet Transplant and Beta-Cell Regeneration, University of British Columbia, Vancouver, BC), were used to assess the functional effect of truncating the distal well-conserved region of the GIP promoter in species other than rat. The distal element used for generating the porcine series, which ranges from -2.1 kb to -4.5 kb relative to the

transcriptional start site, was removed from the pSSBSssGIP(223-237) vector supplied by enGene, Inc.

In vitro Measurement of GIP Promoter Activity

Cell Culture and Transfection

STC-1 cells (from Dr. Daniel Drucker, Department of Medicine, Samuel Lunenfeld Research Institute, Mt. Sinai Hospital, University of Toronto, Toronto, ON; passages 15-32) were cultured in High Glucose Dulbecco's Modified Eagle Medium (GIBCO, Grand Island, NY) supplemented with 10% foetal bovine serum (GIBCO, Grand Island, NY) and 1 x Pen-Strep (GIBCO, Grand Island, NY; final concentration 100 U penicillin and 100 μg streptomycin per 500 mL) at 37°C and 10% CO_2. Cells were seeded at 3 x 10^5 cells / well in 6-well tissue culture-treated plates (Corning Inc., Corning, NY) 24 h prior to transfection. Cells were co-transfected with experimental constructs consisting of firefly (*Photinus pyralis*) luciferase driven by various lengths of the GIP promoter and a co-reporter construct phRL-TK (Promega, Madison, WI) consisting of the HSV-TK promoter driving the expression of sea pansy (*Renilla reniformis*) luciferase using Metafectene Pro (Biontex, Martinsried, Germany) at 0.5 μg/well and 0.016 μg/well, respectively. A Metafectene Pro-to-DNA ratio of 4 μl/μg was used in 2 ml/well of Minimum Essential Medium (GIBCO, Grand Island, NY) for 5 hours. Each plasmid was transfected in triplicate per trial.

Dual Luciferase Reporter Assays

At 48 h post-transfection, cells were lysed in 400 µl lysis buffer supplied with the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). For each reaction, 20 µL of lysate was assayed with 100 µL of luminal and 100 µL of Stop and Glo reagent, according to the kit protocol. Each transfection replicate was assayed twice to ensure technical reproducibility. Luminosity was read on the Infinite®M1000 plate reader (Tecan, Durham, NC) at room temperature. Luminosity readings from the firefly luciferase reaction were normalized to the readings from the *Renilla* luciferase reaction and all readings were then reported as fold change in luminosity from the reading obtained from the pGL4.10 [luc2] promoterless backbone.

Identification of DNA-binding Sites

Isolation of Nuclear Extracts from STC-1 Cells

Nuclear protein extracts were isolated from STC-1 cells using the Schreiber Method (157) with modifications. Briefly, cells were washed with PBS and harvested into microcentrifuge tubes using a cell scraper and pelleted at 1500 g at 4°C for 5 min. The supernatant was removed and replaced with 400 µL Buffer A (10 mM HEPES, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA supplemented with a Complete Mini-EDTA-free protease inhibitor tablet (Roche, Indianapolis, IN) per 10 mL immediately prior to use). Cells were incubated on ice for 15 min and 25 µl of Nonidet P-40 (Roche, Mannheim, Germany) was added and vortexed briefly to release cytoplasmic protein content. Remaining intact nuclei were pelleted by centrifugation (15,000 g at 4°C for 5 s) and the cytoplasmic protein fraction was removed. The pellet was resuspended in 50 µL of Buffer B (20 mM HEPES, 0.4 M NaCl, 1 mM EDTA, 1 mM EGTA, 20% glycerol supplemented with a Complete Mini-EDTA-free protease inhibitor tablet (Roche, Indianapolis, IN) per 10 mL immediately prior to use) and vortexed for 15 min at 4°C to

release nuclear protein content. The supernatant was collected following a centrifugation step (15,000 g at 4°C for 5 min) and protein yield was quantified using the BCA Assay (Pierce, Rockford, IL).

Probes

Overlapping probes (Figure 2) spanning the region of the GIP promoter between -2894 bp and -2574 bp of the rat GIP promoter were ordered as single-stranded oligonucleotides from Integrated DNA Technologies (Coralville, IA) and biotin labelled using the Biotin 3' End DNA Labelling Kit (Pierce, Rockford, IL). A dot blot by hand was used to assess labelling efficiency and DNA crosslinking to membrane was performed using a Hoefer UVC500 commercial UV (254 nm) crosslinker (Holliston, MA) and subsequent detection using the Chemiluminescent Nucleic Acid Detection Module (Pierce, Rockford, IL). Biotin-labelled oligonucleotides were annealed to complementary unlabelled oligos in a thermocycler (95°C for 5 min, -1°C per min until 25°C). Unlabelled probes for competition assays were generated similarly, with the biotin-labelling step omitted.

Electrophoretic Mobility Shift Assays

Electrophoretic mobility shift assays (EMSAs) were performed according to the protocol provided with the LightShift Chemiluminescent EMSA Kit (Pierce, Rockford, IL) using a 6% non-denaturing polyacrylamide gel in 0.5x TBE and the following binding conditions: 50 ng/µL poly(deoxyinosinic-deoxycytidylic) acid, 0.5 mg/mL BSA, 5% glycerol, 100 mM KCl, 5 mM $MgCl_2$, and 0.1% Nonidet P-40. For each probe where a shift was demonstrated, a competition assay containing 200x unlabelled probe over biotin-labelled probe was performed to distinguish specific versus non-specific binding. Binding of nuclear extract proteins to the labelled DNA probe is detected as an upwardly shifted band compared to bands present in a

binding reaction that does not contain nuclear extract, since the protein complex bound to the DNA probe retards its migration down the gel. Specific shifts are defined as upwardly-shifted bands that are no longer detected in a parallel competition assay in which unlabelled probes in 200-fold molar excess of the labelled probes are included in the binding reaction. In these competition assays, the unlabelled probes were allowed to incubate with the nuclear extract for 5 min before the labelled probe was added to further ensure that the majority of protein-DNA complexes are formed with the unlabelled competing probe.

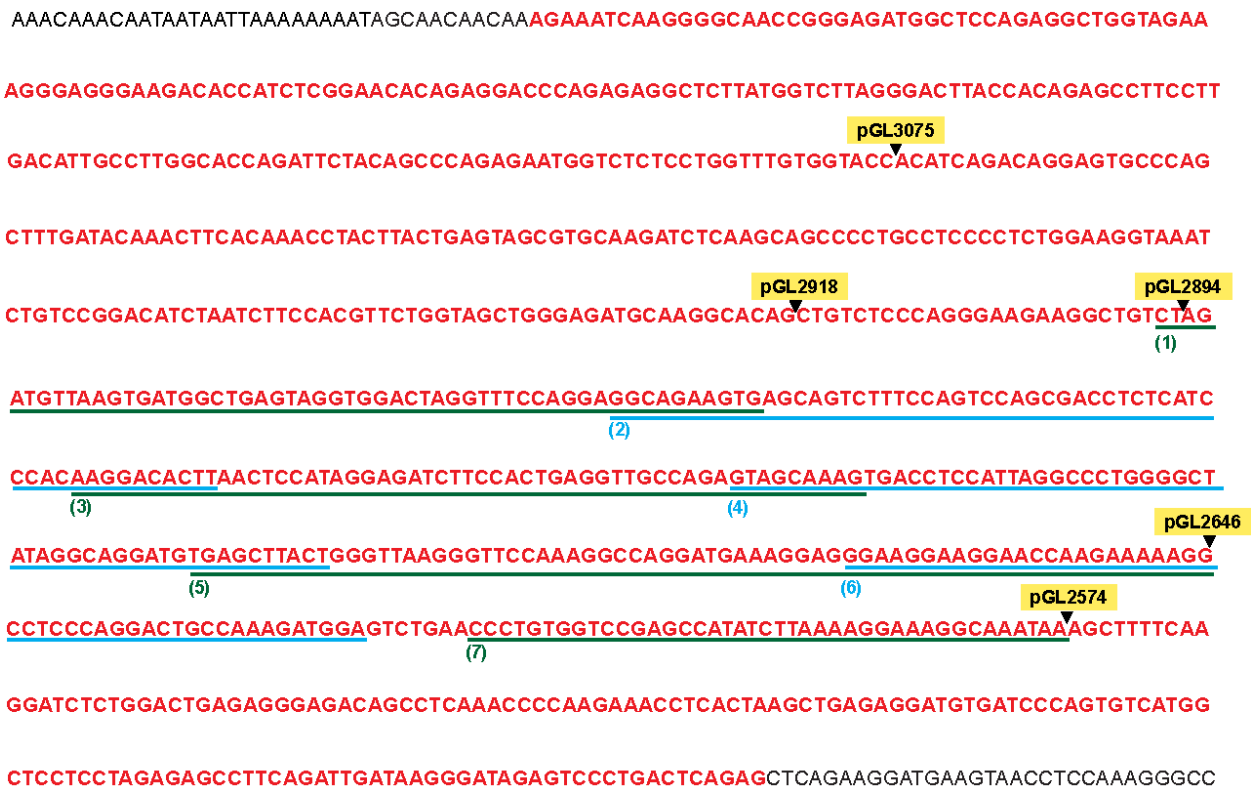


Figure 2. Schematic of overlapping EMSA probes. In this partial sequence of the distal rat GIP promoter from -3301 bp to -2399 bp, the distal well-conserved region is highlighted in red and the distal extents of the rat GIP promoter driving luciferase gene expression in pGL3075, pGL2918, pGL2894, pGL2646, and pGL2574 plasmids are indicated in yellow. Overlapping EMSA probes are indicated in alternating green (odd-numbered probes) and blue (even-numbered probes).

Each probe for which specific binding was demonstrated was subsequently used to perform a series of super-shift EMSAs to screen for transcription factors that may be present in

the binding complex. In super-shift EMSA reactions, antibodies directed against transcription factors that might be bound to the DNA probes were included in the binding reaction to attempt to identify transcription factors involved in certain transcription factor complexes. If the transcription factor that is recognized by the antibody was present in the protein-DNA complex, binding of the antibody to its antigen was detected as a super-shifted band, that is, a band that is shifted even higher than the upwardly shifted band caused by incubation of the DNA probe with nuclear extract alone. The antibodies used for this screen are listed in Table 1.

Table 1. List of antibodies used in super-shift EMSAs.

Antibody	Source	Catalogue Number
rabbit α Pax6 (H-295), polyclonal	Santa Cruz Biotechnology	sc-11357
mouse α Pax6, monoclonal	Millipore	MAB5552
rabbit α Pax6, polyclonal	Covance	PRB-278P
goat α Isl1 (K-20), polyclonal	Santa Cruz Biotechnology	sc-23590
goat α Isl1, polyclonal	R&D Systems	AF1837
rabbit α GATA4 (H-112), polyclonal	Santa Cruz Biotechnology	sc-9053
goat α GATA4 (C-20), polyclonal	Santa Cruz Biotechnology	sc-1237X
rabbit α GATA6 (H-92), polyclonal	Santa Cruz Biotechnology	sc-9055X
rabbit α Pdx1 (N-terminal),	Dr. Joel Habener	N/A
rabbit α Pdx1 (C-terminal)	Dr. Joel Habener	N/A

Assessment of Endogenous Transcription Factor-DNA Interaction in STC-1 Cells

Chromatin Immunoprecipitation (ChIP)

In order to determine whether Pax6 and Pdx1 bind the GIP promoter endogenously in STC-1 cells, ChIP was preformed using the EpiQuik Chromatin Immunoprecipitation Kit (Epigentek, Brooklyn, NY) according to kit protocol. Briefly, STC-1 cells were grown to confluence in 10 cm plates and cross-linked using 1% paraformaldehyde. Cells were lysed and DNA was sheared at 4°C using a sonicator (Artek Systems, Farmingdale, NY) at power 60 for

5 x 15 s with 30 s rest in between. In addition to RNA polymerase (positive control) antibody supplied with the kit, 4 μg of the following antibodies were used for binding to the assay plate: Rabbit IgG (Santa Cruz Biotechnology, sc-2027); Pax6 (Chemicon, MAB5552); Pdx1, N-term (from Dr. Joel Habener). The Pdx1 antibody, supplied as full serum from rabbit was purified using the Melon Gel IgG Spin Purification Kit (Pierce, Rockford, IL), which removes serum proteins from full serum. The concentration of the purified antibody was determined using a BCA Assay (Pierce, Rockford, IL). Following an incubation of the sheared DNA with the bound antibodies, the assay plate was washed to remove unbound DNA, and bound DNA was treated with proteinase K to reverse cross-links. Bound DNA was then used as a template for end-point PCR using primers that amplify the region containing both the putative distal Pax6 and Pdx1 sites (Figure 3).

Figure 3. Region of the distal rat GIP promoter amplified by the ChIP primers. The amplified region contains a putative Pax6 binding site (green) and Pdx1 binding site (yellow).

Isolation and Purification of Primary Mouse GIP-expressing Cells

Animals

Transgenic mice expressing dsRed2 driven by 3.1 kb of the rat GIP promoter (GIP-dsRed2 transgenic mice) (80), on a C57/BL6 background were a generous gift from Dr. Burton Wice (Department of Molecular Biology and Pharmacology, Washington University Medical School, St. Louis, MO). The majority of cell isolations were performed using these

mice. Using these animals as original breeders, animals were crossed to generate GIP-dsRed2 mice on the B6/NOD background by Dr. Majid Mojibian, which was used for one replicate of dsRed2+ cell isolation following the observation of decreased dsRed2 expression in the C57/BL6 line.

Immunohistochemistry

Transgenic mice (13 – 28 weeks old, both males and females) were anaesthetized with isoflurane and perfused with 4% paraformaldehyde through the left ventricle prior to harvesting of the intestine into 4% paraformaldehyde. The tissue was left to fix overnight at 4°C and transferred to 70% EtOH. The tissue was embedded in paraffin and sectioned longitudinally to 5 μm slices at Wax-It Histology Services, Inc (Vancouver, BC). Slides were dewaxed in a series of xylene and ethanol washes. Antigen retrieval was performed using 10 mM citrate buffer (pH 6) for 10 min at 95°C). The sections were blocked with protein blocking solution (DAKO, Carpinteria, CA) for 30 min at room temperature and incubated with primary antibodies (GIP (from Dr. Alison Buchan, 1:10,000); GLP-1 (from Dr. David D'Alessio, 1:10,000); dsRed2 (Chemicon AB3216, 1:100)) overnight at 4°C. The sections were then incubated with donkey Alexafluor secondary antibodies raised against mouse and rabbit IgG (Invitrogen, Carlsbad, CA, 1:1000) for 1 h at room temperature and fixed with Hard Set Vectastain with DAPI (Vector Laboratories, Burlingame, CA). Slides were visualized using a Zeiss Axiovert 200 inverted microscope with the following filter sets (Red: Filter Set 20, excitation: BP 546/12, beam splitter: FT 560, emission: BP 575-640; Green: Filter Set 10, excitation: BP 450-490, beam splitter: FT 510, emission: BP 515-565; Blue: Filter Set 49, excitation: G 365, beam splitter: FT 395, emission: BP 445/50). Images were captured and pseudocoloured using Improvision's OpenLab 5 software.

Isolation of Intestinal Epithelial Cells as a Single Cell Suspension and Fluorescence-activated Cell Sorting (FACS)

A single cell suspension of intestinal epithelial cells was isolated using a modified Weiser's method of isolating intact epithelium (158, 159, 160, 161). Mice were anaesthetized and sacrificed by cervical dislocation and the intestine was removed into Ca^{2+}/Mg^{2+}-free HBSS (GIBCO, Grand Island, NY; supplemented with 25 mM HEPES). The lumen was washed with HBSS containing 250 mM DTT and inverted using a tapered glass rod made by heating a long glass pipette under a flame to create a glass bead on the thicker end of the glass rod. Following inversion, the open ends of the intestinal segment were closed off with sutures and incubated in citrate buffer (96 mM NaCl, 1.5 mM KCl, 27 mM sodium citrate, 8 mM KH_2PO_4, 5.6 mM Na_2HPO_4; pH = 7.4) for 10 min at 37°C. The intestinal segment was then removed and placed in a 1.5 mM EDTA solution in S-MEM (Mediatech, Inc, Manassas, VA; supplemented with 25 mM HEPES and 1 mg/mL BSA; pH = 7.4) for 30 min at 37°C, after which dissociated intact villi were gently removed from the mucosal lining using a cell scraper. The suspension was pelleted (3 min, 0.3 g, 4°C), resuspended in a 0.3 U/mL dispase in S-MEM solution, and incubated in a shaking water bath at 37°C for 10 min. Following dispase incubation, the single cell suspension was pelleted and washed with HBSS and filtered with a 40 μm cell strainer (BD Biosciences, Franklin Lakes, NJ) before resuspension in L-15 (Leibovitz) medium (GIBCO, Grand Island, NY) with 2% FBS for fluorescence-activated cell sorting (FACS). Cells were sorted using a Cytopeia Influx Cell Sorter (BC Cancer Agency Flow Cytometry Core, Vancouver, BC) with a 561 nm laser and 580/30 filter directly into RNAlater (Ambion, Austin, TX) to prevent RNA degradation over the sort period of 5-7 h.

RNA Isolation

RNase-free PBS was added 1:1 v/v to RNAlater in the collection tube and centrifuged at 3000 rpm for 20 min. RNA isolation and DNase I treatment were performed according to kit protocol with the RNAqueous-Micro Kit (Ambion, Austin, TX). The resulting RNA was quantified using the low-range standard curve of the RiboGreen Assay (Molecular Probes, Eugene, OR) and sent to the Centre for Molecular Medicine and Therapeutics (Vancouver, BC) for quality analysis using the Agilent 2100 Bioanalyzer.

Primary Culture of Enriched K-cells

A number of culturing conditions were tested in an attempt to maintain enriched K-cells in culture following FACS. Rat-tail collagen coated tissue-treated 96-well plates (BD Biosciences, Franklin Lakes, NJ) and cover slips (Fisher Scientific, Nepean, ON) were made by dissolving 3.5 mg/mL of rat tail collagen in 0.017 M acetic acid. The solution was allowed to cover the surface of the wells or cover slip for 45 min at room temperature and then aspirated off and left to dry overnight. Plates and coverslips coated in this fashion were stored at 4°C. Commercially-prepared extracellular matrices (collagen I, collagen IV, fibronectin, laminin, poly-d-lysine, Matrigel) in 6-well plates (BD Biosciences, Franklin Lakes, NJ) and coverslips (Fisher Scientific, Nepean, ON) coated with 0.1% poly-l-lysine (Sigma-Aldrich, St. Louis, MO) were also used as culture substrates. Cells were sorted directly from a sorting medium consisting of 2% FBS in PBS into High Glucose Dulbecco's Modified Eagle Medium (GIBCO, Grand Island, NY) supplemented with 10% foetal bovine serum (GIBCO, Grand Island, NY) and 1 x Pen-Strep (GIBCO, Grand Island, NY; final concentration 100 U penicillin and 100 μg streptomycin per 500 mL) and incubated at 37°C and 5% CO_2 at either atmospheric O_2 or low (4%) O_2 conditions.

RESULTS

Previous studies have identified the critical role of several elements in the rat GIP proximal promoter (-193 bp to +19 bp) for transcriptional regulation but the functional contributions of putative elements between -2.5 kb and -193 bp have not yet been characterized. Furthermore, identification of putative *cis*-regulatory elements distal to -2.5 kb of the rat GIP promoter have not yet been pursued, but based on a series of transgenic mouse studies in which reporter genes are driven by varying lengths of either the human or rat GIP promoter, sequences upstream of -2.5 kb of the rat GIP promoter may contribute to tissue-specific expression of GIP. A comparison of 5 kb regions of the GIP promoter upstream to the transcriptional start site in nine species demonstrated that in addition to the well-conserved proximal rat GIP promoter spanning -200 bp upstream of the transcriptional start site, there is a second area of high homology from approximately -2500 bp to -1500 bp (Figure 4).

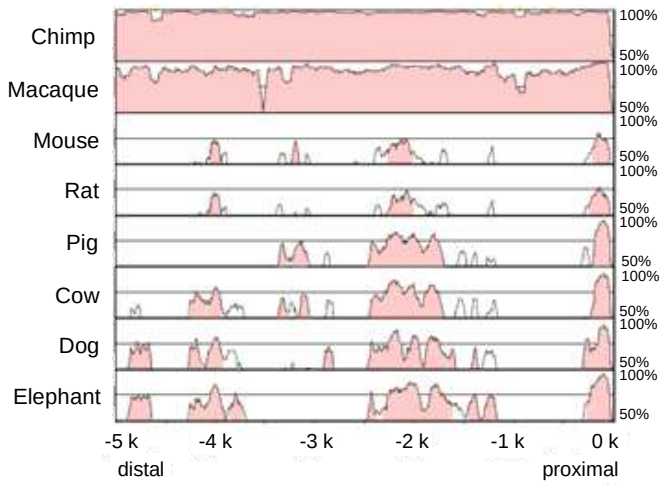


Figure 4. Cross-species homology of the GIP promoter with reference to the human promoter show a region of high homology located at -2500 bp to -1500 bp relative to the transcriptional start site in addition to a region of high homology located between -200 bp and the transcriptional start site in the 5'-upstream promoter. Genomic DNA sequences immediately distal to the transcriptional start site of the GIP gene were compared using mVISTA. Percent-identity plots were plotted with 50% homology shown on the bottom axis and 100% homology on the top axis compared to the human sequence. Homology was calculated over stretches of 100 bp and regions of homology > 70% are highlighted in pink.

A nucleotide-level alignment of the area of high homology spanning -2500 bp and -1500 bp in the human, rat, and pig GIP promoter sequences is seen in Figure 5. Since the model system for K-cells in this thesis was the mouse-derived mixed endocrine intestinal STC-1 cell line, luciferase reporter constructs driven by the rat GIP promoter were used to characterize the contribution of the distal rat GIP promoter to GIP gene regulation. The rat and mouse GIP promoter show high homology with each other, and the use of a rodent cell line to investigate a rodent gene promoter was chosen to mimic as closely as possible the *in vivo* conditions of rodent transcriptional regulation within our model.

Utilizing a truncation series from -2033 bp (pGL2033), -2574 bp (pGL2574), -3075 bp (pGL3075), and -3795 bp (pGL3795) to +19 bp of the rat GIP promoter driving luciferase gene expression, the relative contribution of distal promoter elements to GIP promoter activity was investigated. The pGL3795 plasmid contains the entire distal high homology area, while the sequence of high homology is progressively shortened in the plasmids pGL3075, pGL2574, and pGL2033 (Figure 5). Truncation of the longest construct to -3075 bp resulted in a 14% increase in promoter activity, although the difference in promoter activity between pGL3795 and pGL3075 was not statistically significant. Further truncation of the promoter to -2574 bp decreased promoter activity from the two longest constructs by approximately 50%, and further truncation to -2033 bp did not drop the activity further (Figure 6). This observation suggests that *cis*-regulatory elements that contribute to the activation of the GIP promoter are present in the nucleotide region spanning -3075 bp to -2574 bp relative to the transcriptional start site. Therefore, three more luciferase reporter truncation constructs with 5'-upstream sequence lengths between -3075 and -2574 bp were made based on available restriction enzyme sites in this region (Figure 7). The pGL3075 plasmid was used as a template.

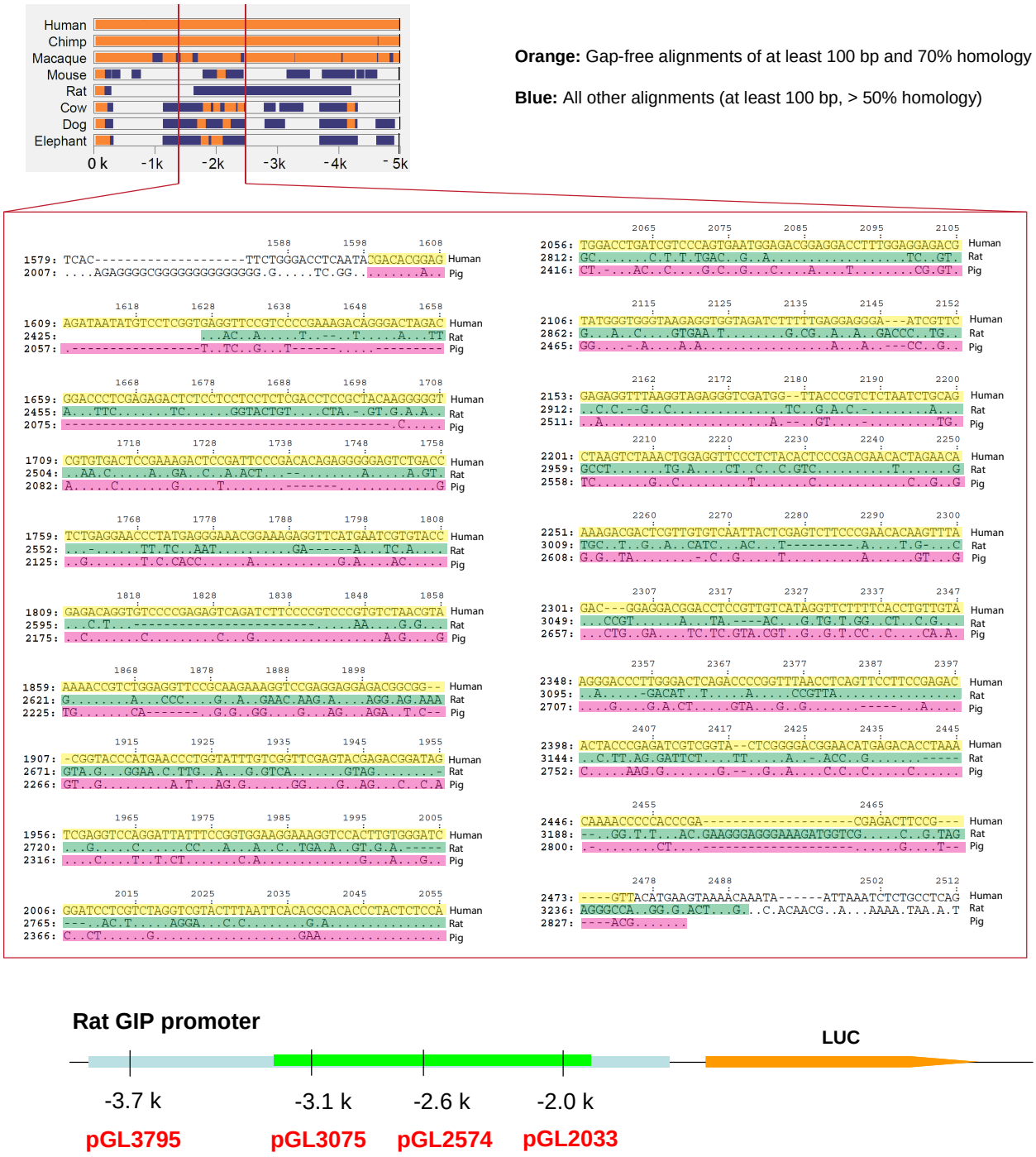


Figure 5. Nucleotide-level alignment of the area of high homology spanning -2500 bp and -1500 bp identified by mVISTA and MultiPipmaker in the human, rat, and pig GIP promoter sequences. Conserved nucleotides relative to the human sequence are indicated by a dot (.) in a homology map between the human (yellow), rat (green), and pig (pink) sequences. The locations of this region in luciferase construct driven by the rat GIP promoter is indicated by green (not to scale). Truncations of this luciferase construct were made at multiple locations (red text, plasmid names correspond to promoter length upstream from the transcriptional start site) to assess the contribution of distal elements to GIP promoter activity.

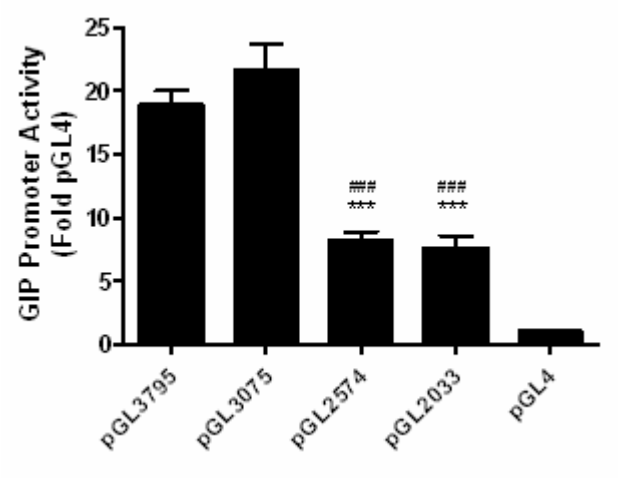


Figure 6. Distal truncation series consisting of various lengths of the rat GIP promoter driving expression of firefly luciferase showed a 50% drop in activity between pGL3075 and pGL2574. STC-1 cells were transfected with the respective luciferase plasmid and luciferase activity was measured 48 h post-transfection using the Dual-Luciferase Reporter assay with 3 technical replicates for each trial (n=4). Luminosity from firefly luciferase activity was normalized to luminosity from renilla luciferase activity and data were expressed as fold activity generated by the promoterless backbone pGL4. Data were expressed as mean fold pGL4 activity and error bars represent standard error of the mean. A one-way ANOVA with Bonferroni post-hoc test was performed. ***$p<0.001$ compared to pGL3795; $^{###}p<0.001$ compared to pGL3075.

As Figure 7 demonstrates, a statistically significant decrease in GIP promoter activity representing 51% of pGL3075 activity was first observed upon truncation of the 5'-upstream sequence to -2646 bp. A less pronounced decrease in activity representing 28% of pGL3075 activity was seen when 157 bp from the distal end of pGL3075 was truncated to make the pGL2918 construct. The activity of the three constructs driven by the longest pieces of the GIP promoter (pGL3075, pGL2918, and pGL2894) did not statistically differ from each other, averaging to a 21-fold increase in promoter activity compared to the promoterless luciferase construct, pGL4. Furthermore, the difference in activity between pGL2646 and pGL2574 was not statistically significant, averaging to 12-fold pGL4 activity. Thus, the search for distal *cis*-regulatory elements was narrowed to the region between 2894 bp and 2646 bp upstream of the transcriptional start site.

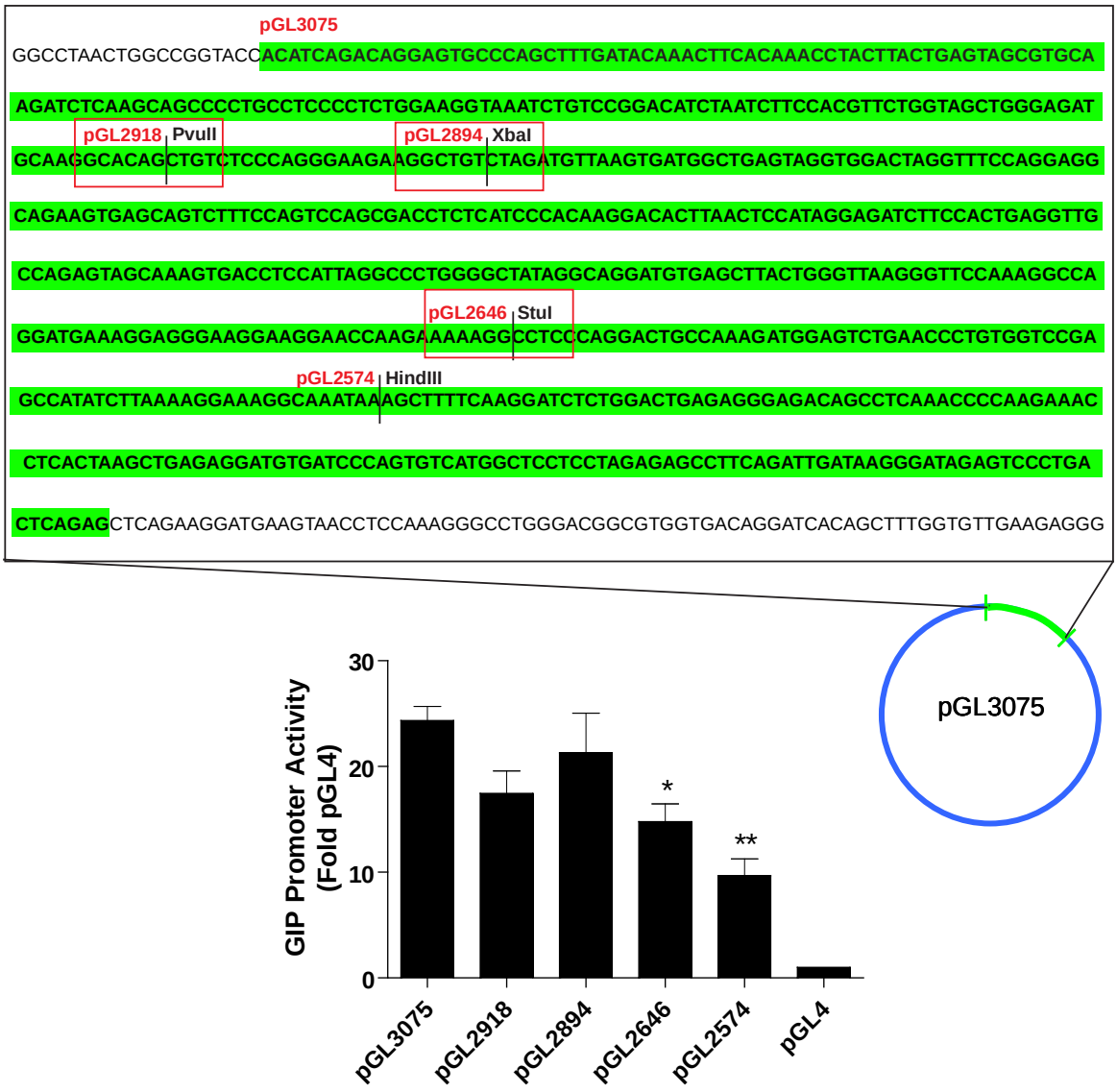


Figure 7. Truncation to pGL2646 produced the first significant drop in activity from pGL3075 in a truncation series generated based on restriction enzyme sites between -3075 bp and -2574 bp. STC-1 cells were transfected with the respective luciferase plasmid and luciferase activity was measured 48 h post-transfection using the Dual-Luciferase Reporter assay with 3 technical replicates for each trial (n=3). Luminosity from firefly luciferase activity was normalized to luminosity from renilla luciferase activity and data were expressed as fold activity generated by the promoterless backbone pGL4. Data were expressed as mean fold pGL4 activity and error bars represent standard error of the mean. A one-way ANOVA with Bonferroni post-hoc test was performed. *$p<0.05$, **$p<0.01$ compared to pGL3075.

In order to determine the location of transcriptional complex binding, seven probes were generated based on the sequence between -2894 bp and -2574 bp (Figure 2) for EMSAs with nuclear extracts from STC-1 cells. Of the seven probes, only one (probe 2) did not show a band shift when nuclear protein extract was included in the binding reaction (Figure 8). Specificity of

the band shifts was determined by a binding reaction containing both nuclear protein extract and a cold probe, that is, an unlabelled version of the probe of interest in 200-fold molar excess. Competition between the cold probe and biotin-labelled probe is expected to remove all specific shifts in this binding reaction; these are indicated as filled arrows in Figure 8.

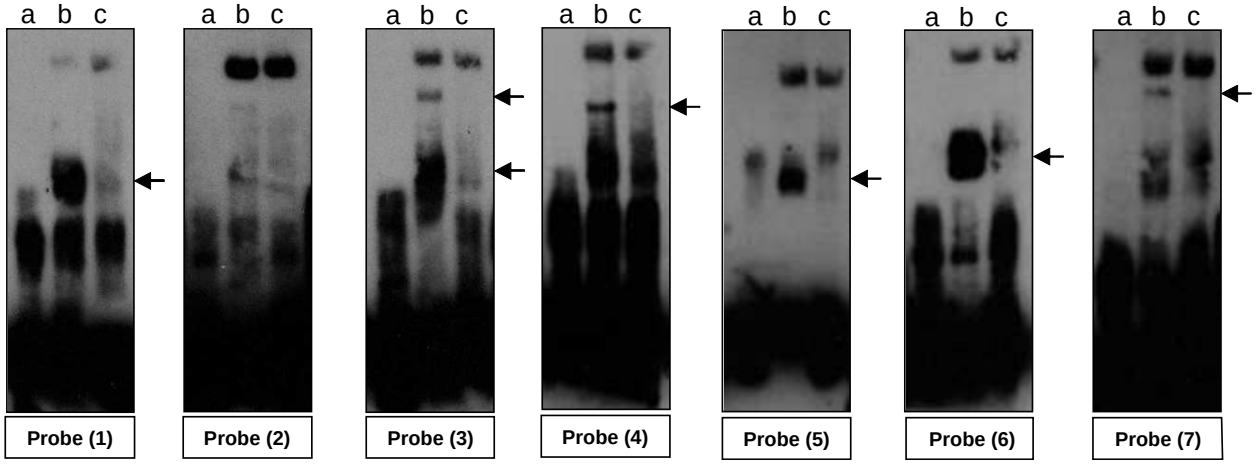


Figure 8. Specific binding of nuclear protein to DNA was observed in all seven probes except probe 2. Biotin-labelled probes 1-7 were incubated in a binding reaction a) alone, b) with nuclear protein extract from STC-1 cells, and c) with both nuclear protein extract from STC-1 cells and 200-fold molar excess of the same unlabelled probe. Filled black arrows indicate specific band shifts.

The probes that produced specific band shifts were entered into the transcription factor binding site prediction software, ALGGEN-PROMO. A visual scan of the probe sequences was also performed to identify consensus sequences of candidate transcription factors which have been shown to bind proximal promoter regulatory elements. Putative binding sites for Pax6, GATA transcription factors, and Pdx1 were identified on probes 1, 3, 4, 6, and 7. Super-shift assays (Figures 9 and 10) were performed to assess if these factors were a part of the complexes producing the band shift observed in Figure 8.

In the initial screen (Figure 9), the biotin-labelled probes were added to the binding reaction containing the nuclear extract prior to the antibody. This order of addition allowed binding of the antibody to the bound form of their respective transcription factor antigens.

However, certain epitopes could be hidden in the DNA-bound conformation of the transcription factor, resulting in a lack of super-shift despite the presence of actual transcription factor binding. Therefore, binding reactions for which no super-shift bands were observed were subjected to a second screen (Figure 10) in which the antibody was added to the binding reaction containing the nuclear extract prior to the biotin-labelled probe. This allowed antibodies to bind unbound transcription factors before participating in the binding reaction with the biotin-labelled DNA probe. Shifted bands resulting from specific binding are indicated with filled black arrows while super-shifted bands are indicated with open black arrows in Figures 9 and 10.

The two super-shift EMSAs screens demonstrated that antibodies directed against Pax6 and Pdx1 produced super-shifted bands in all the probes that were screened (lanes e, l, and m in Figures 9A, 9B, 9D, 9E) except probe 4, for which only a binding reaction containing Pdx1 antisera produced a super-shifted band (lane h in Figure 10C). In addition, GATA4 antisera produced a super-shifted band in probe 3 (lane i in Figure 9B) and GATA6 antisera produced a super-shifted band in probes 4 (lane k in Figure 9C) and 6 (lane h in Figure 10D). In screen two of probe 1, all of the lanes appear to produce super-shifted bands (lanes d to h in Figure 10A), but since the pattern of specific band shifts (black filled arrows) differ from those observed in Figure 9A even without antibody addition, the bands currently marked as super-shifted bands (open arrows, Figure 10A) may be false positives. In any case, there were no binding sites identified in probe 1 that matched the factors where possible binding was demonstrated in Figure 10A. Only Pax6 and Pdx1 antisera produced super-shifted bands in binding reactions containing probes in which putative binding sites for the corresponding transcription factor were identified (green arrows, Figure 9 and 10).

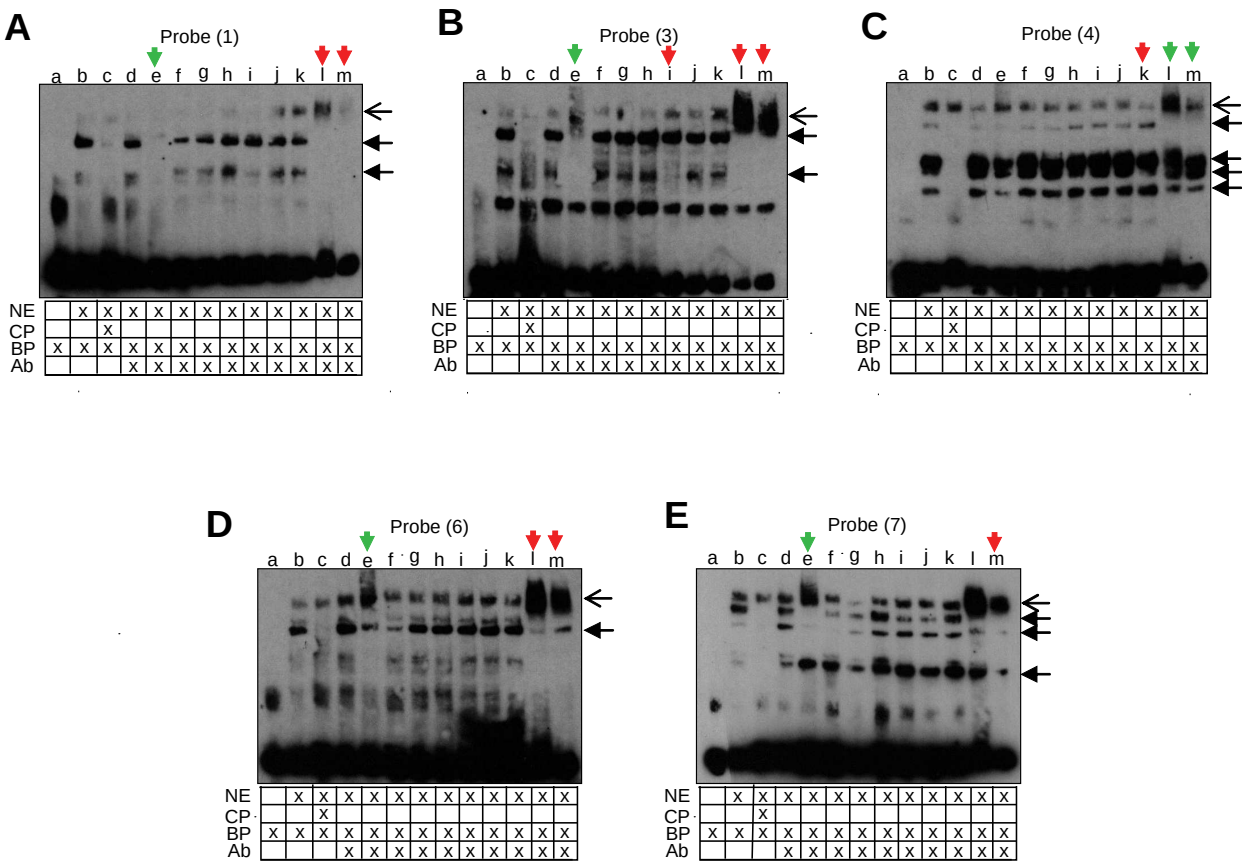


Figure 9. Super-shift EMSAs demonstrate presence of Pax6 and Pdx1 binding in probes 1, 3, 6, and 7. For EMSA probe 1, 3, 4, 6, and 7 (A-E, respectively), biotin-labelled probes were incubated in binding reactions a) alone, b) with nuclear extracts from STC-1, c) with both nuclear extracts from STC-1 and 200-fold molar excess of the same unlabelled probe. To assess if candidate transcription factors form part of the NE-DNA binding complex, a panel of antibodies were added to binding reactions containing the biotin-labelled probe and nuclear extracts. The factors screened include: Pax6 (d: sc-11357, e: MAB5552, f: PRB-278P); Isl1 (g: sc-23590, h: AF1837); GATA4 (i: sc-9053, j: sc-1237X); GATA6 (k: sc-9055X); Pdx1 (l: N-term, m: C-term). NE: nuclear extract; CP: cold probe; BP: biotinylated probe; Ab: antibody. Green arrows indicate the presence of a super-shifted band in a binding reaction for which binding sites have been identified on the probe for the particular transcription factor. Red arrows indicate the presence of a super-shifted band in a binding reaction for which there is no predicted binding site for the transcription factor represented. Filled black arrows indicate specific band shifts and open black arrows indicate super-shifted bands.

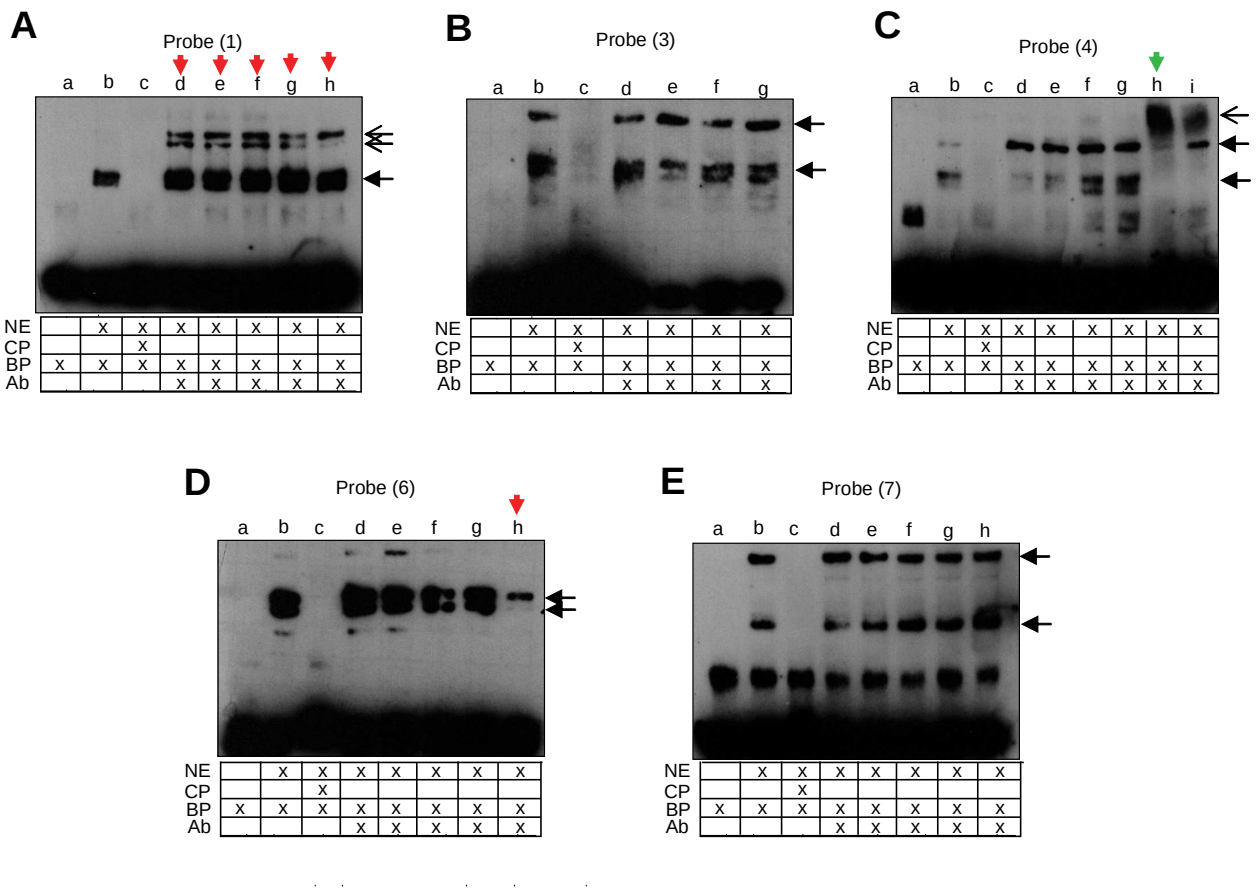


Figure 10. Super-shift EMSAs detected additional Pdx1 binding in probe 4. For EMSA probe 1, 3, 4, 6, and 7 (A-E, respectively), a second screen was performed in which the order of addition of antibody and biotin-labelled probe was reversed. Biotin-labelled probes were incubated in binding reactions a) alone, b) with nuclear extracts from STC-1, c) with both nuclear extracts from STC-1 and 200-fold molar excess of the same unlabelled probe. To assess if candidate transcription factors form part of the NE-DNA binding complex, a panel of antibodies were added to binding reactions containing nuclear extracts for 5 min followed by the addition of the biotin-labelled probes. The factors screened were as follows. (A, D, E): Isl1 (d: sc-23590, e: AF1837); GATA4 (f: sc-9053, g: sc-1237X); GATA6 (h: sc-9055X), (B) Isl1 (d: sc-23590, e: AF1837); GATA4 (f: sc-1237X); GATA6 (g: sc-9055X), (C) Isl1 (d: sc-23590, e: AF1837); GATA4 (f: sc-9053, g: sc-1237X); Pdx1 (h: N-term, i: C-term). NE: nuclear extract; CP: cold probe; BP: biotinylated probe; Ab: antibody. Green arrows indicate the presence of a super-shifted band in a binding reaction for which binding sites have been identified on the probe for the particular transcription factor. Red arrows indicate the presence of a super-shifted band in a binding reaction for which there is no predicted binding site for the transcription factor represented. Filled black arrows indicate specific band shifts and open black arrows indicate super-shifted bands.

Of particular interest is the super-shift patterns observed for probes 3 and 4. Probe 3 appears to bind Pax6 (lane e, Figure 9B) and contains a Pax6 binding site predicted by the ALGGEN-PROMO transcription factor site prediction software (AGGACA) and probe 4 appears to bind Pdx1 (lane h, Figure 10C) and contains the Pdx1 consensus sequence (CATTAG). Both Pdx1 antisera also produced super-shifted bands when incubated with probe 3 (lanes l & m,

Figure 9B). Of the antibodies used in the super-shift EMSAs (Table 1), the Pdx1 antisera were the only full serum antibodies, and the observation that one or both of the Pdx1 antisera produced super-shifted bands in every single probe tested raised concern about the specificity of the antibody. To assess whether the super-shifted bands produced by the Pdx1 anti-sera was specific and not a false positive produced by the non-specific binding of serum proteins to the probe, the binding reactions were repeated by including a binding reaction that contained pre-immune serum from the animal used to generate the antibody. The removal of the specific binding band (indicated with a filled black arrow, Figure 11A) in probe 3 in pre-immune reaction (lane h) and presence of one of the two specific binding bands (indicated with filled black arrows, Figure 11B) in probe 4 in the pre-immune reaction (lane h) suggests that the super-shift bands produced by incubation with Pdx1 antisera was non-specific in probe 3 but specific in probe 4. Whereas GATA4 was previously seen to produce a super-shifted band in probe 3 (lane i, Figure 9B) and GATA 6 was previously seen produce a super-shifted band in probe 4 (lane k, Figure 9C), these observations were not reproduced in Figure 11.

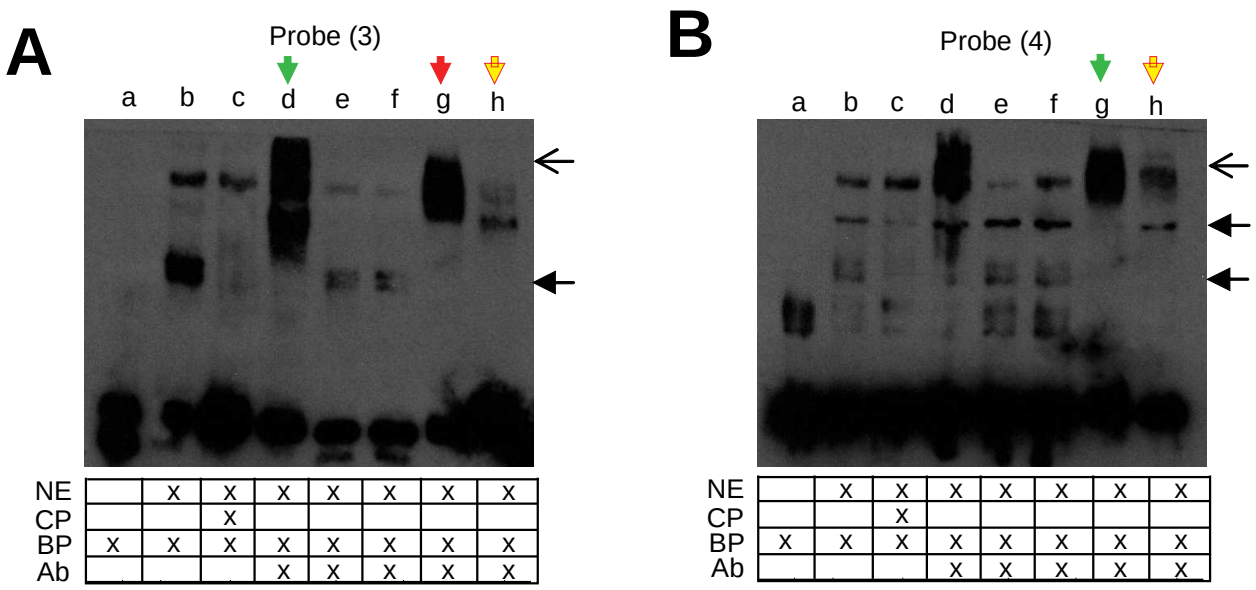


Figure 11. Super-shift EMSAs demonstrated Pax6 and Pdx1 binding in probes 3 and 4. Biotin-labelled probes were incubated in binding reactions a) alone, b) with nuclear extracts from STC-1, c) with both nuclear extracts from STC-1 and 200-fold molar excess of the same unlabelled probe. The participation of Pax6, Pdx1, and GATA transcription factors in the complex was assessed in binding reactions containing antibodies directed against d) Pax6 (MAB5552), e) GATA4 (sc-9053), f) GATA6 (sc-9055X), g) Pdx1 (N-term), h) Pdx1 (N-term) pre-immune serum. NE: nuclear extract; CP: cold probe; BP: biotinylated probe; Ab: antibody. Green arrows indicate the presence of a super-shifted band in a binding reaction for which binding sites have been identified on the probe for the particular transcription factor. Red arrows indicate the presence of a super-shifted band in a binding reaction for which there is no predicted binding site for the transcription factor represented. Yellow arrow indicates the pre-immune serum control. Filled black arrows indicate specific band shifts and open black arrows indicate super-shifted bands.

The functional contribution of both the putative Pax6 (AGGACA) and Pdx1 (CATTAG) binding sites were evaluated by mutating these sites to GGGCCC and GAATTC, respectively in the pGL3075 plasmid. These mutated sites introduced the restriction enzyme sites ApaI (GGGCCC) and EcoRI (GAATTC) into the plasmid, incorporating an effective screening tool for the plasmid containing the mutation. In addition, each site changed at least 50% of the nucleotides in the original putative binding site at both the flanking nucleotide and at least one additional nucleotide in the middle of the site, ensuring sufficient alteration of the original sequence. Lastly, 2/3 of the nucleotide changes in creating the Pax6 site mutation and all of the nucleotide changes in creating the Pdx1 site mutation consisted of changing a purine to a

pyrimidine, or vice versa, changing not only the sequence of nucleotides but also the physical structure of the binding site. Mutation of the Pax6 site was sufficient to drop the promoter activity of pGL3075 to a level equivalent to removing 429 bp from the distal end of this construct (Figure 12A). The mutation of the putative Pdx1 site also decreases promoter activity, but to a lesser extent compared to the decrease in activity mediated by mutation of the putative Pax6 site (Figure 12B). Mutation to both putative binding sites simultaneously reduced promoter activity to a level comparable to mutating the putative Pax6 binding site only (Figure 13).

A pitfall of EMSAs is that the binding of transcription factor to DNA to form transcriptional regulatory complexes occurs in relatively artificial conditions. Therefore, ChIP was performed to assess whether Pax6 and Pdx1 binds to the putative binding sites of interest on the endogenous GIP promoter. Sheared DNA fragments ranging from 500-1500 bp (data not shown) were used as input DNA for immunoprecipitation and the DNA that was pulled down in the assay was used as a temple for endpoint PCR using primers that amplified the region containing the putative Pax6 and Pdx1 sites mutated in the aforementioned functional studies (see Material and Methods). The presence of PCR products in the two assays containing antibodies directed against Pax6 and Pdx1 demonstrates endogenous binding of Pax6 and Pdx1 to chromatin isolated from STC-1 cells (Figure 14).

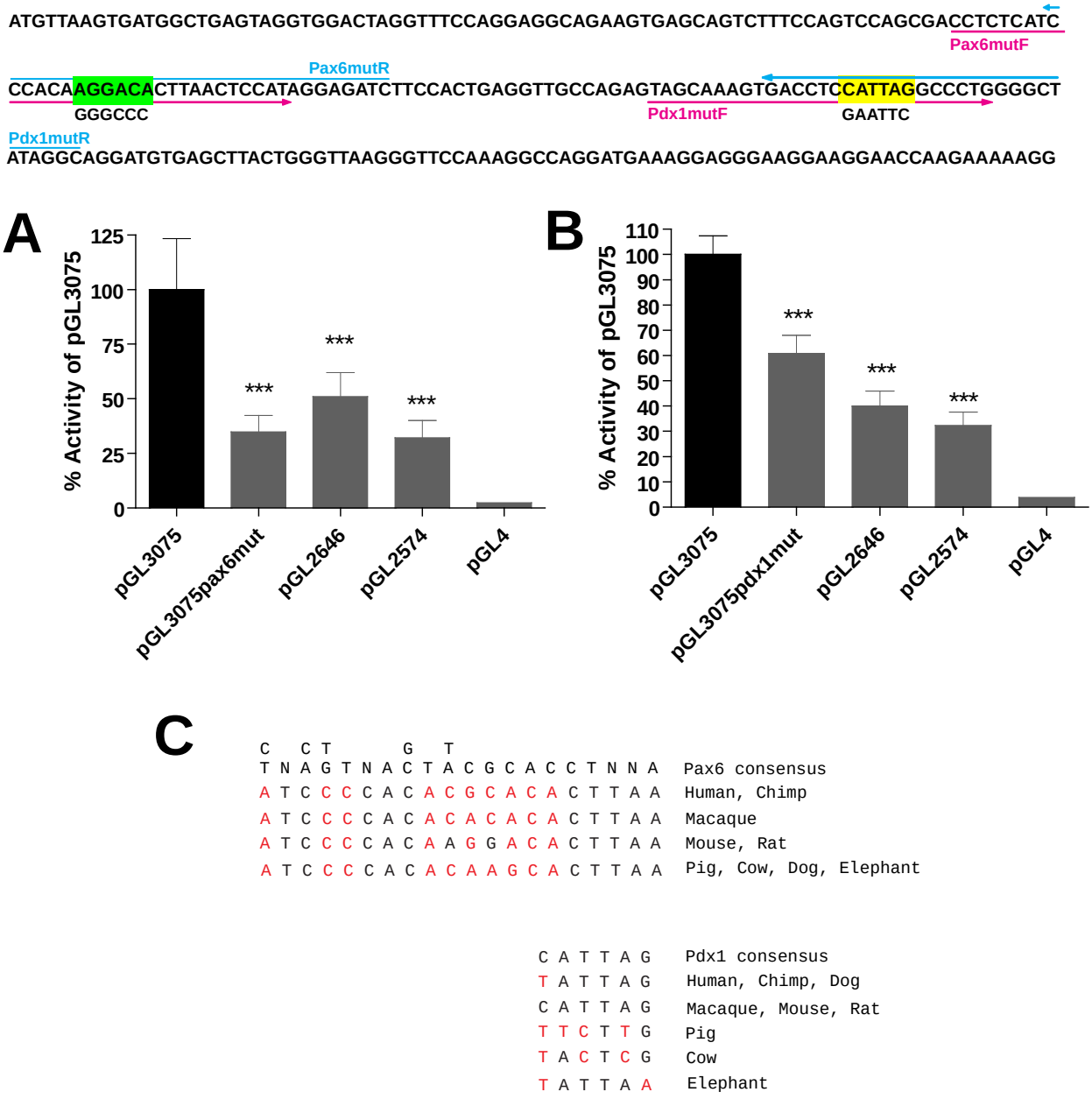


Figure 12. Mutation of putative Pax6 and Pdx1 binding sites significantly reduced promoter activity relative to pGL3075. STC-1 cells were transfected with the respective luciferase plasmid and luciferase activity was measured 48 h post-transfection using the Dual-Luciferase Reporter assay with three technical replicates for each trial (n=3). Mutation of the putative Pax6 binding site from AGGACA to GGGCCC (pGL3075pax6mut) reduced pGL3075 promoter activity to a level equivalent to reducing the promoter length to that in pGL2646 (A). Mutation of the putative Pdx1 binding site from CATTAG to GAATTC (pGL3075pdx1mut) also reduced pGL3075 promoter activity significantly. Luminosity from firefly luciferase activity was normalized to luminosity from renilla luciferase activity and data were expressed as fold activity generated by the promoterless backbone pGL4. Data were expressed as mean fold pGL4 activity and error bars represent standard error of the mean. A one-way ANOVA with Bonferroni post-hoc test was performed. ***$p<0.001$ compared to pGL3075. The homology between the putative Pax6 and Pdx1 binding sites is shown in (C), where black text indicates a nucleotide match to the consensus sequence and red text indicates a mismatch from the consensus sequence.

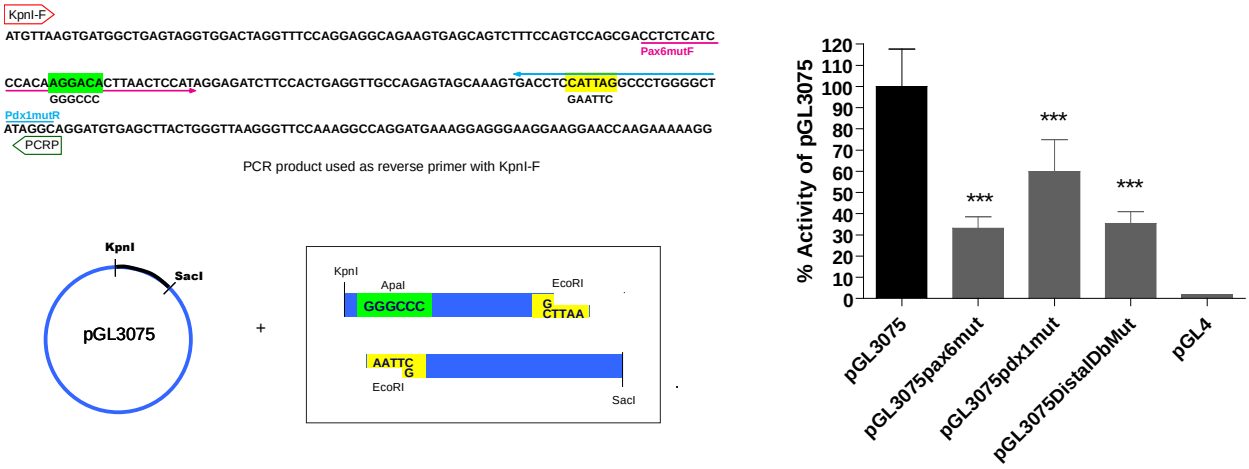


Figure 13. Mutation of both the distal putative Pax6 and Pdx1 binding sites did not drop promoter activity any further than mutation of the putative Pax6 binding site alone. STC-1 cells were transfected with the respective luciferase plasmid and luciferase activity was measured 48 h post-transfection using the Dual-Luciferase Reporter assay with 3 technical replicates for each trial (n=3). Mutation of both the putative Pax6 binding site from AGGACA to GGGCCC and the putative Pdx1 binding site from CATTAG to GAATTC together (pGL3075DistalDbMut) dropped promoter activity level to that of mutating the putative Pax6 binding site alone (pGL3075pax6mut). pGL3075pdx1mut is the plasmid containing the mutated putative Pdx1 binding site alone. Luminosity from firefly luciferase activity was normalized to luminosity from renilla luciferase activity and data were expressed as fold activity generated by the promoterless backbone pGL4. Data were expressed as mean fold pGL4 activity and error bars represent standard error of the mean. A one-way ANOVA with Bonferroni post-hoc test was performed. ***p<0.001 compared to pGL3075.

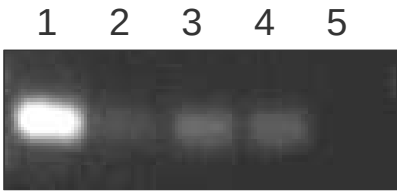


Figure 14. Endogenous binding of Pax6 and Pdx1 to the distal GIP promoter in STC-1 cells. Chromatin immunoprecipitation was performed using IgG antibodies and the chromatin immunoprecipitated by antibodies (Lane 1 – rabbit α RNA polymerase; Lane 2 – rabbit α IgG negative control (sc-2027); Lane 3 – mouse α Pax6 (MAB5552); Lane 4 – rabbit α Pdx1 (N-term); Lane 5 – H_2O (no chromatin) control for PCR) were used as templates for PCR using primers that flank the putative Pax6 and Pdx1 binding sites in the distal GIP promoter.

To evaluate if these data can be generalized to other species, a truncation series of luciferase constructs driven by the human and pig GIP promoters was compared against the truncation series of the rat promoter. Despite a high degree of homology between -2500 bp and -1500 bp as previously shown in Figure 5, the elements present in this region do not appear to play the same functional role between species (Figure 15). Although truncation along the well-conserved region indicated by the red horizontal bars (Figure 15) occur at different locations in the three different species, the overall effect of truncations made in this region appear to yield different effects. In stark contrast to the reduction in promoter activity seen after truncation of this region in the rat promoter, as seen throughout the rest of this thesis and in Figure 15b, truncation of the region in the human promoter appear to have no effect (Figure 15a) while truncation in the pig promoter results in an increase in promoter activity (Figure 15c).

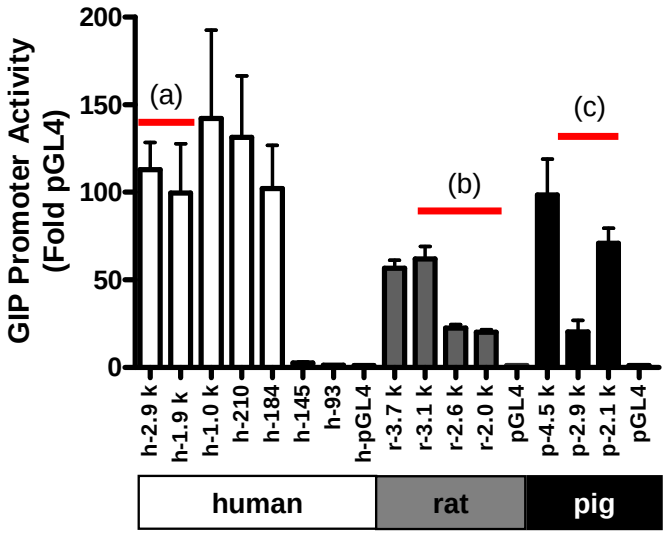


Figure 15. Cross-species comparison of the effect of truncating the distal well-conserved region (-2500 bp to -1500 bp) of the GIP 5'-upstream flanking sequence indicates a different functional role of regulatory elements in this region in different species. STC-1 cells were transfected with the respective luciferase plasmid and luciferase activity was measured 48 h post-transfection using the Dual-Luciferase Reporter assay with 3 technical replicates for each trial (n=3). There appeared to be (a) no effect on promoter activity by truncating the human GIP promoter within the well-conserved distal promoter, (b) a decrease in activity in the rat promoter, and (c) an increase in activity in the pig promoter.

In addition to determining the contribution of distal regulatory elements by identification of sites where candidate transcription factors may bind, a method was developed to isolate primary intestinal GIP-expressing cells for microarray-based non-biased screening for transcription factors that are highly expressed in GIP-expressing cells over non-GIP-expressing cell types. A single cell suspension of intestinal epithelial cells was isolated from transgenic mice expressing dsRed2 driven by 3.1 kb of the rat GIP promoter (80) and sorted based on this fluorescent marker. As Figure 16 demonstrates, immunohistochemical analysis of tissue from these transgenic animals indicated that a majority of dsRed2-immunoreactive cells expressed GIP as expected, indicating that the purity of the cell sample isolated by this method was high (94%). A large number of dsRed2 immunoreactive cells also co-stained for GLP-1 (78%), although examples of dsRed2 immunoreactive cells that are not immunoreactive for GIP and dsRed2 immunoreactive cells that are not immunoreactive for GLP-1 were also found (Figure 16). A sample field of view of the background intensity of staining is seen in Figure 17. The lack of cross-reactivity between the GIP and GLP-1 antibodies used in these studies was previously demonstrated by A. Asadi (Laboratory of Molecular and Cellular Medicine, University of British Columbia, Vancouver, BC; data not shown).

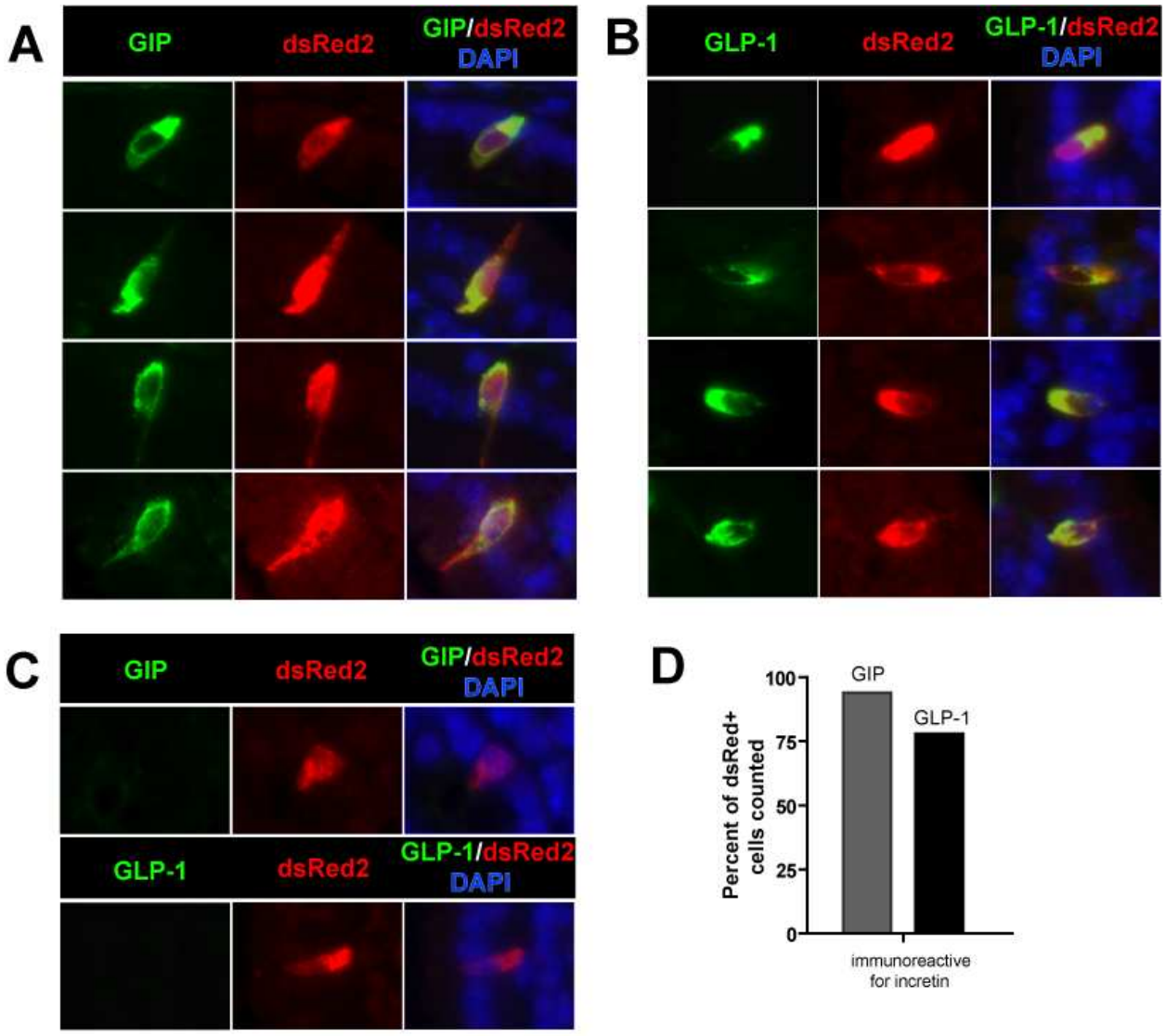


Figure 16. A majority of dsRed2-expressing cells expressed GIP and GLP-1 in GIP-dsRed2 transgenic mouse intestine. Duodenal and jejunal sections from GIP-dsRed2 transgenic mice were co-stained with GIP/dsRed2 and GLP-1/dsRed2 antibodies. In at least three non-adjacent sections from at least two different animals, cells that were dsRed2 immunoreactive were then checked for either GIP or GLP-1 immunoreactivity. (A) Examples of dsRed2 immunoreactive/GIP immunoreactive cells, (B) Examples of dsRed2 immunoreactive/GLP-1 immunoreactive cells, (C) Examples of dsRed2 immunoreactive cells that were not immunoreactive for GIP or GLP-1 (D) 94% of dsRed2 immunoreactive cells were immunoreactive for GIP, compared to 78% that were immunoreactive for GLP-1.

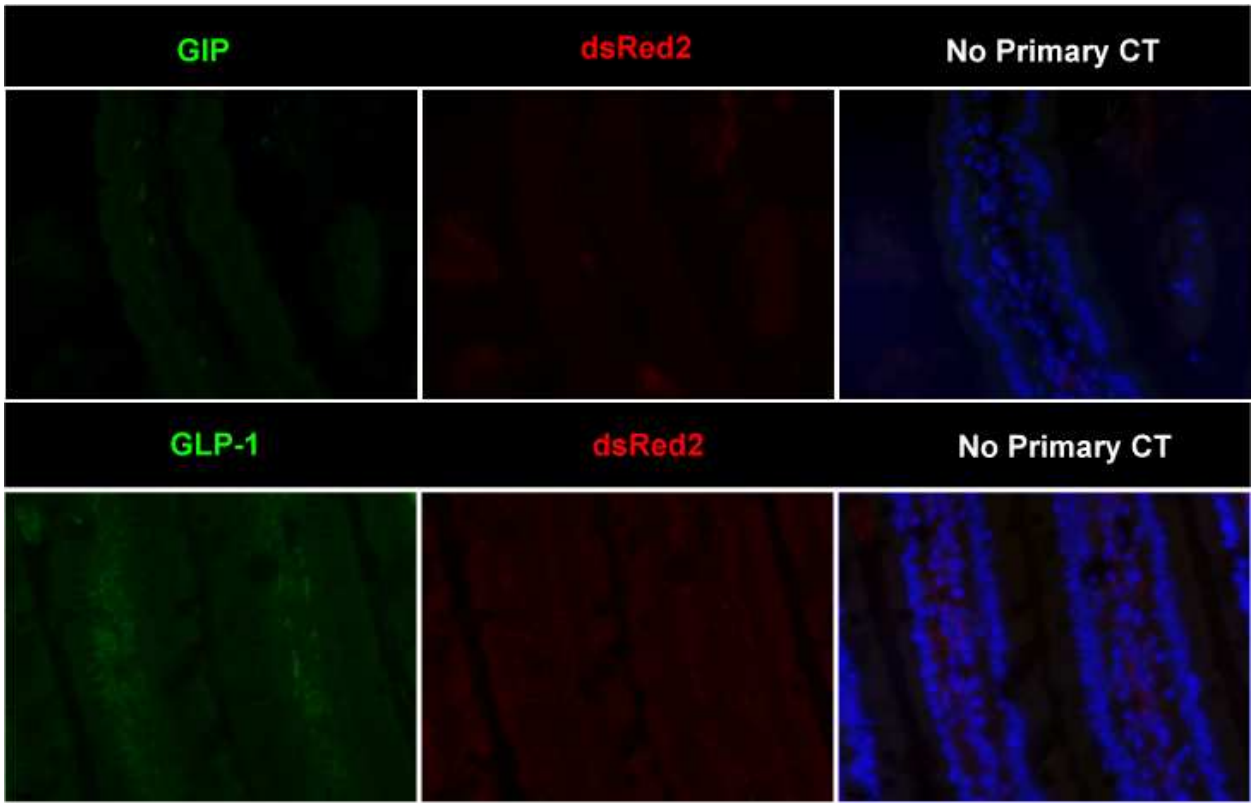


Figure 17. Background intensity of co-staining for GIP/dsRed2 and GLP-1/dsRed2. An example field of view of a section of intestinal tissue that was processed in the same way as stained sections with the exception that blocking solution was used instead of the primary antibody in the overnight incubation.

Intact villi units from these transgenic animals were dissociated from the intestinal mucosa using EDTA (Figure 18A), followed by dispase digestion to disperse the isolated villi units into a single cell suspension (Figure 18B). The total cell yield per animal ranged from 15-30 million and the viability of cells in the single cell suspension assessed by trypan blue exclusion was estimated to range from 50-90%. As evident from these images, dsRed2+ cells are rare among the intestinal epithelial cell population, and the digestion methods used to create the single cell suspension did not interfere with direct visualization of dsRed2 prior to sorting. Such a rare population is reflected by the low yield of dsRed2+ events, which ranged from 0.01% - 0.07% and was observed to vary between sample preparation and individual animals (Figure 19).

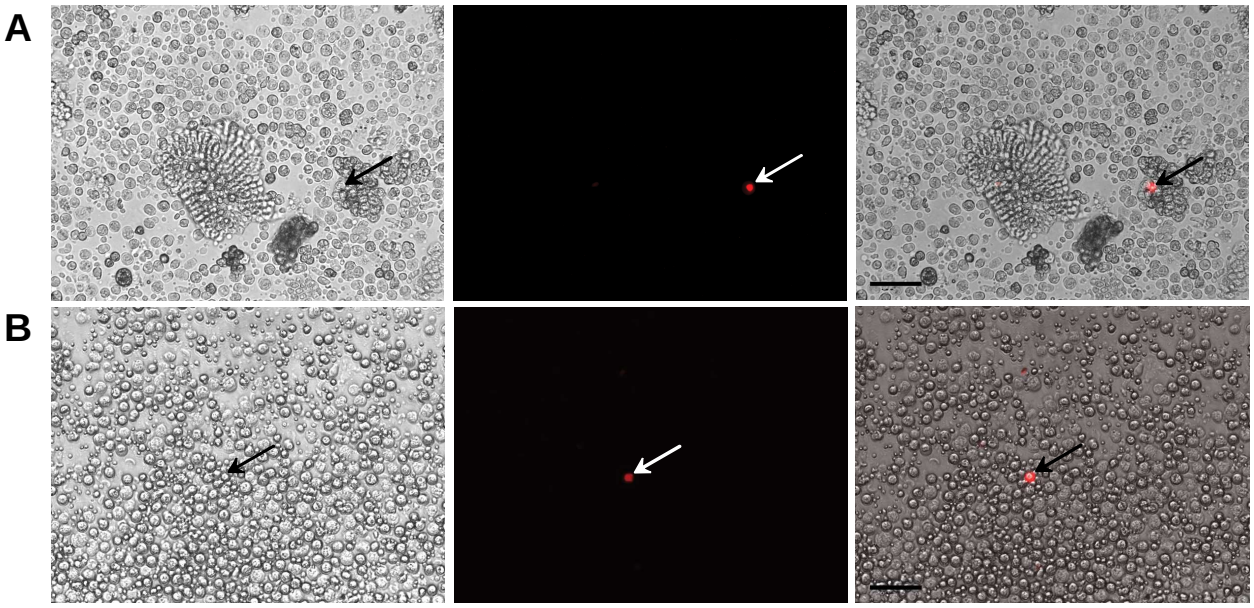


Figure 18. Direct visualization of single cell suspension of primary mouse epithelial cells before FACS. Intact villi units were dissociated from the intestinal mucosa using EDTA (A) and were dispersed into single cells (B). This digestion method did not interfere with direct visualization of dsRed2+ cells, which form a rare population amongst intestinal epithelial cells. Left = brightfield, Center = Filter set 20 (excitation: BP 546/12, beam splitter: FT 560, emission: BP 575-640), Right = overlay. Scale bar = 100 μm

Cells were sorted using the Cytopeia Influx cell sorter, which was fitted with a 561 nm laser and 580/30 filter set that allowed for optimal detection of dsRed2 (Figure 19). Gates were set based on side scatter and forward scatter to remove dead cells, debris, and doublets from the sample to be interrogated. Threshold for dsRed intensity were set based on a primary intestinal cell sample from a wild type mouse that expressed no dsRed2, prepared freshly before each sort session at the same time as the other intestinal cell samples from GIP-dsRed2 transgenic mice.

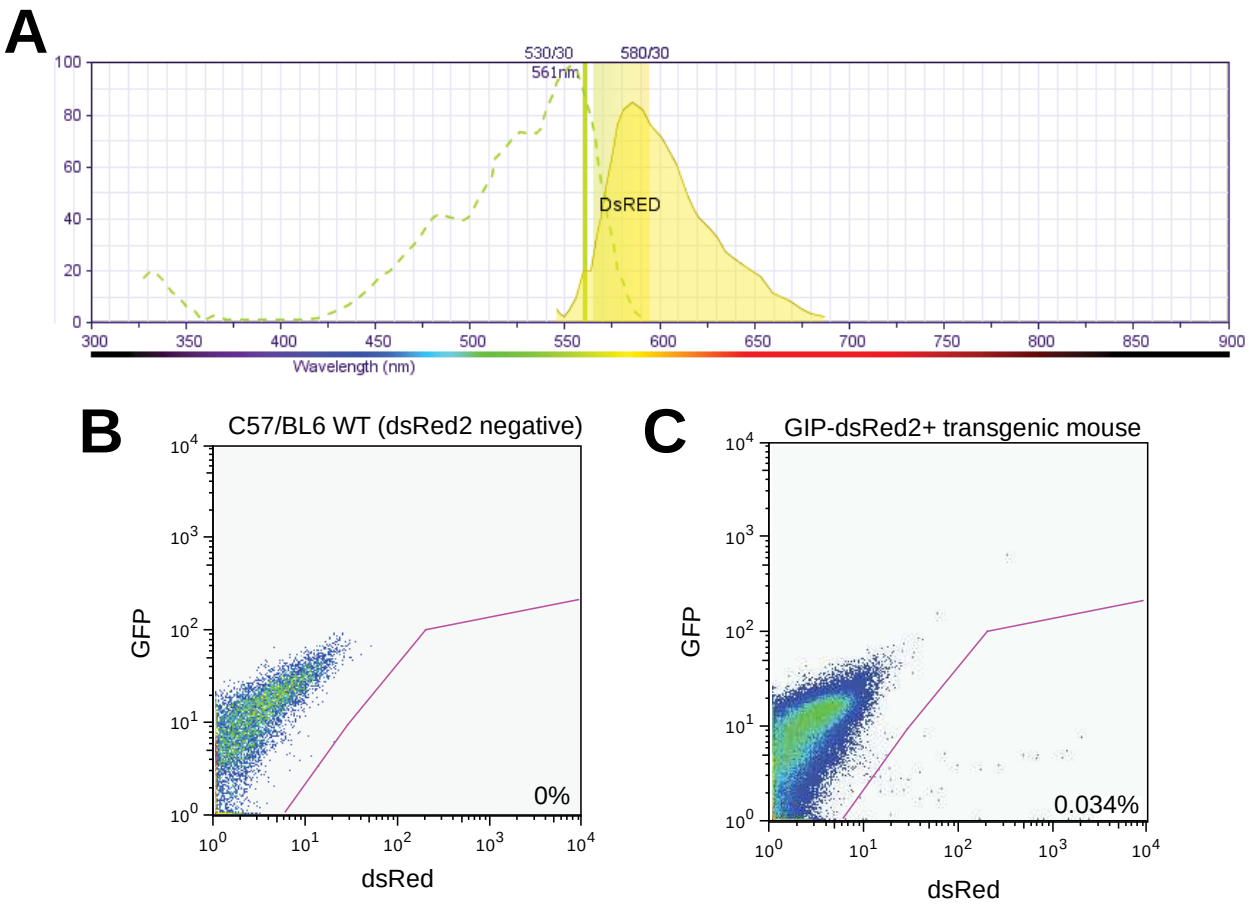


Figure 19. Detection of dsRed2+ events in a single cell suspension of intestinal epithelial cells isolated from GIP-dsRed2 transgenic mice. Cells were sorted using the Cytopeia Influx, which was fitted with a 561 nm laser and 580/30 filter to detect dsRed2 (A). Cells isolated from a C57/BL6 WT animal was used as the negative control to set up gates for sorting (B). Gating based on side scatter (a measure of granularity) and forward scatter (a measure of size) screens out dead cells, debris, and doublets. A representative plot shows a dsRed2+ frequency of 0.034% (C).

These settings on the cell sorter adequately produce a highly enriched population of dsRed2+ that expresses a gradient of dsRed2 intensity (Figure 20). Based on cell morphology, cells appear to be healthy immediately after sorting but efforts to maintain this cell population in culture proved challenging. While cells morphologically appear to survive the sorting procedure, deterioration in cell morphology was observed as early as 1 h incubation at 37°C (regardless of atmospheric or low O_2) and fail to attach when plated to surfaces coated with rat tail collagen, collagen I & IV, fibronectin, laminin, poly-d-lysine, and poly-l-lysine after an overnight incubation at 37°C (regardless of atmospheric or low O_2). Minimum medium volumes were used to promote cell attachment, and in some cases, isolated K-cells were sorted directly onto the culture surface into a droplet of culture medium and allowed to incubate 5 h in a humidified chamber before additional medium was added prior to the overnight incubation. Cells plated to the proprietary Matrigel substrate from BD Biosciences were immobilized on the culture surface as little as 5 h following plating, but the lack of cell flattening and extension of cell processes characteristic of cells in attachment cultures over 3-5 days of culturing did not suggest successful cell adherence. Despite these challenges in culturing isolated K-cells, the healthy morphology of these cells (Figure 20) suggest that high quality RNA could be isolated from them. When cells were sorted directly into RNAlater, 80-140 ng of RNA can be isolated from 20,000-30,000 dsRed2+ events (amounting to 4-7 ng/μl in 20 μl eluant). RNA quality measured by the RNA integrity number (RIN) generated by the Agilent 2100 Bioanalyzer ranged from 6.8 – 7.9, which is sufficient for proceeding with microarray (Figure 21).

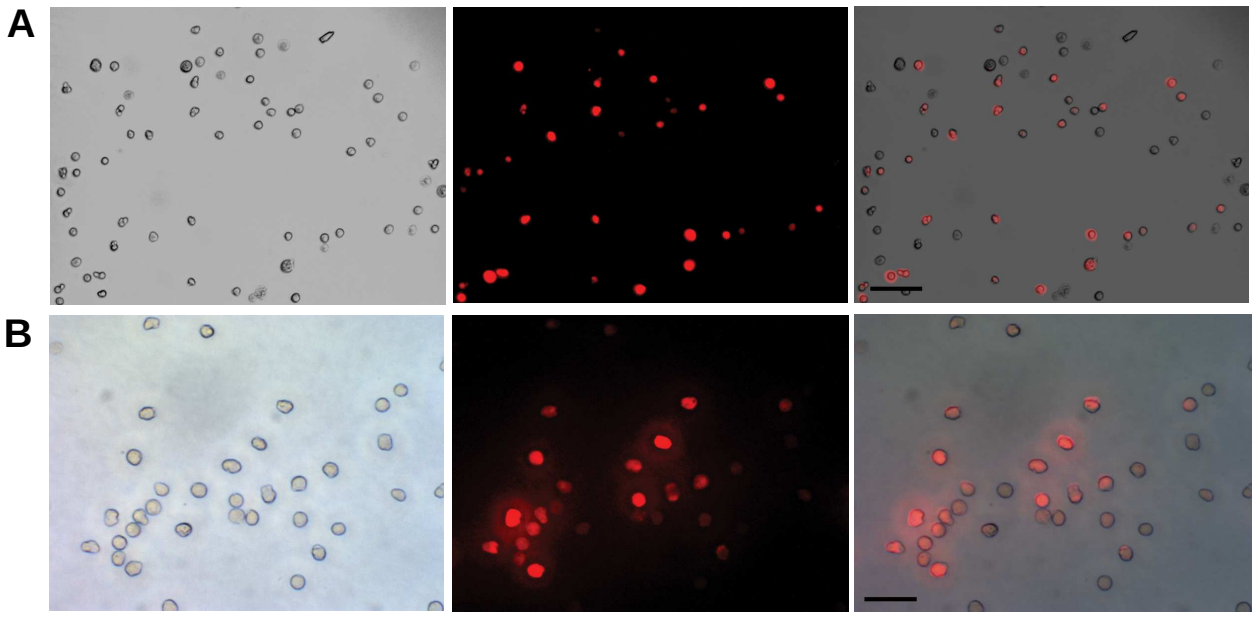


Figure 20. Primary red fluorescent GIP-expressing cells from GIP-dsRed2+ transgenic mice after FACS. The positive fraction purified by FACS represents a population of cells expressing dsRed2 in a gradient of intensities. Left = brightfield, Center = Filter set 20 (excitation: BP 546/12, beam splitter: FT 560, emission: BP 575-640), Right = overlay. (A) scale bar = 100 μm, (B) scale bar = 50 μm.

Due to variances between sample preparations, typically at least three replicates of RNA are prepared for each microarray experiment. However, isolation of sufficient cells for RNA isolation became increasingly difficult as cells isolated from GIP-dsRed2 animals that genotyped positive for the GIP-dsRed2 transgene began expressing fewer dsRed2+ cells. Initially, the lower percentage of dsRed2+ events detected by FACS was assumed to be due to variations in sample preparation and individual variation between mice. However, several different trials using transgenic mice from different litters consistently produced abnormally low yields. One common feature between the low dsRed2-expressing animals was that all were descendants of the same breeding pair, L16. At the time this observation was made, the GIP-dsRed2 line was solely being maintained by breeders (L19 and L20) that were offspring of L16.

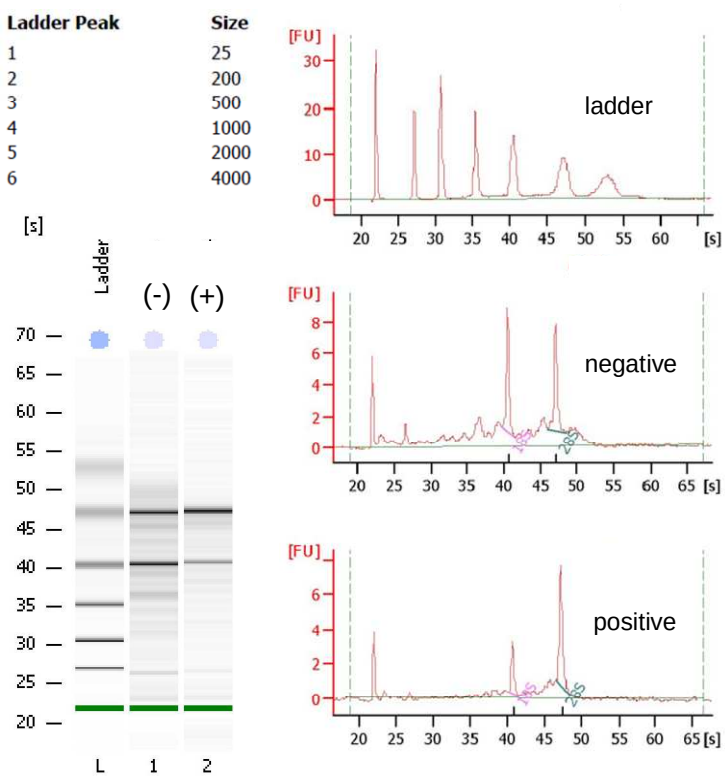


Figure 21. RNA quality check by Agilent 2100 Bioanalyzer indicated that sample quality was sufficient for proceeding with microarray analysis. The quality of multiple preps consisting of pooled dsRed2+ fractions from 3-5 animals was measured by the RNA integrity number (RIN) generated by the Agilent 2100 Bioanalyzer. RIN values were ranged from 6.8 – 7.9, which was sufficient for microarray analysis. A representative plot generated by the Agilent 2100 Bioanalyzer is presented for a dsRed2- and dsRed2+ preparation from GIP-dsRed2 transgenic mice on a C57/BL6 background.

In order to determine whether dsRed2 expression was reduced in the process of breeding, the intestine of one offspring from each of L19 (#300) and L20 (#346) were harvested for immunohistochemcial analysis. The number of GIP immunoreactive cells that were dsRed2 immunoreactive was counted in these tissues and compared to intestinal tissue from an early litter (#88 from L9). As Figure 22 demonstrates, the percent of GIP immunoreactive cells that were dsRed2 immunoreactive had decreased from approximately 60% in the L9 animal to under 20% in recent offspring. This was a matter for concern since the RNA yields from 20,000 – 30,000 dsRed2+ events were already close to the minimum amounts required for microarray analysis. The decrease of dsRed2 expression in GIP-expressing cells by 40% means that almost

twice the number of input cells is required to obtain the same number of red events in the positive fraction. Although the supply of cells is abundant (20 -50 million cells per animal), an elongated sort time from the current 5-7 h would make obtaining sufficient cells for a microarray RNA sample highly inefficient. Furthermore, a longer time required for sorting meant that cells must stay necessarily longer in a single cell suspension, which may decrease the overall health of the cells prior to sorting and could result in poor RNA quality.

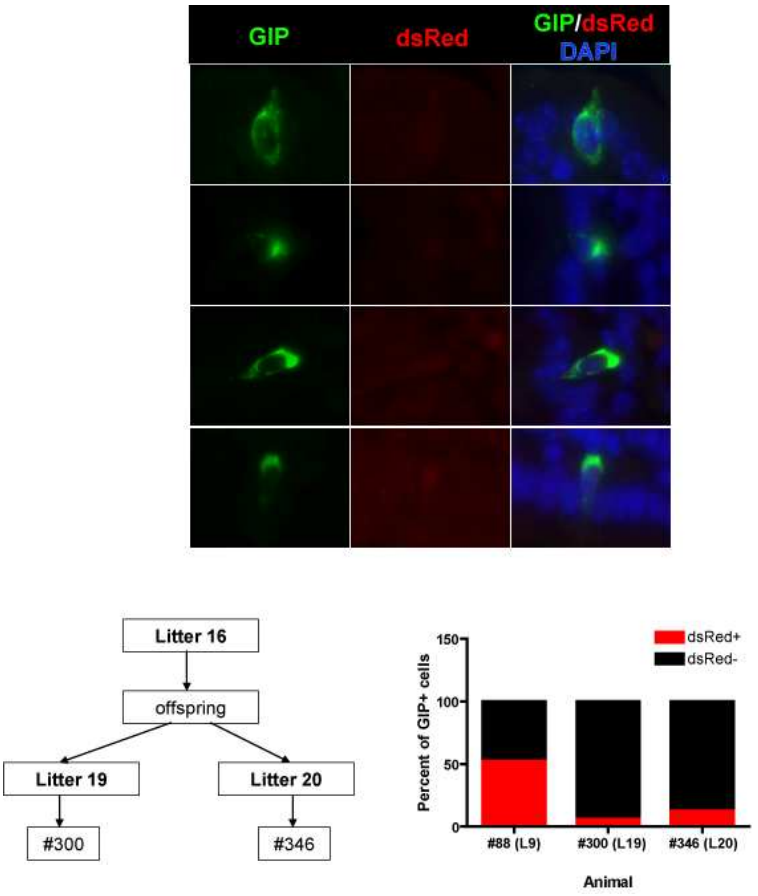


Figure 22. Decreased expression of dsRed2 in the intestine of recent GIP-dsRed2 offspring compared to an offspring from an earlier litter. The percentage of GIP immunoreactive cells that were dsRed2 immunoreactive decreased from 60% in an animal from L9 to under 20% in animals from L19 and L20.

Since the reason for decreased dsRed2 expression is unknown, but appeared to be related with a breeding event that occurred in the offspring of L16 and further propagated to the remaining transgenic animals in the GIP-dsRed2 line, re-establishment of the line with an animal from an earlier litter may be necessary to restore the yields attained to this point. In the interest of time, this avenue was not pursued, but GIP-dsRed2 animals bred on a B6/NOD background were readily available and cell isolation from these animals yielded a 154 ng and 500 ng of RNA from 35,000 dsRed2+ and 100,000 dsRed2- events, respectively (Figure 23).

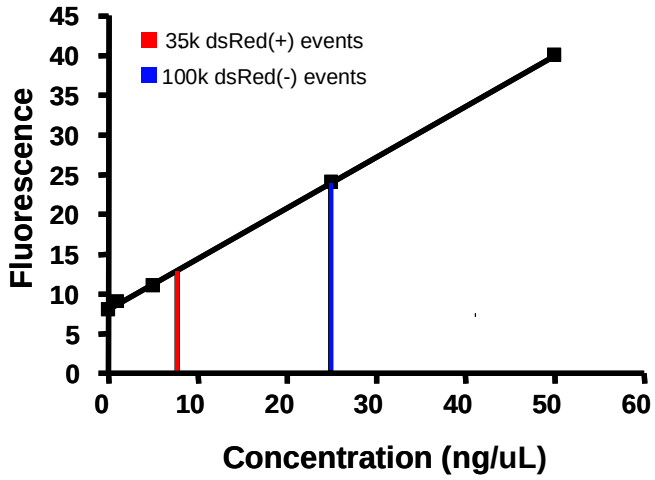


Figure 23. Isolation of RNA from dsRed2+ GIP-expressing cells from GIP-dsRed2 mice on the B6/NOD background. Typical proportions of dsRed2+ events were re-established when cells were isolated from GIP-dsRed2 transgenic animals bred on a B6/NOD background that originated from a GIP-dsRed2 C57/BL6 breeder. RNA was eluted in 20 μL elution buffer.

DISCUSSION

Identification of *Cis*-Regulatory Elements at the Distal GIP Promoter

Genes can generally be categorized into two classes based on how they are transcriptionally regulated. Genes with promoters which contain a TATA box are characterized by transcriptional initiation at a well-defined site involving recruitment of RNA Polymerase II and general transcription factors which form the pre-initiation complex, while transcription initiation in genes with promoters rich in CpG islands tend to display more plasticity and variation (162). The rat GIP gene belongs to the TATA-box containing class of genes, and thus, transcriptional regulation of the GIP gene has to date focussed largely on the proximal promoter (-193 bp to +19 bp), which is situated immediately upstream of the core promoter region at which the initiation complex is formed. While distal promoters upstream of -200 bp tend to play a lesser role in transcription regulation of TATA-box containing genes, the potential for these upstream sequences to influence promoter activity is not completely non-existent. Transgenic mouse studies in which reporter genes are driven by different lengths of the GIP promoter (40, 80, 146), suggest that the distal GIP promoter may contribute to gene regulation, at least in part, to determining the tissue-specificity of GIP expression.

As consensus sequences for *cis*-regulatory elements tend to be conserved between species, a candidate distal region spanning approximately -2500 bp to -1500 bp that demonstrates high homology between the 5'-upstream sequences of the GIP gene from nine species was chosen for analysis on the rat GIP promoter. Within this region, one putative binding site was identified for Pax6 (-2802 bp to -2797 bp) and another putative binding site was identified for Pdx1 (-2742 bp to -2737 bp), which show 60% and 100% to the Pax6 and Pdx1 consensus sequences, respectively (Figure 12). Previous studies in cell adhesion molecule genes (163, 164)

show that elements with even less than 50% of homology with the Pax6 consensus sequence are able to recruit Pax6. Furthermore, the putative Pax6 binding sites corresponding to the rat putative binding site in eight other species displayed 50-60% homology with the Pax6 consensus sequence (Figure 12). Meanwhile, although the putative Pdx1 binding site in rat was an exact match to the Pdx1 consensus sequence as well as to the sequence in the proximal rat GIP promoter shown previously to bind Pdx1 and Isl (135, 144), the putative Pdx1 binding sites corresponding to the rat putative binding site ranged between 33-100% homology with the Pdx1 consensus sequence, suggesting that conclusions generated about the Pax6 putative binding site on the rat GIP promoter may be more generalizable to other species than conclusions drawn about the Pdx1 putative binding site.

Previous studies in the proximal human and rat GIP promoter have demonstrated that Pax6 binds in the region between -193 bp and -138 bp in the human GIP promoter (132) and Pdx1 binds at -156 bp and -151 bp in the rat GIP promoter (135). The human GIP promoter study reported that -210 bp of the GIP promoter produced maximum promoter activity of the constructs tested, corresponding to approximately 55-fold of the activity generated by the promoterless luciferase reporter construct, pGL4 (132). Furthermore, truncation of the promoter driving a luciferase reporter from the distal end of the promoter resulted in a 90% drop in activity between -184 and -145 (132). Similarly, the minimum promoter length required for maximum activity of 200-fold the promoterless luciferase construct, pGL2, was reported to be -193 bp in the rat promoter study, and a mutation to the Pdx1 binding site from -156 bp to -151 bp in this construct resulted in a 85% drop in promoter activity (135). The present distal regulatory element data indicates that mutation to the putative Pax6 binding site reduced promoter activity by 65% while mutation to the putative Pdx1 binding site reduced promoter activity by 40%. Given that removal of proximal regulatory elements already accounts for the majority of GIP

promoter activity observed in these two studies, it is important to clarify the biological significance of the present distal regulatory element data.

The reference promoter from which changes in promoter activity was measured between each of the proximal promoter studies and the current study vary in length. Both of the reference promoters used in the proximal promoter studies were relatively short (~ 200 bp) and highly active (55-fold background activity for human promoter and 200-fold for rat promoter) compared to the reference promoter used in the current study (~3 kb), which displayed 20- to 25-fold background activity (Figure 6 and 7). It is not usual that the fold activity from background is higher for reporter constructs driven by shorter lengths of the promoters than longer promoters, since repressor elements may be present in the more distal regions. This has been demonstrated previously by Fujita *et al.* (132) for the human GIP promoter, in which -1.0 kb, -1.9 kb, and -2.9 kb of the promoter all have progressively lower activity than the -210 bp construct. The location of repressors may not be located far upstream of the proximal promoters, since Someya *et al.* (142) demonstrated, also in the human GIP promoter, that a truncation of -258 bp to -180 bp of the promoter resulted in augmentation of promoter activity. Repressor and enhancer elements are likely interspersed with each other, since progressive shortening of a -1200 bp promoter in the Someya *et al.* (142) study showed first a decrease in promoter activity prior to augmentation. Overall, the percentages expressing decreases in GIP promoter activity cannot be directly compared to each other unless the reference promoter is identical, but useful information may still be gleaned from different studies pertaining to the presence of an enhancer or repressor elements along a given segment of the promoter.

Despite differences in the lengths of the reference promoters, it does appear that the promoter constructs used in the current study tend to have low activity. As described in the Materials and Methods section, the promoter segment used to generate the pGL2574 plasmid

was identical to the -2.5 kb promoter which is described in the text of Boylan *et al.*'s manuscript (143) as having "similar rates of transcription" as the activity mediated by the longest -943 bp promoter presented in the data. However, the activity of this promoter in the current studies displayed only an 8-fold increase from background activity (Figure 6). The luciferase reporter used in the current study is based on the pGL4.10[luc2] backbone, while the Boylan *et al.* (143) study used reporters based on the pGL2-basic backbone. Both constructs are manufactured by Promega, and according to the technical specifications from the company, the pGL4.10[luc2] backbone should, if anything be expressed even better in mammalian systems (165). Two other factors may account for the low expression of the reporters in this study. Dedifferentiation of STC-1 cells, which has been shown to occur (166), may account for decreased GIP promoter-driven reporter activity and differences in transfection amount and efficiency may also account for lowered expression. These same factors may also explain why the -2.5 kb plasmid used in the Boylan *et al.* (143) study demonstrated only a 4-fold increase from background activity in a later paper published by the same group (145).

What can be said of the distal promoter elements identified in the current study is that the putative Pax6 binding site appears to be more important for rat GIP promoter activity than the putative Pdx1 site, since mutation of both sites simultaneously reduced GIP promoter to the same level as mutating the putative Pax6 site alone (Figure 13). Overall, whether these distal elements identified interact with the proximal promoter elements or the nature of this interaction will require further investigation, such as, by assessing how promoter activity differs amongst a promoter in which the proximal element has been mutated, a promoter in which the distal element has been mutated, a promoter in which both proximal and distal elements together are mutated, and a promoter in which both proximal and distal elements are left intact.

Pax6 and Pdx1 as *Trans*-Regulatory Elements at the Distal Promoter

Mutant mice deficient for Pax6 (124) and Pdx1 (127, 135) have suggested that both these transcription factors are important to determining GIP-expressing cell numbers. However, only Pdx1 appears to be necessary for the maintenance of GIP-expression, since examples of GIP-expressing cells lacking Pax6, but no examples of GIP-expressing cells lacking Pdx1 have been shown in the adult mouse intestine (132, 135). The ability of Pdx1 overexpression alone to differentiate immature rat intestinal IEC6 cells into GIP-expressing cells (140) also showed that Pdx1 is sufficient for GIP expression. Therefore, Pax6 may behave in the capacity of GATA4 (145) in regards to GIP expression in that it modulates GIP expression levels but is not required for expression. Another hypothesis is that Pax6 may only be expressed at certain intervals of the life span of K-cells.

Several lines of evidence suggest that these two transcription factors may be the transcription factors that bind the *cis*-regulatory elements identified in the distal rat GIP promoter. Firstly, the consensus sequences of these two factors are located in these elements (Figure 12). Secondly, incubation of biotinylated probes containing these consensus sites with nuclear extract from STC-1 cells and anti-sera against Pax6 and Pdx1 show the presence of super-shift bands (Figure 9-11). Finally, endogenous binding of Pax6 and Pdx1 in STC-1 cells was demonstrated using chromatin immunoprecipitation (Figure 14). Additional experiments to provide even more evidence that Pax6 and Pdx1 act on the distal regulatory elements at -2802 bp to -2797 bp and -2742 bp to -2737 bp of the rat GIP promoter could consist of co-transfection experiments of rat GIP reporter plasmids with Pax6 and Pdx1 expression plasmids. Of interest is whether overexpression of Pax6 and Pdx1 in varying amounts would produce a dose-dependent augmentation of GIP promoter activity. The dose-dependency of Pax6 and Pdx1 has previously been shown in similar overexpression studies of the rat and human GIP promoters (132, 145).

Furthermore, the effects of overexpression of Pax6 and Pdx1 together could be determined to assess whether these factors reciprocally enhance the effects of each other. Concerted binding of Pax6 and Pdx1 for optimal somatostatin activity has been demonstrated at two regulatory elements found -450 bp upstream of the somatostatin gene (167). If Pax6 and Pdx1 indeed act on the *cis*-regulatory elements identified in the present study, the effects of Pax6 and Pdx1 are expected to be abolished or reduced when the same co-transfection experiment is carried out with the pGL3075pax6mut, pGL3075pdx1mut, or pGL3075DistalDbMut constructs.

Other transcription factors in addition to Pax6 and Pdx1 were tested for binding to *cis*-regulatory elements (Figures 9-11). In probes 3 and 4, which span the distal element studied, binding of GATA4 to probe 3 (Figure 9B) and binding of GATA6 to probe 4 (Figure 9C) was observed. However, this result could not be replicated (Figure 11). In both cases, no GATA binding sites ((A/T)GATA(A/G)) were predicted on the probe, but the presence of super-shifted bands may indicate complex formation with Pax6 and Pdx1. Should an interaction between Pax6 and Pdx1 be suggested by co-transfection studies such as the ones described above, the transcription factors which participate in complex formation to mediate this interaction may be of interest. The occupancy of transcription factors on their specific binding sites have been estimated to range from milliseconds to ~100 s in fluorescence recovery after photobleaching experiments (109), and such a dynamic interaction may be one reason why the observation of GATA4 and GATA6 binding to their respective probes was not replicated. A more representative view of whether GATA4 and GATA6 participates in complex formation at the *cis*-regulatory elements represented by probes 3 and 4 could be obtained by performing more replicates of the EMSA experiment as well as demonstrating the binding of these factors using other techniques.

Evaluation of the STC-1 Model System

For reasons already outlined in the Introduction, utilizing the STC-1 tumour cell line as a model has some limitations, including the uncertainty of whether these results can be extrapolated to the *in vivo* system. The sensitivity of transcription regulation to the cellular milieu of other genes and signalling pathways suggest that studies of transcription regulation carried out in tumour cell lines, which have undergone altered gene expression and signalling in the process of immortalization (168), suggests that conclusions may not be physiologically relevant. While the current studies do not contribute any additional knowledge to whether or which part of the rat distal GIP promoter contributes to restricting GIP expression in specific tissues, an attempt to assess the role of the high homology distal region from -2500 bp to -1500 bp (Figure 4) in other species was made. Reporter plasmids driven by varying lengths of the human and pig GIP promoter were compared against the activity of the rat GIP promoter in STC-1 cells (Figure 15). It is difficult to assess whether the divergent effects of removing segments of the highly conserved distal region identified in Figure 4 in the different species are physiological. All three promoter constructs were tested in a rodent cell line, such that species-specific factors influencing the consequences of truncating regulatory elements within this region may be misrepresented. There is also the possibility that the function of well-conserved sequence may differ between species, since structural conservation may not necessarily equate to functional conservation across species (169).

Quality of Isolated Primary GIP-expressing Cell Preparations

To address the limitations of performing studies in transcriptional regulation in tumour cell lines, Parker and colleagues devised a new model for studying K-cell biology (42) and used this model to assess glucose and nutrient sensing using a candidate gene approach. However,

one of the limitations of a candidate gene approach, like the one used by Parker and colleagues and the present strategy of identifying *trans*-regulatory elements discussed so far in this thesis, is that it is only possible to screen already-identified genes and transcription factors. The potential of discovering new transcription factors not previously known to be expressed in K-cells is low, and therefore, the process of selecting genes and transcription factors as candidates is subject to experimenter bias. RNA of sufficient quantity and quality was isolated for quantitative RT-PCR using the fluorescence-based approach described by Parker *et al.* (42), and the studies described in this thesis demonstrate that RNA could also be isolated using a similar strategy of sufficient quantity and quality for microarray. The minimum requirement of 25 ng total RNA per 11 μL of water (requires amplification prior to cDNA synthesis) was met by isolating RNA from 20,000 – 35,000 dsRed2+ cells. This high-throughput and unbiased alternative to screening for genes that are expressed in higher levels in GIP-expressing cells over non-GIP-expressing cells may expand the list of candidate factors that may act at *cis*-regulatory elements found to be important for GIP expression. Additionally, microarray data may also provide a starting point for investigating other aspects of the K-cell life cycle and potential as a gene therapy target.

Methodological Limitations to Sample Purity

The modified Weiser's protocol (see Materials and Methods) removed epithelial cells from both the villi and crypts into a single cell suspension, which encompasses all the regions where GIP-expressing cells are found in the intestine (35). Critical to the purity of the sample obtained by this method was the assumption that dsRed2 is specifically expressed in GIP-expressing cells. This assumption was evaluated by immunohistochemical analysis of tissue from GIP-dsRed2 animals. As Figure 16D demonstrates, 94% of dsRed2 immunoreactive cells were simultaneously GIP immunoreactive. The lack of GIP immunoreactivity in the remaining 6% of dsRed2+ cells does not necessarily represent non-specific expression of dsRed2 in the

GIP-dsRed2 animals, since the GIP antibody used (170) recognizes an epitope in the C-terminus region of GIP. Alternative processing of GIP to shorter biologically active versions removes this epitope, rendering these forms of GIP invisible to detection by this antiserum. Characterization of this C-terminus antibody in mouse intestine (unpublished data, A. Asadi) shows that the antibody detects upwards of 90% of GIP immunoreactive cells detected by a second antibody that detects both GIP_{1-42} and shorter variants. This finding is consistent with the proportion of dsRed2+ cells immunoreactive for GIP using the C-terminus antibody. An additional hypothesis accounting for the lack of GIP immunoreactivity in some dsRed2 immunoreactive cells is that GIP expression may be downregulated in GIP-expressing cells near the end of their life cycle. Meanwhile, the dsRed2 protein has a long half-life of 4.6 days (171), which could still be detected as immunoreactivity after the gene expression machinery has been downregulated. Given the well-established migration of GIP-expressing cells from the crypt to the tip of the villus as they age, this hypothesis can be tested by seeing if the cells that are only immunoreactive for dsRed2 are those located close to the extrusion zone at the tip of the villi.

A second technical limitation to purity is inherent in FACS. FACS operates on the basis that events detected in a single cell stream are subsequently deflected by electrically-charged plates to collection tubes. The purity of the sample depends, therefore, on how well the input sample is dispersed into a single cell suspension and fluidics parameters which affect the stream of cells to be interrogated and how they are deflected into the collection tube. Given the rarity of the cell samples, a representative purity check was performed once to check the amount of error caused by these factors, after which the settings for cell isolation and fluidics were kept constant for subsequent preparations. The purity check consisted of passing a small sample from the dsRed2+ fraction back through the flow cytometer for fluorescence detection. Keeping all gating the same from the initial sort, the purity of the sample was determined to be 88%.

Improvements to Current Isolation Procedure

The range of dsRed2+ events recorded by FACS from 0.01 – 0.07% is consistent with the finding that GIP-expressing cells are one of at least 15 enteroendocrine cells, which all together make up < 1% of the intestinal epithelium (35). Nevertheless, Figure 22 indicates that dsRed2 is not expressed in every GIP-expressing cell. Given the closeness of the dsRed2+ cell yield to the proportion of K-cells expected to be in the intestinal epithelium population, this observation is likely due to a discrepancy between dsRed2+ cells and dsRed2 immunoreactive cells, which are dependent on the dsRed2 antibody used. It is also well known that dsRed2, despite having a faster maturation rate than its predecessor, dsRed, still requires approximately 20 h to attain optimal fluorescence (172, 173). Given the rapid turnover of enteroendocrine cells of 3-5 days (35), the population of GIP immunoreactive cells that are not immunoreactive for dsRed2 (Figure 22) could represent newly-differentiated GIP-expressing cells in which insufficient time following dsRed2 protein synthesis has elapsed for dsRed2 maturation. A third hypothesis that may address the lack of dsRed2+ in some GIP-expressing cells relies on the heterogeneity within the K-cell population. The existence of a subpopulation of GIP-expressing cells that express GLP-1 (29, 30, 36), and xenin (174), is well-documented. Whether the requirements for GIP gene expression differ among these subpopulations are not known, and thus, the length of the GIP promoter driving dsRed2 expression may be insufficient for activation in some of these K-cells.

As the current amount of RNA isolated per preparation falls quite close to the minimum required for microarray with RNA amplification prior to cDNA synthesis, improvements in the isolation procedure which would lead to a higher cell yield is desired. Currently, the rate-limiting step in sample preparation occurs at the FACS, and enrichment of the input sample prior to sorting could significantly increase the proportion of dsRed2+ cells within the interrogated

population and thus, the overall yield. In order to perform pre-FACS enrichment, positive markers that recognize a cell population that includes K-cells are required. Alternatively, negative markers that are specific for even some cells that do not express GIP could decrease the number of events that require interrogation, and thus, decrease the overall FACS time significantly. Known positive-selection markers include endocrine markers, such as chromogranin A and B (175, 176) and negative-selection markers include Lgr5 (177), an intestinal stem cell-specific marker, and enterocytin (178), an enterocyte-specific marker. As these are not cell-surface markers, the use of these as enrichment markers would require labelling.

One of the challenges encountered with establishing this cell isolation method was the finding that the number of dsRed2+ cells in the intestine of the GIP-dsRed2 line of mice declined over time (Figure 22). Although the mechanism of this decreased expression was not identified, the observation that reduced dsRed2+ cell numbers was seen in animals that were all descendants to the same breeding pair led to the unconfirmed speculation that decreased dsRed2 expression may be a result of decreased transgene copy numbers being passed onto the next generation in a specific breeding event. The data also suggest that strain differences in dsRed2 expression levels may occur, since isolation of cells from GIP-dsRed2 animals on the B6/NOD, which originated from the same litter in which low dsRed2 expression was first observed, resulted in a cell yield comparable to isolation from initial preparations from cells derived from GIP-dsRed2 on the C57/BL6 background (Figure 23).

Limitations to Maintaining a Primary Enriched K-cell Culture

Attempts to maintain primary isolated GIP-expressing cells in culture were unsuccessful, even though previous studies show that primary mixed intestinal cultures can be kept for as long

as several weeks (179, 180, 181, 182). Generally, cell attachment to the basement membrane via extracellular matrix molecules is required for cell survival (182). The disruption of these extracellular matrix interactions can trigger apoptosis in a process called anoikis (182). Unfortunately, a basic requirement for FACS is the generation of an input single cell solution. Therefore, it is likely that the process of preparing primary intestinal epithelial cells for FACS may destroy the ability of these cells to survive in culture. Indeed, there has not yet been a report of successful culturing of primary intestinal epithelial cells that have been subjected to flow cytometry. However, the isolation of K-cells by fluorescent labelling still has its merits in that terminal studies can still be done on highly purified populations of K-cells.

Potential Issues that May Arise in Data Interpretation

As mentioned earlier, GIP-expressing cells do not represent a homogenous population of cells within the enteroendocrine cell population. In addition to subpopulations which co-express other hormones (29, 30, 36, 174) post-translational cleavage processing of preproGIP produces a number of biologically active variants (18). The heterogeneity in the GIP-expressing cells naturally means that any cell isolation method based on the ability of cells to express GIP, such as the fluorescence-based method described in this thesis, will produced an enriched population reflecting the same diversity of GIP-expressing cells. Figure 16 demonstrates that a subpopulation of dsRed2-expressing cells expresses GLP-1, consistent with reports that a majority of GIP-expressing cells also express GLP-1 (29, 30, 36). Distinguishing features between GIP-expressing cells that only express GIP and those that express GLP-1 in addition to GIP have not yet been identified. The functional significance of enteroendcrine cells that only express one incretin hormone versus those that express both remain to be clarified, and even whether these cells represent samples of distinct populations or whether they represent a single population at different stages of the life cycle has not been studied.

In the current cell isolation method for preparing samples for microarray, the experimental sample would consist of a mixture of all these subpopulations of GIP-expressing cells, while the criterion for highly expressed genes, set as those genes which are expressed above a certain threshold more in the dsRed2+ population than genes in the dsRed2- cell population treats the dsRed2+ as a single homogenous population. Therefore, while microarray data may suggest that a particular gene is highly expressed in the K-cells as a pool, potentially non-random variations within the K-cell population in expression level for this gene may exist. The availability of a transgenic animal which expresses the Venus yellow-fluorescent protein in cells that specifically express GLP-1 (161) may provide the first steps to distinguishing between different subpopulation of GIP-expressing cells. Breeding of the GIP-dsRed2 transgenic line with this Venus-expressing line could provide a useful model for studying the differences between GIP-expressing cells that express GLP-1 and those that do not, since each of the GIP-only, GIP- and GLP-1, and GLP-1-only expressing cells could then be detected using a different FACS gate.

Another consideration that would highly impact the interpretation of microarray data is the negative control, which provides the reference for comparing gene expression levels. At present, negative samples consist of RNA isolated from dsRed2- fraction from input cells isolated from GIP-dsRed2 animals, which means that it is a mixed population of enterocytes, goblet cells, and enteroendocrine cells that do not express GIP. The use of so many different cell types in the negative control generates noise, which would reduce the sensitivity of detection of genes that are only moderately more in GIP-expressing cells than other cells. For example, if a certain gene is expressed 5-fold more in GIP-expressing cells than any other intestinal epithelial cell, the fold increase of gene expression when the negative control consists of three different non-GIP-expressing cell types, may only be detected as a 0.4 fold increase. The heterogeneity of

the negative control increases the stringency for the specificity of the gene marker for GIP-expressing cells, which may be an advantage for isolating candidate genes that have a functional role specific to K-cells.

Conclusions and Future Directions

Two *cis*-regulatory elements were identified in a well-conserved distal region of the rat GIP promoter at -2802 bp to -2797 bp and at -2742 bp to -2737 bp relative to the transcription start site. Mutational analysis of these two elements showed that these elements contributed to 40-65% of GIP promoter activity. Pax6 and Pdx1 were identified as the *trans*-regulatory elements that bind at these sites. In particular, Pax6 binding to the more distal *cis*-regulatory element appeared to be the functionally more important of the two sites analyzed, and conservation of the Pax6 consensus sequence across eight different species also appear to be higher than for the other site. The focus in the present analysis have been exclusively on the 5' upstream promoter, based on the categorization of the GIP gene as a TATA-box containing gene, but intronic sequences may also regulate transcriptional activity in TATA-box containing genes. Intronic regulation, for example, has already been demonstrated for the proglucagon gene (183). Furthermore, single-nucleotide polymorphisms of the transcription factor TCF-4 which are associated with increased risk of developing diabetes have almost occurred exclusively in intronic regions (184). Recognizing that transcriptional behaviour observed in STC-1 cells may not reflect physiological function, efforts to isolated primary GIP-expressing cells were described, although primary cultures of isolated K-cells did not survive. However, the fluorescence-based isolation method for GIP-expressing cells would allow for non-biased screening for GIP-expressing cell-specific markers, including transcription factors, which contribute to regulation of GIP gene transcription.

REFERENCES

1. Papaspyros N (1952) *The History of Diabetes* (Robert Stockwell Ltd., London).

2. International Diabetes Federation Diabetes Atlas. [http://www.diabetesatlas.org, accessed Nov. 15, 2009].

3. Ryan JG (2009) Cost and policy implications from the increasing prevalence of obesity and diabetes mellitus. *Gend. Med.* 6:86-108.

4. Bayliss W, Starling E (1902) The Mechanism of Pancreatic Secretion. *J. Physiol.* 28:325-353.

5. Moore B (1906) On the Treatment of Diabetes Mellitus by Acid Extract of Duodenal Mucous Membrane. *Biochem. J.* 1:28-38.

6. Creutzfeldt W (2005) The [pre-] history of the incretin concept. *Regul. Pept.* 128:87-91.

7. Rehfeld JF (2004) A centenary of gastrointestinal endocrinology. *Horm. Metab. Res.* 36:735-741.

8. Unger R, Eisentraut A (1969) Entero-insular Axis. *Arch. Intern. Med.* 123:261-266.

9. McIntyre N, Holdsworth C, Turner D (1964) New interpretation of oral glucose tolerance. *Lancet* 2:20-21.

10. Perley M, Kipnis D (1967) Plasma Insulin Responses to Oral and Intravenous Glucose Studies in Normal and Diabetic Subjects. *J. Clin. Invest.* 46:1954-1962.

11. Nauck M, Bartels E, Orskov C, Ebert R, Creutzfeldt W (1993) Additive insulinotropic effects of exogenous synthetic human gastric inhibitory polypeptide and glucagon-like-peptide-1-(7-36) amide infused at near-physiological insulinotropic hormone and glucose concentrations. *J. Clin. Endocrinol. Metab.* 76:912-917.

12. Nauck M et al. (1986) Incretin Effects of Increasing Glucose Loads in Man Calculated from Venous Insulin and C-Peptide Responses. *J. Clin. Endocrinol. Metab.* 63:492-498.

13. Dupre J, Beck J (1966) Stimulation of release of insulin by an extract of intestinal mucosa. *Diabetes* 15:555-559.

14. Brown J, Mutt V, Pederson R (1970) Further Purification of a Polypeptide Demonstrating Enterogastrone Activity. *J. Physiol.* 209:57-64.

15. Brown J, Pederson R (1976) GI hormones and insulin secretion. *Endocrinol. Proc. 5th Int. Congr. Endocrinol.* 2:568-570.

16. Brown J (1987) Role of gastric inhibitory polypeptide in regulation of insulin release. *Front. Horm. Res.* 16:157-166.

17. Dupre J, Ross S, Watson D, Brown J (1973) Stimulation of Insulin Secretion By Gastric Inhibitory Polypeptide in Man. *J. Clin. Endocrinol. Metab.* 37:826-828.

18. McIntosh CH, Widenmaier S, Kim S (2009) Glucose-dependent Insulinotropic Polypeptide (Gastric Inhibitory Polypeptide; GIP). *Vitam. Horm.* 80:409-471.

19. Inagaki N et al. (1989) Gastric inhibitory polypeptide: structure and chromosomal localization of the human gene. *Mol. Endocrinol.* 3:1014-1021.

20. Higashimoto Y, Liddle R (1993) Isolation and Characterization of the Gene Encoding Rat Glucose-Dependent Insulinotropic Peptide. *Biochem. Biophys. Res. Commun.* 193:182-190.

21. Higashimoto Y, Opara E, Liddle R (1995) Dietary regulation of glucose-dependent insulinotropic peptide (GIP) gene expression in rat small intestine. *Comp. Biochem. Phys. C, Pharmacol. Toxicol. Endocrinol.* 110:207-214.

22. Schieldrop P et al. (1996) Isolation of a murine glucose-dependent insulinotropic polypeptide (GIP) cDNA from a tumor cell line (STC6-14) and quantification of glucose-induced increases in GIP mRNA. *Biochim. Biophys. Acta* 1308:111-113.

23. Baggio L, Drucker D (2007) Biology of incretins: GLP-1 and GIP. *Gastroenterology* 132:2131-2157.

24. Kieffer T, McIntosh C, Pederson R (1995) Degradation of glucose-dependent insulinotropic polypeptide and truncated glucagon-like peptide 1 in vitro and in vivo by dipeptidyl peptidase IV. *Endocrinology* 136:3585-3596.

25. Deacon C, Danielsen P, Klarskov L, Olesen M, Holst J (2001) Dipeptidyl peptidase IV inhibition reduces the degradation and clearance of GIP and potentiates its insulinotropic and antihyperglycemic effects in anesthetized pigs. *Diabetes* 50:1588-1597.

26. Deacon C, Nauck M, Meier J, Hucking K, Holst J (2000) Degradation of endogenous and exogenous gastric inhibitory polypeptide in healthy and in type 2 diabetic subjects as revealed using a new assay for the intact peptide. *J. Clin. Endocrinol. Metab.* 85:3575-3581.

27. Cataland S, Crockett S, Brown J, Mazzaferri E (1974) Gastric inhibitory polypeptide (GIP) stimulation by oral glucose in man. *J.Clin. Endocrinol. Metab.* 39:223-228.

28. Elliot R et al. (1993) Glucagon-like peptide-1 (7-36)amide and glucose-dependent insulinotropic polypeptide secretion in response to nutrient ingestion in man: acute post-prandial and 24-h secretion patterns. *J. Endocrinol.* 138:159-166.

29. Mortensen K, Christensen L, Holst J, Orskov C (2003) GLP-1 and GIP are colocalized in a subset of endocrine cells in the small intestine. *Regul. Pept.* 114:189-196.

30. Mortensen K, Petersen L, Orskov C (2000) Colocalization of GLP-1 and GIP in human and porcine intestine. *Ann. N.Y. Acad. Sci.* 921:469–472.

31. Polak J, Bloom S, Kuzio M, Brown J, Pearse A (1973) Cellular localization of gastric inhibitory polypeptide in the duodenum and jejunum. *Gut* 14:284–288.

32. Cho G et al. Upregulation of glucose-dependent insulinotropic polypeptide and its receptor in the retina of streptozotocin-induced diabetic rats. *Curr. Eye Res.* 25:381-388.

33. Nyberg J et al. (2005) Glucose-dependent insulinotropic polypeptide is expressed in adult hippocampus and induces progenitor cell proliferation. *J. Neurosci.* 25:1816-1825.

34. Tseng C et al. (1995) Glucose-dependent insulinotropic peptide (GIP) gene expression in the rat salivary gland. *Mol. Cell. Endocrinol.* 115:13-19.

35. Höcker M, Wiedenmann B (1998) Molecular mechanisms of enteroendocrine differentiation. *Ann. N.Y. Acad. Sci.* 859:160-174.

36. Theodorakis MJ et al. (2006) Human duodenal enteroendocrine cells: source of both incretin peptides, GLP-1 and GIP. *Am. J. Physiol. Endocrinol. Metab.* 290:E550-E559.

37. Moriya R, Shirakura T, Ito J, Mashiko S, Seo T (2009) Activation of sodium-glucose cotransporter 1 ameliorates hyperglycemia by mediating incretin secretion in mice. *Am. J. Physiol.Endocrinol. Metab.* 297:E1358-E1365.

38. Jetton TL et al. (1994) Analysis of upstream glucokinase promoter activity in transgenic mice and identification of glucokinase in rare neuroendocrine cells in the brain and gut. *J. Biol. Chem.* 269:3641-3654.

39. Moates J, Nanda S, Cissell M, Tsai M, Stein R (2003) BETA2 Activates Transcription From the Upstream Glucokinase Gene Promoter in Islet β-cells and Gut Endocrine Cells. *Diabetes* 52:403-408.

40. Cheung AT et al. (2000) Glucose-Dependent Insulin Release from Genetically Engineered K Cells. *Science* 290:1959-1962.

41. Murphy R et al. (2009) Glucokinase, the pancreatic glucose sensor, is not the gut glucose sensor. *Diabetologia* 52:154-159.

42. Parker HE, Habib AM, Rogers GJ, Gribble FM, Reimann F (2009) Nutrient-dependent secretion of glucose-dependent insulinotropic polypeptide from primary murine K cells. *Diabetologia* 52:289-298.

43. Fujita T (1991) Taste cells in the gut and on the tongue. Their common, paraneuronal features. *Physiol. Behav.* 49:883-885.

44. Höfer D, Püschel B, Drenckhahn D (1996) Taste receptor-like cells in the rat gut identified by expression of alpha-gustducin. *Proc. Natl. Acad. Sci. U. S. A.* 93:6631-6634.

45. Jang H et al. (2007) Gut-expressed gustducin and taste receptors regulate secretion of glucagon-like peptide-1. *Proc. Natl. Acad. Sci. U. S. A.* 104:15069-15074.

46. Wong G, Gannon K, Margolskee R (1996) Transduction of bitter and sweet taste by gustducin. *Nature* 381:796-800.

47. Fujita Y et al. (2009) Incretin release from gut is acutely enhanced by sugar but not by sweeteners in vivo. *Am. J. Physiol. Endocrinol. Metab.* 296:E473-9.

48. Alumets J, Håkanson R, O'Dorisio T, Sjölund K, Sundler F (1978) Is GIP a glucagon cell constituent? *Histochemistry* 58:253-257.

49. Larsson L, Moody A (1980) Glicentin and gastric inhibitory polypeptide immunoreactivity in endocrine cells of the gut and pancreas. *J. Histochem. Cytochem.* 28:925–933.

50. Wideman RD, Kieffer TJ (2009) Mining incretin hormone pathways for novel therapies. *Trends Endocrinol. Metab.* 20:280-286.

51. Musson MC et al. (2009) Expression of Glucose-Dependent-Insulinotropic Polypeptide (Gip) in the Zebrafish. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 297:R1803-R1812.

52. Irwin DM (2009) Molecular evolution of mammalian incretin hormone genes. *Regul. Pept.* 155:121-130.

53. Irwin D, Zhang T (2006) Evolution of the vertebrate glucose-dependent insulinotropic polypeptide (GIP) gene. *Comp. Biochem. Phys. D. Genomics Proteomics* 1:385-395.

54. Kieffer TJ, Habener JF (1999) The glucagon-like peptides. *Endocr. Rev.* 20:876-913.

55. Nauck M, Stöckmann F, Ebert R, Creutzfeldt W (1986) Reduced incretin effect in type 2 (non-insulin-dependent) diabetes. *Diabetologia* 29:46-52.

56. Vilsbøll T, Krarup T, Deacon CF, Madsbad S, Holst JJ (2001) Reduced postprandial concentrations of intact biologically active glucagon-like peptide 1 in type 2 diabetic patients. *Diabetes* 50:609-613.

57. Meier J, Nauck M, Schmidt W, Gallwitz B (2002) Gastric inhibitory polypeptide: the neglected incretin revisited. *Regul. Pept.* 107:1–13.

58. Mudaliar S, Henry RR (2009) Incretin therapies: effects beyond glycemic control. *Eur. J. Intern. Med.* 20:S319-S328.

59. Kim W, Egan JM (2008) The Role of Incretins in Glucose Homeostasis and Diabetes Treatment. *Pharmacol. Rev.* 60:470 -512.

60. Farilla L et al. (2003) Glucagon-like peptide 1 inhibits cell apoptosis and improves glucose responsiveness of freshly isolated human islets. *Endocrinology* 144:5149-5158.

61. De León DD et al. (2003) Role of endogenous glucagon-like peptide-1 in islet regeneration after partial pancreatectomy. *Diabetes* 52:365-371.

62. Tourrel C, Bailbe D, Meile M, Kergoat M, Portha B (2001) Glucagon-Like Peptide-1 and Exendin-4 stimulate β-cell neogenesis in streptozotocin-treated newborn rats resulting in persistently improved glucose homeostasis at adult age. *Diabetes* 50:1562-1570.

63. Holz G, Kühtreiber W, Habener J (1993) Pancreatic beta-cells are rendered glucose-competent by the insulinotropic hormone glucagon-like peptide-1(7-37). *Nature* 361:362-365.

64. Weir G, Bonner-weir S (2004) Five Stages of Evolving β-Cell Dysfunction During Progression to Diabetes. *Diabetes* 53:S16-S21.

65. Wang Y et al. (1996) GIP regulates glucose transporters, hexokinases, and glucose-induced insulin secretion in RIN 1046-38 cells. *Mol. Cell. Endocrinol.* 116:81-87.

66. Hinke SA et al. (2003) Glucose-dependent insulinotropic polypeptide receptor null mice exhibit compensatory changes in the enteroinsular axis. *Am. J. Physiol. Endocrinol. Metab.* 284:E931-939.

67. Roberge J, Brubaker P (1993) Regulation of intestinal proglucagon-derived peptide secretion by glucose-dependent insulinotropic peptide in a novel enteroendocrine loop. *Endocrinology* 133:233-240.

68. Herrmann-Rinke C, Vöge A, Hess M, Göke B (1995) Regulation of glucagon-like peptide-1 secretion from rat ileum by neurotransmitters and peptides. *J. Endocrinol.* 147:25-31.

69. Damholt AB, Buchan AM, Kofod H (1998) Glucagon-like-peptide-1 secretion from canine L-cells is increased by glucose-dependent-insulinotropic peptide but unaffected by glucose. *Endocrinology* 139:2085-2091.

70. Kim S et al. (2005) Glucose-dependent insulinotropic polypeptide (GIP) stimulation of pancreatic beta-cell survival is dependent upon phosphatidylinositol 3-kinase (PI3K)/protein kinase B (PKB) signaling, inactivation of the forkhead transcription factor Foxo1, and down-regu. *J. Biol. Chem.* 280:22297-22307.

71. Trümper A, Trümper K, Hörsch D (2002) Mechanisms of mitogenic and anti-apoptotic signaling by glucose-dependent insulinotropic polypeptide in β (INS-1)-cells. *J. Endocrinol.* 174:233-246.

72. Trümper A et al. (2001) Glucose-Dependent Insulinotropic Polypeptide Is a Growth Factor for (INS-1) Cells by Pleiotropic Signaling. *Mol. Endocrinol.* 15:1559-1570.

73. Maida A, Hansotia T, Longuet C, Seino Y, Drucker DJ (2009) Differential Importance of GIP Versus GLP-1 Receptor Signaling for Beta Cell Survival in Mice. *Gastroenterology*:Epub ahead of print [10.1053/j.gastro.2009.09.004].

74. Irwin N, Flatt PR (2009) Therapeutic potential for GIP receptor agonists and antagonists. *Best Pract. Res. Clin. Endocrinol. Metab.* 23:499-512.

75. Meier JJ et al. (2001) Reduced insulinotropic effect of gastric inhibitory polypeptide in first-degree relatives of patients with type 2 diabetes. *Diabetes* 50:2497-2504.

76. Yip RG, Wolfe M (2000) Minireview GIP biology and fat metabolism. *Science* 66:91-103.

77. Rodbard HW et al. (2009) Statement by an American Association of Clinical Endocrinologists/American College of Endocrinology Consensus Panel on Type 2 Diabetes Mellitus: An Algorithm for Glycemic Control. *Endocr. Pract.* 15:541-559.

78. Yip RG, Boylan MO, Kieffer TJ, Wolfe MM (1998) Functional GIP Receptors are Present on Adipocytes. *Endocrinology* 139:4004-4007.

79. Miyawaki K et al. (2002) Inhibition of gastric inhibitory polypeptide signaling prevents obesity. *Nat. Med.* 8:738-742.

80. Althage M et al. (2008) Targeted ablation of glucose-dependent insulinotropic polypeptide-producing cells in transgenic mice reduces obesity and insulin resistance induced by a high fat diet. *J. Biol. Chem.* 283:18365-18376.

81. McClean PL et al. (2009) GIP receptor antagonism reverses obesity, insulin resistance, and associated metabolic disturbances induced in mice by prolonged consumption of high-fat diet. *Am. J. Physiol. Endocrinol. Metab.* 293:E1746-E1755.

82. Gault V et al. (2005) Chemical Ablation of Gastric Inhibitory Polypeptide Receptor Action by Daily (Pro3)GIP Administration Improves Glucose Tolerance and Ameliorates Insulin Resistance and Abnormalities of Islet Structure in Obesity-Related Diabetes. *Diabetes* 54:2436-2446.

83. Parker JC et al. (2007) Metabolic effects of sub-chronic ablation of the incretin receptors by daily administration of (Pro3)GIP and exendin(9-39)amide in obese diabetic (ob/ob) mice. *Biol. Chem.* 388:221-226.

84. Orban T et al. (2009) Pancreatic Islet Autoantibodies as Predictors of Type 1 Diabetes in the Diabetes Prevention Trial–Type 1. *Diabetes Care* 32:2269-2274.

85. Wright EE (2009) Overview of insulin replacement therapy. *J. Fam. Pract.* 58:S3-9.

86. Bliss M (2000) *The Discovery of Insulin.* (University of Toronto Press, Toronto)3rd editio.

87. Hanaire H et al. (2008) Treatment of diabetes mellitus using an external insulin pump: the state of the art. *Diabetes Metab.* 34:401-23.

88. Hartman I (2008) Insulin analogs: impact on treatment success, satisfaction, quality of life, and adherence. *Clin. Med. Res.* 6:54-67.

89. Kelly W, Lillehei R, Merkel F, Idezuki Y, Goetz F (1967) Allotransplantation of the pancreas and duodenum along with the kidney in diabetic nephropathy. *Surgery* 61:827-837.

90. Kemp C, Knight M, Scharp D, Ballinger W, Lacy P (1973) Effect of transplantation site on the results of pancreatic islet isografts in diabetic rats. *Diabetologia* 9:486-491.

91. Ballinger W, Lacy P, Scharp D, Kemp C, Knight M (1973) Isografts and allografts of pancreatic islets in rats. *Br. J. Surg.* 60:313.

92. Ballinger W, Lacy P (1972) Transplantation of intact pancreatic islets in rats. *Surgery* 72:175-186.

93. Scharp DW et al. (1990) Insulin independence after islet transplantation into type I diabetic patient. *Diabetes* 39:515-518.

94. Shapiro A et al. (2000) Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosupressive regimen. *New Engl. J. Med.* 343:230-238.

95. Ryan E et al. (2005) Five-year follow-up after clinical islet transplantation. *Diabetes* 54:2060-2069.

96. Harlan DM, Kenyon NS, Korsgren O, Roep BO (2009) Current advances and travails in islet transplantation. *Diabetes* 58:2175-2184.

97. Paveglio SA, Olson DE, Huey PU, Porter MH, Thulé PM (2003) Glucose-responsive hepatic insulin gene therapy of spontaneously diabetic BB/Wor rats. *Hum. Gene Ther.* 14:1401-1413.

98. Chen NK, Sivalingam J, Tan SY, Kon OL (2005) Plasmid-electroporated primary hepatocytes acquire quasi-physiological secretion of human insulin and restore euglycemia in diabetic mice. *Gene Ther.* 12:655-667.

99. Yasutomi K et al. (2003) Intravascular insulin gene delivery as potential therapeutic intervention in diabetes mellitus. *Biochem. Biophys. Res. Commun.* 310:897-903.

100. Alam T, Sollinger H (2002) Glucose-regulated insulin production in hepatocytes. *Transplantation* 74:1781-1787.

101. Dong H, Woo SL (2001) Hepatic insulin production for type 1 diabetes. *Trends Endocrinol. Metab.* 12:441-446.

102. Choi KS et al. (2005) In vitro trans-differentiation of rat mesenchymal cells into insulin-producing cells by rat pancreatic extract. *Biochem. Biophys. Res. Commun.* 330:1299-1305.

103. Yang L et al. (2002) In vitro trans-differentiation of adult hepatic stem cells into pancreatic endocrine hormone-producing cells. *Proc. Natl. Acad. Sci. U. S. A.* 99:8078-8083.

104. Tang D et al. (2004) In Vivo and In Vitro Characterization of Insulin-Producing Cells Obtained From Murine Bone Marrow. *Diabetes* 53:1721-1732.

105. Oh S et al. (2004) Adult bone marrow-derived cells trans-differentiating into insulin-producing cells for the treatment of type I diabetes. *Lab. Invest.* 84:607-617.

106. Kleinjan DA, van Heyningen V (2005) Long-Range Control of Gene Expression: Emerging Mechanisms and Disruption in Disease. *Am. J. Hum. Genet.* 76:8-32.

107. Smale ST, Kadonaga JT (2003) The RNA polymerase II core promoter. *Annu. Rev. Biochem.* 72:449-479.

108. Gilchrist DA, Fargo DC, Adelman K (2009) Using ChIP-chip and ChIP-seq to study the regulation of gene expression: genome-wide localization studies reveal widespread regulation of transcription elongation. *Methods* 48:398-408.

109. Hager GL, McNally JG, Misteli T (2009) Transcription dynamics. *Mol. Cell* 35:741-753.

110. Cheng H, Leblon (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. *Am. J. Anat.* 141:537-561.

111. Cheng H (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. II. Mucous cells. *Am. J. Anat.* 141:481-501.

112. Cheng H, Leblond C (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. III. Entero-endocrine cells. *Am. J. Anat.* 141:503-519.

113. Cheng H (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. IV. Paneth cells. *Am. J. Anat.* 141:521-535.

114. Cheng H, Leblond C (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. I. Columnar cell. *Am. J. Anat.* 141:461-479.

115. Lee CS, Kaestner KH (2004) Development of gut endocrine cells. *Best Pract. Res. Clin. Endocrinol. Metab.* 18:453-462.

116. Yang Q, Bermingham N, Finegold M, Zoghbi H (2001) Requirement of Math1 for Secretory Cell Lineage Commitment in the Mouse Intestine . *Science* 294:2155-2158.

117. Lee CS, Perreault N, Brestelli JE, Kaestner KH (2002) Neurogenin 3 is essential for the proper specification of gastric enteroendocrine cells and the maintenance of gastric epithelial cell identity. *Genes Dev.* 16:1488-1497.

118. Gradwohl G, Dierich A, Lemeur M, Guillemot F (2000) neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc. Natl. Acad. Sci. U. S. A.* 97:1607-1611.

119. Jenny M et al. (2002) Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO J.* 21:6338-6347.

120. Ishibashi M et al. (1995) Targeted disruption of mammalian hairy and Enhancer of split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. *Genes Dev.* 9:3136-3148.

121. Jensen J et al. (2000) Control of endodermal endocrine development by Hes-1. *Nat. Genet.* 24:36-44.

122. Mutoh H et al. (1997) The basic helix-loop-helix transcription factor BETA2/NeuroD is expressed in mammalian enteroendocrine cells and activates secretin gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 94:3560-3564.

123. Naya FJ et al. (1997) Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice. *Genes Dev.* 11:2323-2334.

124. Larsson LI, St-Onge L, Hougaard DM, Sosa-Pineda B, Gruss P (1998) Pax 4 and 6 regulate gastrointestinal endocrine cell development. *Mech. Dev.* 79:153-159.

125. Larsson L, Tingstedt J, Madsen O, Serup P, Hougaard D (1995) The LIM-homeodomain protein Isl-1 segregates with somatostatin but not with gastrin expression during differentiation of somatostatin/gastrin precursor cells. *Endocrine* 3:519-524.

126. Larsson LI, Madsen OD, Serup P, Jonsson J, Edlund H (1996) Pancreatic-duodenal homeobox 1 -role in gastric endocrine patterning. *Mech. Dev.* 60:175-184.

127. Offield M et al. (1996) PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development* 122:983-995.

128. Heller RS et al. (2004) The role of Brn4/Pou3f4 and Pax6 in forming the pancreatic glucagon cell identity. *Dev. Biol.* 268:123-134.

129. Gosmain Y, Avril I, Mamin A, Philippe J (2007) Pax-6 and c-Maf functionally interact with the alpha-cell-specific DNA element G1 in vivo to promote glucagon gene expression. *J. Biol. Chem.* 282:35024-35034.

130. Ding J et al. (2009) Pax6 haploinsufficiency causes abnormal metabolic homeostasis by down-regulating glucagon-like peptide 1 in mice. *Endocrinology* 150:2136-2144.

131. Hill M, Asa S, Drucker D (1999) Essential Requirement for Pax6 in Control of Enteroendocrine Proglucagon Gene Transcription. *Mol. Endocrinol.* 13:1474–1486.

132. Fujita Y et al. (2008) Pax6 and Pdx1 are required for production of glucose-dependent insulinotropic polypeptide in proglucagon-expressing L cells. *Am. J. Physiol. Endocrinol. Metab.* 295:E648-E657.

133. Trinh DK, Zhang K, Hossain M, Brubaker PL, Drucker DJ (2003) Pax-6 activates endogenous proglucagon gene expression in the rodent gastrointestinal epithelium. *Diabetes* 52:425-433.

134. Jonsson J, Garlsson L, Edlund T, Edlund H (1994) Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature* 371:606-609.

135. Jepeal LI et al. (2005) Cell-specific expression of glucose-dependent-insulinotropic polypeptide is regulated by the transcription factor PDX-1. *Endocrinology* 146:383-391.

136. Boyer DF et al. (2006) Complementation rescue of Pdx1 null phenotype demonstrates distinct roles of proximal and distal cis-regulatory sequences in pancreatic and duodenal expression. *Dev. Biol.* 298:616-631.

137. Marshak S et al. (2000) Functional Conservation of Regulatory Elements in the pdx-1 Gene: PDX-1 and Hepatocyte Nuclear Factor 3β Transcription Factors Mediate β-Cell-Specific Expression. *Mol. Cell Biol.* 20:7583-7590.

138. Gerrish K, Van Velkinburgh JC, Stein R (2004) Conserved transcriptional regulatory domains of the pdx-1 gene. *Mol. Endocrinol.* 18:533-548.

139. Gerrish K et al. (2000) Pancreatic β Cell-specific Transcription of the pdx-1 Gene. *J. Biol. Chem.* 275:3485-3492.

140. Yamada S et al. (2001) Differentiation of immature enterocytes into enteroendocrine cells by Pdx1 overexpression. *Am. J. Physiol. Gastrointest. Liver Physiol.* 281:G229-G236.

141. Chen C, Fang R, Davis C, Maravelias C, Sibley E (2009) Pdx1 Inactivation Restricted to the Intestinal Epithelium in Mice Alters Duodenal Gene Expression in Enterocytes and Enteroendocrine Cells. *Am. J. Physiol. Gastrointest. Liver Physiol.*:Epub ahead of print [10.1152/ajpgi.90586.2008].

142. Someya Y, Seino Y, Inagaki N, Ishii S, Maekawa T (1993) Two 3',5'-cyclic-adenosine monophosphate response elements in the promoter region of the human gastric inhibitory polypeptide gene. *FEBS Lett.* 317:67-73.

143. Boylan M, Jarboe L, Jepeal L, Wolfe M (1997) Cell-specific expression of the glucose-dependent insulinotropic polypeptide gene in a mouse neuroendocrine tumor cell line. *J. Biol. Chem* 272:17438-17443.

144. Jepeal L, Boylan M, Wolfe M (2003) Cell-specific expression of the glucose-dependent insulinotropic polypeptide gene functions through a GATA and an ISL-1 motif in a mouse neuroendocrine tumor cell line. *Regul. Pept.* 113:139-147.

145. Jepeal LI, Boylan MO, Michael Wolfe M (2008) GATA-4 upregulates glucose-dependent insulinotropic polypeptide expression in cells of pancreatic and intestinal lineage. *Mol. Cell. Endocrinol.* 287:20-29.

146. Yeung C, Wong CK, Chung SK, Chung SS, Chow BK (1999) Glucose-dependent insulinotropic polypeptide gene expression in the stomach: revealed by a transgenic mouse study, in situ hybridization and immunohistochemical staining. *Mol. Cell. Endocrinol.* 154:161 - 170.

147. Rindi G et al. (1990) Development of neuroendocrine tumors in the gastrointestinal tract of transgenic mice. Heterogeneity of hormone expression. *Am. J. Pathol.* 136:1349–1363.

148. Gajic D, Drucker D (1993) Multiple cis-acting domains mediate basal and adenosine 3',5'-monophosphate-dependent glucagon gene transcription in a mouse neuroendocrine cell line. *Endocrinology* 132:1055-1062.

149. Goldhamer DJ, Faerman A, Shani M, Emerson CP (1992) Regulatory Elements That Control the Lineage-Specific Expression of myoD . *Science* 256:538-542.

150. Waschek J, Hsu C, Eiden L (1988) Lineage-specific regulation of the vasoactive intestinal peptide gene in neuroblastoma cells is conferred by 5.2 kilobases of 5'-flanking sequence. *Proc. Natl. Acad. Sci. U. S. A.* 85:9547-9551.

151. Agoston D, Bravo D, Waschek J (1990) Expression of a chimeric VIP gene is targeted to the intestine in transgenic mice. *J. Neurosci. Res.* 27:479-486.

152. Portela-Gomes GM, Johansson H, Olding L, Grimelius L (1999) Co-localization of neuroendocrine hormones in the human fetal pancreas. *Eur. J. Endocrinol.* 141:526-533.

153. Schwartz S et al. (2000) PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res.* 10:577-586.

154. Schwartz S et al. (2003) MultiPipMaker and supporting tools: alignments and analysis of multiple genomic DNA sequences. *Nucleic Acids Res.* 31:3518-3524.

155. Farre D et al. (2003) Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res.* 31:3651-3653.

156. Messeguer X et al. (2002) PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* 18:333-334.

157. Schreiber E, Matthias P, Muller M, Schaffner W (1989) Rapid detection of octamer binding proteins with'mini extracts', prepared from a small number of cells. *Nucleic Acids Res.* 17:6419.

158. Weiser M (1973) Intestinal Epithelial Cell Surface Membrane Glycoprotein Synthesis. *J. Biol. Chem* 248:2536-2541.

159. Ferraris R, Villenas S, Diamond J (1992) Regulation of brush-border enzyme activities and enterocyte migration rates in mouse small intestine. *Am. J. Physiol. Gastrointest. Liver Physiol.* 262:1047-1059.

160. Dekaney CM, Rodriguez JM, Graul MC, Henning SJ (2005) Isolation and characterization of a putative intestinal stem cell fraction from mouse jejunum. *Gastroenterology* 129:1567-1580.

161. Reimann F et al. (2008) Glucose sensing in L cells: a primary cell study. *Cell Metab.* 8:532-539.

162. Antequera F (2003) Structure, function and evolution of CpG island promoters. *Cell. Mol. Life Sci.* 60:1647-1658.

163. Meech R, Kallunki P, Edelman GM, Jones FS (1999) A binding site for homeodomain and Pax proteins is necessary for L1 cell adhesion molecule gene expression by Pax-6 and bone morphogenetic proteins. *Proc. Natl. Acad. Sci. U. S. A.* 96:2420-2425.

164. Holst BD, Wang Y, Jones FS, Edelman GM (1997) A binding site for Pax proteins regulates expression of the gene for the neural cell adhesion molecule in the embryonic spinal cord. *Proc. Natl. Acad. Sci. U. S. A.* 94:1465-1470.

165. Promega (2009) pGL4 Luciferase Reporter Vectors Technical Manual.

166. Kieffer T et al. (1995) Gastric inhibitory polypeptide a tumor-derived cell line release from a tumor-derived cell line. *Am. J. Physiol. Endocrinol. Metab.* 32:E316-E322.

167. Andersen FG et al. (1999) Pax6 and Pdx1 form a functional complex on the rat somatostatin gene upstream enhancer. *FEBS Lett.* 445:315-320.

168. Heeg S, Doebele M (2006) In Vitro Transformation Models. *Cell Cycle* 5:630-634.

169. Doniger SW, Fay JC (2007) Frequent gain and loss of functional transcription factor binding sites. *PLoS Comput. Biol.* 3:932-942.

170. Buchan A, Ingman-Baker J, Levy J, Brown J (1982) A comparison of the ability of serum and monoclonal antibodies to gastric inhibitory polypeptide to detect immunoreactive cells in the gastroenteropancreatic system of mammals and reptiles. *Histochemistry* 76:341-349.

171. Verkhusha V et al. (2003) Accelerated Publications High Stability of Discosoma DsRed As Compared to Aequorea EGFP. *Biochemistry* 42:7879-7884.

172. Baird GS, Zacharias D, Tsien RY (2000) Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. U. S. A.* 97:11984-11989.

173. Bevis BJ, Glick BS (2002) Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed). *Nature Biotechnol.* 20:83-87.

174. Anlauf M, Weihe E, Hartschuh W, Hamscher G, Feurle GE (2000) Localization of xenin-immunoreactive cells in the duodenal mucosa of humans and various mammals. *J. Histochem. Cytochem.* 48:1617-1626.

175. Natori S et al. (1998) Chromogranin B (secretogranin I), a neuroendocrine- regulated secretory protein, is sorted to exocrine secretory granules in transgenic mice. *EMBO J.* 17:3277-3289.

176. Hendy G, Bevan S, Mattei M, Mouland A (1995) Chromogranin A. *Clin. Invest. Med.* 18:47-65.

177. Barker N et al. (2007) Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* 449:1003-1007.

178. Parnis S, Nicoletti C, Ollendorff V, Massey-Harroche D (2004) Enterocytin: A new specific enterocyte marker bearing a B30.2-like domain. *J. Cell. Physiol.* 198:441-451.

179. Kedinger M, Haffen K, Simon-Assmann P (1987) Intestinal tissue and cell cultures. *Differentiation* 36:71-85.

180. Evans G, Flint N, Somers A, Eyden B, Potten C (1992) The development of a method for the preparation of rat intestinal epithelial cell primary cultures. *J. Cell Sci.* 101:219-231.

181. Booth C, Patel S, Bennion G, Potten C (1995) The isolation and culture of adult mouse colonic epithelium. *Epithelial Cell Biol.* 4:76-86.

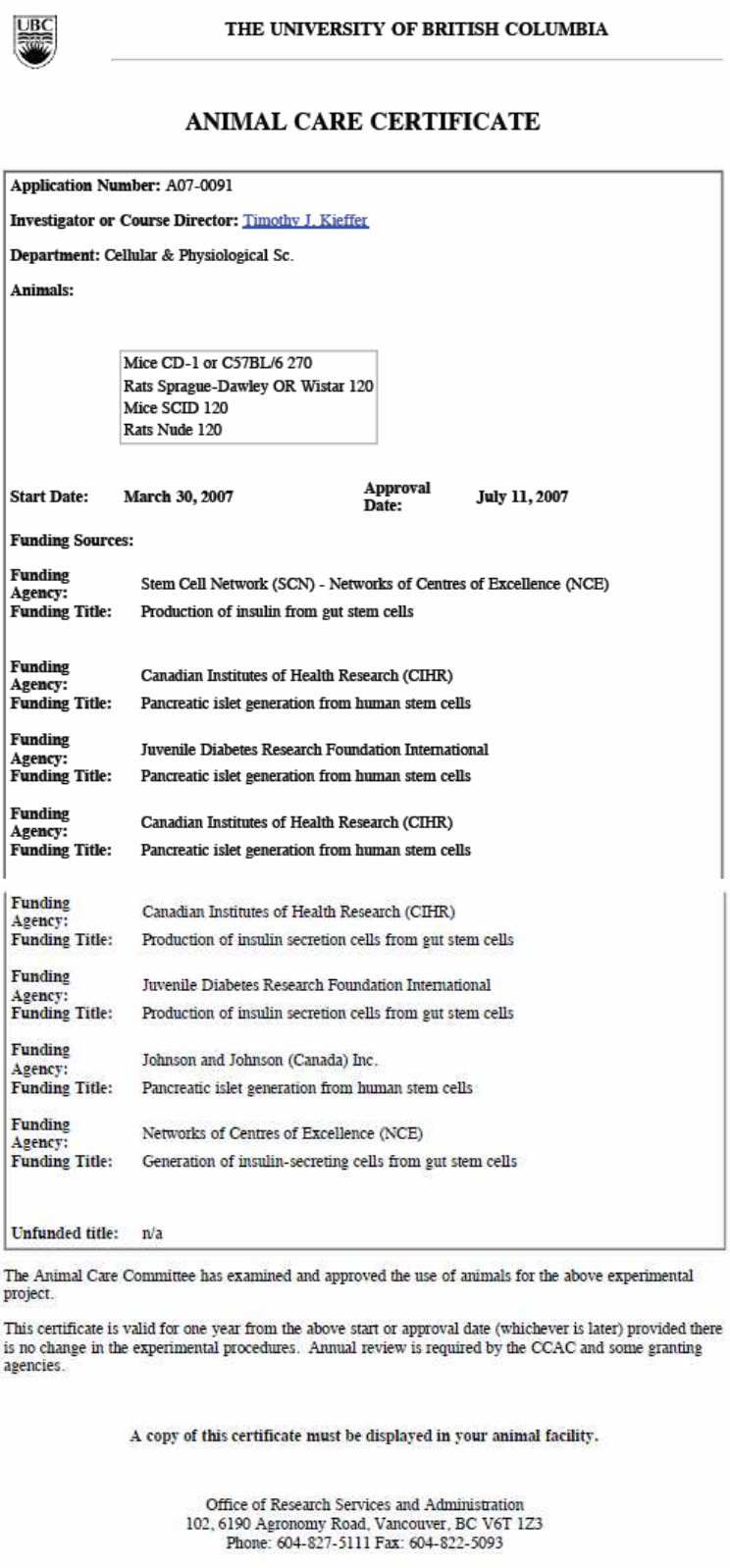
182. Kaeffer B (2002) Mammalian intestinal epithelial cells in primary culture: a mini-review. *In Vitro Cell. Dev. Biol. Anim.* 38:123-134.

183. Zhou L, Nian M, Gu J, Irwin DM (2006) Intron 1 sequences are required for pancreatic expression of the human proglucagon gene. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 290:R634-R641.

184. Grant S et al. (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* 38:320-323.

APPENDIX

UBC Research Ethics Board Certificates of Approval



THE UNIVERSITY OF BRITISH COLUMBIA

ANIMAL CARE CERTIFICATE

Application Number: A07-0091

Investigator or Course Director: Timothy J. Kieffer

Department: Cellular & Physiological Sc.

Animals:

Mice CD-1 or C57BL/6 270
Rats Sprague-Dawley OR Wistar 120
Mice SCID 120
Rats Nude 120

Start Date: March 30, 2007 **Approval Date:** July 11, 2007

Funding Sources:

Funding Agency: Stem Cell Network (SCN) - Networks of Centres of Excellence (NCE)
Funding Title: Production of insulin from gut stem cells

Funding Agency: Canadian Institutes of Health Research (CIHR)
Funding Title: Pancreatic islet generation from human stem cells

Funding Agency: Juvenile Diabetes Research Foundation International
Funding Title: Pancreatic islet generation from human stem cells

Funding Agency: Canadian Institutes of Health Research (CIHR)
Funding Title: Pancreatic islet generation from human stem cells

Funding Agency: Canadian Institutes of Health Research (CIHR)
Funding Title: Production of insulin secretion cells from gut stem cells

Funding Agency: Juvenile Diabetes Research Foundation International
Funding Title: Production of insulin secretion cells from gut stem cells

Funding Agency: Johnson and Johnson (Canada) Inc.
Funding Title: Pancreatic islet generation from human stem cells

Funding Agency: Networks of Centres of Excellence (NCE)
Funding Title: Generation of insulin-secreting cells from gut stem cells

Unfunded title: n/a

The Animal Care Committee has examined and approved the use of animals for the above experimental project.

This certificate is valid for one year from the above start or approval date (whichever is later) provided there is no change in the experimental procedures. Annual review is required by the CCAC and some granting agencies.

A copy of this certificate must be displayed in your animal facility.

Office of Research Services and Administration
102, 6190 Agronomy Road, Vancouver, BC V6T 1Z3
Phone: 604-827-5111 Fax: 604-822-5093